

2025

2024

2023

2022

2020

2019

2018

2017

2016

2015
20132014
2012

2023

hEART 2023: 11th Symposium of the European Association for Research in Transportation, September 6-8, 2023

Venue: Eidgenössische Technische Hochschule Zürich

Organization: Prof. Kay Axhausen

Anders Fjendbo Jensen and Jeppe Rich	Empirical analysis of cycling trends in two of Europe's most bicycle-friendly regions: Identifying the successes and the setbacks	PDF
Sander Van Cranenburgh and Francisco Garrido Valenzuela	Using computer vision-enriched discrete choice models to assess the visual impact of transport infrastructure renewal projects: A case study of the Delft railway zone	PDF
Stephane Hess and Sander Van Cranenburgh	Combine and conquer: model averaging for out-of-distribution forecasting	PDF
Sebastian Hörl and Puchinger Jakob	Modeling the ecological and economic footprint of last-mile parcel deliveries using open data: A case study for Lyon	PDF
Alessio Daniele Marra and Francesco Corman	Evaluating real-time information systems on public transport disturbances	PDF
Rowan Hoogervorst, Evelien van der Hurk, Philine Schiewe, Anita Schöbel and Reena Urban	The Bus Rapid Transit Investment Problem	PDF
Patrick Stokkink, André de Palma and Nikolas Geroliminis	Carpooling with Transfers and Travel Time Uncertainty	PDF
Tom Haering and Michel Bierlaire	A Spatial Branch and Bound Algorithm for Continuous Pricing with Advanced Discrete Choice Demand Modeling	PDF
Ruben A. Kuipers and Michelle Ochsner	The Impact of Weather Phenomena on Passenger Volumes for Commuter Trains	PDF
Takara Sakai, Takashi Akamatsu and Koki Satsukawa	Welfare impacts of remote and flexible working policies in the bottleneck model	PDF

Arkadiusz Drabicki, Oded Cats and Rafał Kucharski	urban public transport – before vs. after COVID-19 pandemic	PDF
Thomas Hancock, Charisma Choudhury, Joan Walker and Stephane Hess	Quantum choice models leap out of the laboratory: capturing real-world behavioural change.	PDF
Konstantin Krauss	Shifting to sharing: Are external costs reduced or merely redistributed?	
Bastián Henríquez-Jara, C. Angelo Guevara and Angel Jimenez-Molina	Identifying instant utility (latent emotion) triggers using psychophysiological indicators with an Experience-Based Choice Model in a travel experiment	PDF
Joris Wagenaar, Evelien van der Hurk, Marie Schmidt and Evelien van der Hurk	A model for Robust Rolling Stock Scheduling	PDF
Rui Yao and Renming Liu	Can Bayesian Optimization be the Last Puzzle for Automatic Estimation of Neural Network Discrete Choice Models? An experiment	PDF
Gabriel Nova and C. Angelo Guevara	In-depth, Breath-first or Both? Toward the Development of a RUM-DFT Discrete Choice Model	PDF
Carlos Gaete-Morales, Julius Jöhrens, Florian Heining and Wolf-Peter Schill	Power sector effects of alternative options for electrifying heavy-duty vehicles	PDF
Leon Weinsziehr, Frederik Bachmann, Antonios Tsakarestos and Klaus Bogenberger	Detection of Bus Bunching through the Analysis of Prevalent Public Transport Control Data	PDF
Rasha Bowirrat, Karel Martens and Yoram Shiftan	Explaining Walking in Cities – a Machine Learning Approach	PDF
Mads Paulsen and Jeppe Rich	Optimal bicycle network expansions with endogenous demand	PDF
Felix Hofinger and Martin Fellendorf	Lane change behavior on motorways based on naturalistic trajectory data	PDF
Hannes Wallimann, Kevin Blättler and Widar von Arx	Do price reductions attract customers in urban public transport? A synthetic control approach	PDF
Muhamad Rizki, Tri Basuki Joewono and Yusak Susilo	Exploring the Effect of Apps Evolution and Users' Personality on Mobile Apps Adoption and Post-Adoption Pattern Over Time: Evidence from Super-Apps Users in Indonesian Cities	PDF
Florian Fuchs, Viera Klasovitá and Francesco Corman	Routing Passengers while Timetabling Based on Promises from Line Planning: A Logic-Based Benders Approach	PDF
Anupriya Anupriya, Daniel Graham and Prateek Bansal	Quantification of non-linear effects in agglomeration economies for transport appraisals	PDF
Georgios Kapousizis, Rumana Sarker, Baran Ulak and Karst Geurs	Acceptance of new technologies affecting safety on electric bicycles: evidence from five European countries	PDF
Gunnar Flötteröd	Improved precision in a heuristic for particle-based and stochastic dynamic traffic assignment	PDF
Magdalena Schilling, Marvin V. Baumann, Jörg Sonnleitner, Markus Friedrich and Peter Vortisch	Design hourly volume estimation at freeway nodes using floating car data	PDF

Claudia Bandiera, Richard Connors and Francesco Viti	Mobility Service Providers' Equilibrium Strategies in Multi-modal Networks	PDF
Thomas Schatzmann, Felix Zwick and Kay Axhausen	Investigating the preferences for the use of urban ridepooling	PDF
Ilka Dubernet and Dennis Seibert	Investigating preferences for powertrains when buying a car in Germany	PDF
Dimitrios Pappelis, Emmanouil Chaniotakis, Tim Hillel and Maria Kamargianni	Modelling Travel Time Anticipation Under Rational Inattention and Endogenous Information Constraints	PDF
Daniel Hörcher and Daniel Graham	The spatial variation of travel time valuations: A general equilibrium model and application in project appraisal	PDF
Ying-Chuan Ni, Michail Makridis and Anastasios Kouvelas	Investigating Link- and Network-level Bicycle Traffic Flow Characteristics using a Microsimulation Approach	PDF
Farnoud Ghasemi, Arkadiusz Drabicki and Rafał Kucharski	Dynamics of the Ride-Sourcing Market: A Coevolutionary Model of Competition between Two-Sided Mobility Platforms	PDF
Louis Balzer and Ludovic Leclercq	Cooperation between Ride-Hailing and Public Transportation with Tradable Credit Schemes	PDF
Manon Seppecher and Ludovic Leclercq	An auctioning process for large-scale ride-hailing vehicles repositioning	PDF
Menno Yap, Howard Wong and Oded Cats	Public Transport Crowding Valuation in a Post-Pandemic Era	PDF
Jingyi Cheng and Shadi Sharif Azadeh	A data-driven dynamic demand hotspots forecasting framework for on-demand meal delivery platforms	PDF
Shadi Haj Yahia, Omar Mansour and Tomer Toledo	Incorporating Domain Knowledge in Deep Neural Networks for Mode Choice Analysis	PDF
David Kohlrantz and Tobias Kuhnimhof	Modeling the Demand for Bicycle Parking Facilities	PDF
Sara Momen, Bart van Arem and Shadi Sharif Azadeh	Dynamic location for charging operations of shared free-floating e-scooters	PDF
Mingye Luan, S.Travis Waller and David Rey	A non-additive path-based reward credit scheme for traffic congestion management	PDF
Joanna Ji, Qin Zhang, Ana Tsui Moreno and Rolf Moeckel	The impact of social networks and coordinated destination choice on the spread of epidemics using Episim	PDF
Laurent Cazor, Mirosława Łukawska, Mads Paulsen, Thomas Rasmussen and Otto Nielsen	Whose preferences matter more? Handling unbalanced panel data for choice modelling	PDF
Christina Iliopoulou, Michail Makridis and Anastasios Kouvelas	Resilience-Oriented Design for Public Transport Networks	PDF
Filippos Adamidis, Sara Moghavam Ghaffari and Constantinos Antoniou	Acceptance of car-reducing measures: observed factors and latent attitudes	PDF
Han Zhou, Yashar Araghi, Bachtijar Ashari and Maaïke Snelder	An activity-based latent class modelling approach to assess the impact of hybrid working on travel demand in the Netherlands after COVID-19	PDF
Léa Ricard, Guy Desaulniers, Andrea	The stochastic multiple depot electric vehicle scheduling	

Lodi and Louis-Martin Rousseau	problem with recourse	PDF
Inneke Van Hoeck and Pieter Vansteenwegen	A heuristic approach to improve the robustness of a railway timetable in a bottleneck area	PDF
Niousha Bagheri Khoulenjani, Milad Ghasri and Michael Barlow	Post-hoc explanation methods for deep neural networks in choice analysis	PDF
Vishal Mahajan, Guido Cantelmo and Constantinos Antoniou	An open-source framework for the robust calibration of large-scale traffic simulation models	PDF
Benoit Matet, Etienne Côme, Angelo Furno, Sebastian Hörl and Latifa Oukhellou	Use of Origin-Destination data for calibration and spatialization of synthetic travel demand	PDF
Federico Bigi, Nicola Schwemmler and Francesco Viti	Evaluating the impact of Free Public Transport using agent-based modeling: the case-study of Luxembourg	PDF
Gülin Göksu Başaran, Jesper Bláfoss Ingvardson and Otto Anker Nielsen	Influence of station characteristics, urban surroundings and perceived safety on satisfaction and public transport ridership	PDF
Rong Cheng, Andreas Fessler, Allan Larsen, Otto Anker Nielsen and Yu Jiang	Assessing the Impacts of Public Transport-Based Crowdshipping: A Case Study in Nørrebro District in Copenhagen	PDF
Hassan Idoudi, Mostafa Ameli, Cyril Nguyen Van Phu, Mahdi Zargayouna and Abderrezak Rachedi	Enhancing Evacuation Planning and Management through Vehicular Communication	PDF
Liang Ma, Daniel J. Graham and Marc E.J. Stettler	Using Explainable Machine Learning to Interpret the Effects of Policies on Air Pollution: COVID-19 Lockdown in London	PDF
Nicola Ortelli, Matthieu de Lapparent and Michel Bierlaire	Faster estimation of discrete choice models via weighted dataset reduction	PDF
Lubing Li, Ka Fai Ng, Jacob Lo and Hong Lo	Adaptive Traffic Signal Control: A Novel Modelling Approach	PDF
Fábio Hipólito, Jeppe Rich and Peter Bach Andersen	Charging demand for the unserved — an agent-based model approach	PDF
Yiru Jiao, Simeon Calvert, Sander van Cranenburgh and Hans van Lint	Varying critical time to collision: a perspective of driver space	PDF
Xiaowei Zhu, Anupriya Anupriya and Daniel Graham	Understanding the cycle traffic impacts of Cycle Superhighways in London	PDF
Khashayar Khavarian, Shaghayegh Vosough and Claudio Roncoli	How do electric bikes affect the route choice of cyclists? A case study of Greater Helsinki	PDF
Léon Sobrie, Marijn Verschelde and Bart Roets	Explainable predictions for real-time employee workload management in railway control rooms	PDF
Camila Balbontin, John Nelson, David Hensher and Matthew Beck	Identifying main drivers for students and staff members' mode choice or to work/study from home: A case study in Australia	PDF
Michał Bujak and Rafał Kucharski	Assessing expected ride-pooling performance with non-deterministic, heterogeneous travellers' behaviour.	PDF
Gabriel Hannon, Joanna Ji, Qin Zhang,	Implementing an Agent-Based Formation of Social	PDF

Ana Tsui Moreno and Rolf Moeckel	Networks for Joint Travel	
Emily Morey, R. Eddie Wilson and Kevin Galvin	Simulation of Mixtures of Legacy and Autonomous Mainline Rail Operations	PDF
Kenan Zhang, Andres Fielbaum and Javier Alonso-Mora	What do walking and e-hailing bring to scale economies? A general microeconomic model for on-demand mobility	PDF
Giancarlo Parady, Yuki Oyama and Makoto Chikaraishi	Text-aided Group Decision-making Process Observation Method (x-GDP): A novel methodology for observing the joint decision-making process of travel choices	PDF
Tai-Yu Ma, Yumeng Fang, Richard Connors, Francesco Viti and Haruko Nakao	A fast algorithm to optimize meeting-point-based electric first-mile feeder services with capacitated charging stations	PDF
Dimitrios Argyros, Renming Liu, Ravi Seshadri, Felipe Rodrigues and Carlos Lima Azevedo	Bayesian Optimization of Road Pricing using Agent-based Mobility Simulation	PDF
Hao Yin and Elisabetta Cherchi	A stated choice experiment to estimate preference for fully automated taxis: comparison between immersive virtual reality and online surveys	PDF
Fernanda Guajardo and Sebastián Raveau	Travel mode choice modelling of visually impair people through latent variables	PDF
Milad Malekzadeh, Dimitrios Troullinos, Ioannis Papamichail and Markos Papageorgiou	Microscopic Simulation-based Testing of Internal Boundary Control of Lane-free Automated Vehicle Traffic	PDF
Cristian Domarchi and Elisabetta Cherchi	Changes in car ownership due to life events: Insights from the UK Longitudinal Household Survey	PDF
Caio Vitor Beojone and Nikolas Geroliminis	Providing a Revenue-forecasting Scheme to Relocate Groups of Ride-Sourcing Drivers	PDF
Severin Diepolder, Andrea Araldo, Tarek Chouaki, Santa Maiti, Sebastian Horl and Costantinos Antoniou	On the Computation of Accessibility Provided by Shared Mobility	PDF
Andrea Pellegrini and John Rose	On allowing endogenous minimum consumption bounds in the Multiple Discrete Continuous Choice Model: An application to expenditure patterns	PDF
Hamoun Pourroshanfekr Arabani, Mattias Ingelström, Mats Alaküla and Francisco J. Márquez-Fernández	MATSim-based assessment of fast charging infrastructure needs for a full-electric passenger car fleet on long-distance trips in Sweden	PDF
Elham Hajhashemi, Patricia Lavieri and Neema Nassir	Applying a latent class cluster analysis to identify consumer segments of electric vehicle charging styles	PDF
Janody Pougala, Tim Hillel and Michel Bierlaire	Modelling the impact of activity duration on utility-based scheduling decisions: a comparative analysis	PDF
Haoye Chen, Jan Kronqvist, Wilco Burghout, Erik Jenelius and Zhenliang Ma	Mixed Integer Formulation with Linear Constraints for Integrated Service Operations and Traveler Choices in Multimodal Mobility Systems	PDF
Kacper Rossa, Andrew Smith, Richard Batley and Phillip Hudson	The valuation of arrival and departure delays in the UK passenger rail using satisfaction survey data	PDF
Kailin Chen, Daniel Graham, Richard		

Anderson, Anupriya Anupriya and Prateek Bansal	Understanding the Capacity of Airport Runways	PDF
Bogdan Kapatsila, Dea van Lierop, Francisco J. Bahamonde-Birke and Emily Gris�	The Effect of Incentives on the Actions Transit Riders Make in Response to Crowding	PDF
Nejc Gerzini�, Maurizio van Dalen, Barth Donners and Oded Cats	The impact of covid-19 on modal shift in long-distance travel	PDF
Ryota Okazaki, Yuki Oyama, Naoto Imura and Katsuhiko Nishinari	Day-to-day delivery demand management: Evaluation based on routing efficiency and customer satisfaction	PDF
Anna Reiffer and Peter Vortisch	Estimating Household-Level Time-Use within a Week Activity Scheduling Framework – Application of the MDCEV Model	PDF
Aurore Sallard and Milos Balac	Bayesian Networks for travel demand generation: An application to Switzerland	PDF
Nico Kuehnel, Shivam Arora, Felix Zwick and Qin Zhang	Simulated Annealing in a Co-Evolutionary, Agent-Based Transport Modeling Framework - The Example of Ride-pooling Driver Supply Optimization	PDF
Sijia Sun, Hossam Abdelghaffar, S�rgio Batista, M�nica Men�ndez and Yuanqing Wang	Analyzing Network-wide Energy Consumption of Electric Vehicles in a Multimodal Traffic Context: Insights from Drone Data	PDF
Gaurav Malik and Chris Tamp�re	Application of a Metamodel-Based Optimization Approach for Toll Optimization and its comparison with Metaheuristics-based Model Optimization via a Case Study.	PDF
Bingyuan Huang, Hans W�st and Mathijs de Haas	Assessing the Long-term Impact of E-bikes on Sustainable Mobility: A National-Level Study in the Netherlands	PDF
Lorena Torres Lahoz, Francisco Camara Pereira, Georges Sfeir, Ioanna Arkoud, Mayara Moraes Monteiro and Carlos Lima Azevedo	Attitudes and Latent Class Choice Models using Machine Learning	PDF
Ana Tsui Moreno, Matthias Langer and Rolf Moeckel	How mobile are persons with mobility restrictions? Analysis of number of days with activities using one-week activity schedules in Germany	PDF
Saumya Bhatnagar, Rongge Guo, Jihui Ma and Mauro Vallati	Prediction of Passengers Demand for Customized Bus Systems	PDF
Josephine Grau, Lea Fuchs, Torben Lelke and Peter Vortisch	City-wide bottleneck and deficiency analysis on a road network generated from the Open Street Map road network using Floating Car Data (FCD)	PDF
Benjamin Gramsch Calvo and Kay W. Axhausen	Exploring the impact of the social network geography on the individual's activity space using structural equation models	PDF
Adrien Nicolet and Bilge Atasoy	Choice-driven Service Network Design and Pricing in Intermodal Transport	PDF
Ida Kristoffersson and Chengxi Liu	Estimation of demand models for long-distance cross-border travel	PDF
Lynn Fayed, Gustav Nilsson and	On the dynamic pricing of pool ride-hailing services in bus	PDF

Nikolas Geroliminis	lanes	
Jaime Soza-Parra and Oded Cats	Who is ready to live a car-independent lifestyle? A latent class cluster analysis of attitudes towards car ownership and usage	PDF
Marko Maljkovic, Gustav Nilsson and Nikolas Geroliminis	On fair discounted charging in electric ride-hailing markets with limited budgets	PDF
Daniel Heimgartner and Kay W. Axhausen	Contributions of Can, May and Want to the Home Office Frequency Decision	PDF
Margarita Gutjar, Chiara Calastri and Matthias Kowald	Householdfleet adaptation as reaction to price regulations: A stated adaptation experiment on the promotion of electric vehicles	PDF
Jing Lyu, Feixiong Liao and Soora Rasouli	Modeling Visit Probabilities within Space-Time Prisms of Daily Activity-Travel Patterns	PDF
Gijsbert Koen de Clercq, Maaïke Snelder, Arjan van Binsbergen and Bart van Arem	Analysing the Effects of Adding Shared Electric Bicycles as a New Mode on the Modal Split of Multimodal Trips between Delft and Rotterdam Using an Unlabelled Multimodal Supernetwork	PDF
Allister Loder and Klaus Bogenberger	MobilityCoins - an integrated multimodal Wardropian model for policy analysis	PDF
Gabriel Nova, C. Angelo Guevara, Stephane Hess and Thomas O. Hancock	Random Utility Maximization model considering the information search process	PDF
Jose Holguin-Veras, Diana Ramirez-Rios and Trilce Encarnacion	Who is Responsible for the Externalities Produced by Freight Carriers? Hint: The Answer is Not as Simple as it Seems...	PDF
Eui-Jin Kim and Prateek Bansal	A new flexible and interpretable choice model with monotonicity constraints, non-linearity, and taste heterogeneity	PDF
Nicole Adler, Gianmarco Andreana and Gerben de Jong	Competing on Emissions Charges	PDF
Chris ten Dam, Francisco Bahamonde-Birke, Dick Ettema, Gert Jan Kramer and Vinzenz Koning	The influence of the built environment on real world car energy efficiency	PDF

Empirical analysis of cycling trends in two of Europe’s most bicycle-friendly regions: Identifying the successes and the setbacks

Anders Fjendbo Jensen*¹ and Jeppe Rich²

¹Associate Professor, DTU Management, Technical University of Denmark

²Professor, DTU Management, Technical University of Denmark

SHORT SUMMARY

We present a combined longitudinal and socioeconomic study of cycling demand in the Netherlands and Denmark from 2010 to 2021. The countries are comparable in demography and both countries have well-developed cycling cultures. The longitudinal data allow us to study successes and setbacks related to cycling uptake over time. E-bikes are successfully promoting longer cycling trips and increased cycling among elderly people, particularly in the Netherlands. However, in rural areas of Denmark, we see setbacks in the form of significant reductions in cycling among children. By applying an econometric model of the combined selection effect of bicycling and the mileage effect conditional on travelling by bike, we analyse how these effects are related to, e.g. increased distances to school, increased work distance, car ownership, changes in urbanisation and other socio-economic factors. We show that these factors cannot explain all of the decline for childrens cycling in Denmark.

Keywords: Cycling demand Mode choice Cycling behaviour Longitudinal analysis Cohort study

1 INTRODUCTION

Research show that cycling leads to better health, reduced congestion and better accessibility (Rich et al., 2021) and thus Bicycle research has received increasing awareness in recent years (Heinen & Götschi, 2022). Several studies have studied cycling across countries (Schneider et al., 2022; Goel et al., 2022; Buehler & Goel, 2022; Pucher & Buehler, 2008). Longitudinal studies are rare and typically from few countries with available data. Harms et al. (2014) found increasing cycling volumes in the Netherlands in urban areas and decreasing volumes in rural areas between 1996-2012. Van Goeverden et al. (2013) found that national bicycle shares on average were stable, whereas Harms et al. (2014) showed that educational trips experienced a significant increase from 1994-1996 to 2007-2009. A more recent study by Kroesen & van Wee (2021) found that the level of urbanisation was the strongest predictor of cycling distance.

It is yet unclear how different effects are related to the cycling uptake for different groups in areas with different degree of urbanisation when considered in a longitudinal perspective. Better insight is needed to meet unfavourable tendencies with the right policy measures. As an example, if younger age groups have a low cycling uptake locally, but not globally, it suggest that there are local conditions that prevent young people from cycling. Global trends on the other hand, may signal wider behavioural tendencies that could result from social media usage (Meyer et al., 2021) or other factors. These tendencies require a different set of actions.

In this paper we present a combined longitudinal and socioeconomic study of cycling demand in the Netherlands and Denmark in the period from 2010 to 2021. The two countries are among the most successfully cycling nations in the world and are by all means comparable economically and socially. The two countries also host and maintain two of the most comprehensive trip diaries in the world (Christensen, 2020; Boonstra et al., 2022). The trip diaries are very similar in structure and scope and are of high quality. This allows a comparison of the cycling uptake across socioeconomic groups and over time.

Year	NL	Trips	Trips	Trips	Km	Km	DK	Trips	Trips	Trips	Km	Km
	N	All	Bike	E-bike	Bike	E-bike	N	All	Bike	E-Bike	Bike	E-Bike
2010	44165	2.81	0.76	NA	3.28	NA	23705	2.90	0.43	0.01	3.08	5.72
2011	38722	2.72	0.78	NA	3.44	NA	18009	2.92	0.44	0.01	2.94	4.58
2012	43307	2.72	0.79	NA	3.31	NA	9696	2.97	0.46	0.01	3.14	5.94
2013	42350	2.73	0.76	NA	3.41	NA	8912	2.72	0.44	0.01	3.45	5.47
2014	42600	2.73	0.81	NA	3.60	NA	9581	2.76	0.45	0.02	3.62	5.02
2015	37350	2.64	0.74	NA	3.45	NA	8720	2.81	0.41	0.02	3.39	6.44
2016	37229	2.59	0.72	NA	3.45	NA	8794	3.03	0.44	0.02	3.21	4.10
2017	38127	2.54	0.70	NA	3.48	NA	9920	3.11	0.41	0.03	3.36	5.16
2018	57260	2.8	0.76	0.15	3.46	6.01	11087	2.90	0.42	0.03	3.41	4.38
2019	53380	2.73	0.66	0.17	3.39	5.86	10204	2.84	0.39	0.03	3.31	4.71
2020	62940	2.34	0.51	0.18	3.53	6.34	12161	2.98	0.38	0.05	3.53	4.71
2021	67083	2.52	0.50	0.23	3.56	5.83	10153	3.14	0.33	0.05	3.41	3.88
Total	564513	2.66	0.71	0.18	3.45	6.01	140942	2.92	0.42	0.03	3.32	5.01

Table 1: Number of observations and average trips per year in the two trip diaries. Trips are calculated as weighted averages based on the official sample weights. 'Bike km' represent the corresponding weighted bicycle mileage per biked trip excluding e-bikes.

2 DATA

In Denmark, The Danish National Travel Survey (TU) (Christensen, 2020) constitutes a representative cross-sectional sample of the Danish population between the ages of 10 and 84. The survey collects information concerning the daily travel habits of approximately 1,000 Danish respondents per month. In the Netherlands, the Dutch national travel survey has been conducted since 1985, but only since 2018 in its current format. During the course of the whole year, approximately 45,000 individuals aged 6 and older currently participate, which corresponds to approximately 0.2% of the Dutch population. In both surveys one individual per household is selected. In Table 1 we present key bicycle descriptive variables in the period between 2010-2021.

It is notable that e-bikes constitute an increasing share of bicycles and travel significantly longer. However, the share in Denmark is small compared to the Netherlands. It is also interesting to see that, by-and-large, trip frequencies for bicycles (when combining electric bicycles and conventional bicycles) are largely constant, while bicycle mileage seems to increase slightly.

3 METHODOLOGY

With the absence of panel data a first and straightforward measurement of bicycle trends over time can be based on age-cohorts (Rich et al., 2022). The idea is to study how a group of people at a given age (e.g., 10-20 years of age) behave in different years. This allows us to examine if 10-20 year olds in the year 2010 behave differently compared to those in the same age group observed in the year 2021.

In Rich et al. (2022), bicycling demand was decomposed into a selection effects and a mileage effect. The selection was measured as a binary variable if the person travelled by bike on the day of the interview. The mileage effect was measured as the total bicycle mileage on the day provided that the interview person travelled by bike. In the paper, the two effects was modelled in a Hurdle-type model (Cragg, 1971) and it was found that the selection effect was the main driver of changes while the conditional mileage effect were largely inconclusive. This paper is inspired by the same idea but develops models for both countries separately.

In addition, we include more explanatory variables in order to assess whether the trends in bicycle demand can be explained by external factors. These factors include car ownership, household characteristics, urbanisation and distance to school and work. In case we still see a significant trends, it suggest that other factors not included in the model are at play.

When modelling the development in bicycle demand, we use a decomposition of the probability of cycling on a given day, and conditionally on people cycling, the total mileage. This selection process is estimated using a logistic regression as shown in Eq. (1). In Eq. (1) $y = 0$ indicates that a person did not cycle, whereas $y > 0$ indicates that the person engaged in cycling activities

on the day.

The mileage model in Eq. (2) is modelled as a standard generalized linear model. The aim is to classify the variables that drive the bicycle demand rather than predicting bicycle demand (Efron, 2020).

$$P_{n,t}(y = 0|x) = 1 - \frac{1}{1 + e^{-(-x_{n,t} \cdot \beta_s)}} \quad (1)$$

$$g_{n,t}(y) = x_{n,t} \cdot \beta_m + u_{n,t}, \quad u \sim N(0, \sigma_u) \quad (2)$$

In this case we apply a link-function $g(y) = \ln(y)$, which corresponds to the assumptions in a log-normal Hurdle model.

For each country, we estimate a logistic regression (Eq. 1) and a mileage model (Eq. 2) for all individuals $n = 1, \dots, N$ and all years $t = 1, \dots, T$. In all models and for all countries, the following set of explanatory variables are used:

$$\begin{aligned} x_{n,t}\beta = & k_t + \beta_f \cdot Female_{t,n} + \beta_c \cdot Cars_{t,n} + \beta_{sd} \cdot School_dist_{t,n} \\ & + \beta_{ch} \cdot Children_dist_{t,n} + \beta_{wd} \cdot Work_dist_{t,n} + \beta_e \cdot Ebiket_{t,n} \\ & + \beta_{u1} \cdot Urban_{n,2} + \dots + \beta_{u4} \cdot Urban_{n,5} + \beta_{a1,t1} \cdot Age_group_{n,t1,1} \\ & + \dots + \beta_{a4,t11} \cdot Age_group_{n,t11,4} \end{aligned} \quad (3)$$

With the variables defined as.

- k_t : Dummy variables for each year, of which $T - 1$ can be identified. Hence, 2010 is set as the reference.
- $Female_n$: If the respondent is a female person.
- $Cars_{n,t}$: The number of cars in the household in which n belongs.
- $School_dist_{n,t}$: The distance to school for those respondents who go to elementary school (age ≤ 16 years).
- $Children_{n,t}$: If small children under the age of 6 is present in the household.
- $Ebiken, t$: For the selection model, this variable cannot be included as we do not know the type of bicycle owned by the respondent. However, for the mileage model, it makes sense because we know if the specific trip is carried out by electric bike.
- $Work_dist_{n,t}$: The distance to work for those who work.
- $Urban_{n,t,u}$: Urban classification dummies with $u = 1$ representing densely populated areas and $u = 5$ rural areas. $u = 1$ is the reference level.
- $Age_group_{n,t,a}$: Age cohorts combined with years. In the simple model, we only consider age cohorts with $a = 1$ corresponding to [10-20], $a = 2$ to [20-30], $a = 3$ to [30-66] $a = 4$ to [66-]. However, in a generalized model, we allow for different age cohort parameters over the years, in order to identify if behavior is changing.

As the underlying data are micro-data we track any changes in these underlying variables over the period. Hence, if more people are moving to the cities, this will be measured in the model through the urbanization variables. If people in cities in 2010 and 2020 behave largely identically, we should be able to model these changes with a main-effect variable for urbanization.

The models help us to understand which factors that explain the development in bicycle demand over time. However, most importantly, the models will allow us to see if age groups change behavior over time when accounting for urbanization degree, car ownership, and distances to work and school. If this is the case, it will suggest that there are other factors that need to be investigated in the future, e.g. measures related to perceived or actual safety as an example.



Figure 1: Selection model parameters for different age-groups and years for Denmark. Vertical bars represent 5% confidence level.

4 RESULTS AND DISCUSSION

An interesting hypothesis is if the parameters for the main effect for age groups by year change over the period. This should suggest if there are unexplained effects (beyond the attributes that are included in the simple reference) for a given age group in a given year change over the period. Therefore, for both the selection and mileage models, we combine the parameters $k_{2011}, \dots, k_{2021}$ with the four age groups as well. Hence, for a given year i , we estimate not one (e.g., k_{2011}), but four main effects corresponding to the four age groups (e.g., $Age_{group1} * k_{2011}$, $Age_{group2} * k_{2011}$, $Age_{group3} * k_{2011}$ and $Age_{group4} * k_{2011}$).

The results from the selection models indicate that distance to work and school indeed influence the preference for bicycle in a negative way. Gender, has opposite sign in the two countries. Hence, females are more likely to bike in Denmark, whereas the opposite is true in the Netherlands. We also find that the presence of cars in the household significantly reduced the likelihood of travelling by bike. Similarly, there is a increasing likelihood of using bicycle in urbanised areas. Across all years, the older people gets, the less likely they are to choose the bicycle. The same pattern is seen for both countries.

Rather than including the full table of parameters, we present only the specific main effect parameters in Figure 1 and 2. The vertical bars represent a 5% level of confidence.

It is interesting to see that for Denmark, for all age groups over the age of 20, there is no significant difference in parameters. However, for young respondents between 10-20 years of age there is a significant decline in the parameters over the years. This suggests that there are other factors in play. For the Netherlands, the pattern is slightly different. While the younger respondents also display a decline, the other age groups show an increase. Most notably for the oldest age group, where we see a significant increase.

The mileage models suggests that the conditional mileage for the youngest age group declines slightly over the period, although the parameters are not statistically different. The only systematic change is seen for the oldest age group in the Netherlands where the tendency to travel longer increases. Presumably, this trend is driven by an increase in electric bicycles for this age group.

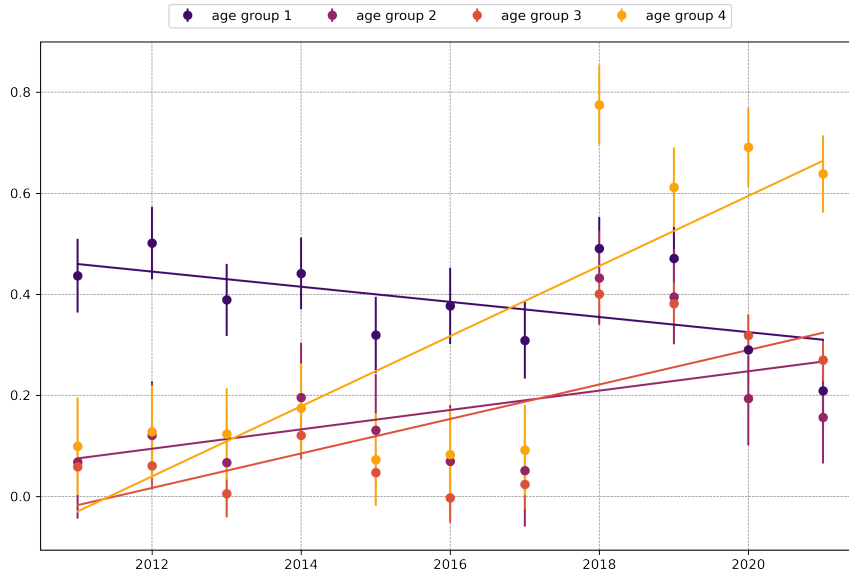


Figure 2: Selection model parameters for different age-groups and years for the Netherlands. Vertical bars represent 5% confidence level.

5 CONCLUSIONS

Based on longitudinal data from The Netherlands and Denmark, which are both countries with a high cycling demand compared to other countries, we have applied an econometric model to study how different factors are related to cycling uptake and use in these countries. In both countries, an increasing distance to school, more cars in the household, the share of households in rural areas, and higher age is related to a lower probability that a person will use a bicycle. Even when these factors are taken into account, there is a tendency of a decline in preference for cycling for younger people in Denmark. We cannot pinpoint the origin of these changes, but a possible explanation is that safety perception play an increasingly important role for kids and their parents. For the Netherlands, on the other hand, there are other positive unknowns affecting the oldest group. Whether this is related to an increasing awareness of the positive health effects from cycling is difficult to say. We find that the conditional mileage for the youngest age group declines slightly over the period, although the parameters are not statistically different. The only systematic change is for the oldest age group in the Netherlands where the tendency to travel longer increases. Presumably, this trend is driven by an increase in electric bicycles for this age group.

ACKNOWLEDGEMENTS

We would like to thank Josephine M. Tönqvist and Negin Farhang for their work with the preparation of data. Furthermore, we would like to thank Kira Janstrup and Maarten Kroesen for fruitful discussions during the analysis process based on their great knowledge on the cycling and the national travel surveys in Denmark and The Netherlands respectively.

REFERENCES

- Boonstra, H. J., van den Brakel, J., & Wüst, H. (2022). Modelling mobility trends - update including 2021 odin data and covid effects. , 1-312. Retrieved from <https://www.cbs.nl/en-gb/background/2022/42/time-series-modelling-of-mobility-trends-1999-2021>
- Buehler, R., & Goel, R. (2022). Chapter seven - a global overview of cycling trends. In E. Heinen & T. Götschi (Eds.), *Cycling* (Vol. 10, p. 137-158). Academic Press. Retrieved from <https://>

www.sciencedirect.com/science/article/pii/S2543000922000191 doi: <https://doi.org/10.1016/bs.atpp.2022.04.007>

- Christensen, H. (2020). National travel survey 2020. , 1–29. Retrieved from www.cta.man.dtu.dk/english/national-travel-survey
- Cragg, J. G. (1971, 9). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, *39*(5), 829. Retrieved from <https://www.jstor.org/stable/1909582?origin=crossref> doi: 10.2307/1909582
- Efron, B. (2020). Prediction, estimation, and attribution. *Journal of the American Statistical Association*, *115*(530), 636-655. Retrieved from <https://doi.org/10.1080/01621459.2020.1762613> doi: 10.1080/01621459.2020.1762613
- Goel, R., Goodman, A., Aldred, R., Nakamura, R., Tatah, L., Garcia, L. M. T., ... Woodcock, J. (2022). Cycling behaviour in 17 countries across 6 continents: levels of cycling, who cycles, for what purpose, and how far? *Transport Reviews*, *42*(1), 58-81. Retrieved from <https://doi.org/10.1080/01441647.2021.1915898> doi: 10.1080/01441647.2021.1915898
- Harms, L., Bertolini, L., & te Brömmelstroet, M. (2014). Spatial and social variations in cycling patterns in a mature cycling country exploring differences and trends. *Journal of Transport and Health*, *1*(4), 232-242. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2214140514000802> (Walking and Cycling: The contributions of health and transport geography) doi: <https://doi.org/10.1016/j.jth.2014.09.012>
- Heinen, E., & Götschi, T. (2022). Chapter one - cycling: Past, current and future. In E. Heinen & T. Götschi (Eds.), *Cycling* (Vol. 10, p. 1-6). Academic Press. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2543000922000130> doi: <https://doi.org/10.1016/bs.atpp.2022.04.001>
- Kroesen, M., & van Wee, G. P. (2021). Autobezit en autogebruik onder jongeren en visies ten aanzien van deelmobiliteit. *Working paper - Delft University of Technology*.
- Meyer, S., Kolbenstvedt, M., & Ellis, I. O. (2021, December). *Tiltak for barns egenstyrte mobilitet og aktivitet* (Tech. Rep. No. 1864/2021). Transportøkonomisk Institutt.
- Pucher, J., & Buehler, R. (2008). Making cycling irresistible: Lessons from the netherlands, denmark and germany. *Transport Reviews*, *28*(4), 495-528. Retrieved from <https://doi.org/10.1080/01441640701806612> doi: 10.1080/01441640701806612
- Rich, J., Jensen, A. F., Pilegaard, N., & Hallberg, M. (2021). Cost-benefit of bicycle infrastructure with e-bikes and cycle superhighways. *Case Studies on Transport Policy*, *9*(2), 608-615. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2213624X21000353> doi: <https://doi.org/10.1016/j.cstp.2021.02.015>
- Rich, J., Myhrmann, M. S., & Mabit, S. M. (2022). Our children cycle less - a danish pseudo-panel analysis. *Journal of Transport Geography*.
- Schneider, F., Jensen, A. F., Daamen, W., & Hoogendoorn, S. (2022, July). Empirical analysis of cycling distances in three of Europe's most bicycle-friendly regions within an accessibility framework. *International Journal of Sustainable Transportation*, 1–15. Retrieved 2023-01-11, from <https://www.tandfonline.com/doi/full/10.1080/15568318.2022.2095945> doi: 10.1080/15568318.2022.2095945
- Van Goeverden, C., Van Arem, B. G., T.Vehlow, C., Reinhardt, T., & Weiskopf, D. (2013). The unexpected stable market share of the bicycle in the netherlands. *WCTR 2013: 13th World Conference on Transport Research, Rio de Janeiro, Brazil, 15-18 July 2013*. Retrieved from <http://resolver.tudelft.nl/uuid:31098da0-bcd3-450f-b7ab-704f9beb691f>

Using computer vision-enriched discrete choice models to assess the visual impact of transport infrastructure renewal projects:

A case study of the Delft railway zone

Sander van Cranenburgh¹

Francisco Garrido Valenzuela¹

¹CityAI lab, Transport and Logistics Group, Delft University of Technology

Abstract

a computer vision-enriched discrete choice model to investigate the impact of the redevelopment of the Delft railway zone on the visual environment. Using computer vision-enriched discrete choice models, we evaluate the changes in the utility levels derived from the visual environment by analysing over 70k street-view images from periods before and after the redevelopment of the railway zone in Delft. We find evidence that the visual appearance of the railway zone has considerably improved after the redevelopment project. This finding highlights the potential of using computer vision-enriched discrete choice models to quantitatively evaluate and monitor changes to the visual environment arising from new transport infrastructure projects.

Keywords

Discrete choice modelling, Appraisal, Computer vision, Street-view images

1. Introduction

Cost-Benefit Analysis (CBA) involves tallying up all costs of a (transport) project and subtracting that amount from the total projected benefits of the project. Since the benefits are often not expressed in euros, the latter requires a monetisation step to convert these to euros. Some benefits of transport projects are comparatively easy to monetise, such as travel time savings. Over the years, an extensive practice has been established (Small 2012; Kouwenhoven et al. 2014). Other effects are still more challenging to monetise, often because of their abstract or enigmatic nature. One key example of a hard-to-monetise benefit involves changes to the visual appearance and environment. Transport projects often have a major (visual) impact on the landscape. Their visual impact often plays a crucial role in the political debate leading to the decision to build new transport infrastructure.

However, the visual impacts of a transport project are typically merely assessed qualitatively. As a result, they are not included in either of the indicators of CBA that are decisive in the political process: the benefit-cost ratio or the net present value (Annema and Koopmans 2015). This weak position of impacts on the visual environment and landscapes, more generally, can lead to poor land-use decisions that cause welfare losses which undermine public support.

Recently, Computer Vision-enriched Discrete Choice Models (henceforth: CV-enriched DCMs) have been proposed (Van Cranenburgh and Garrido-Valenzuela 2023). This model extends the application of discrete choice models towards visual preferences. Van Cranenburgh and Garrido-Valenzuela (2023) demonstrate their use in residential location choices – showing how trade-offs are captured between monthly cost, travel time and street-level factors, such as openness, building typology and greenness (as embedded in images).

This study applies the CV-enriched DCM trained by (Van Cranenburgh and Garrido-Valenzuela 2023) to investigate changes to the visual environments resulting from transport infrastructure projects.

Specifically, we focus on the Delft railway zone, which was put underground in the period 2014-2015. This transformed the visual appearance of the whole railway zone. Using the trained CV-enriched DCM, we compute utility levels for over 70k Google street-view images from before and after the redevelopment of the railway zone. Thereby, we aim to provide a rigorous quantitative underpinning of the benefits to the visual environment arising from the redevelopment of the railway zone. With this work, we contribute to the stream of research that capitalises on street-view images as a source of information about the urban environment (Naik et al. 2017; Rossetti et al. 2019; Ma et al. 2021; Ramírez et al. 2021; Garrido-Valenzuela et al. 2022)

2. Methodology

Our method involves the following four steps. First, we collect street-view images before and after the infrastructure renewal in the surroundings of Delft railway station. Second, we apply the CV-enriched DCM trained by Van Cranenburgh and Garrido-Valenzuela to the images to produce a utility level per image. Third, we aggregate the utilities across spatial hexagons for spatial analysis. Fourth, we analyse the changes in aggregate utility levels before and after the infrastructure renewal.

2.1. Delft railway station area and image data collection

After years of fierce political debate, in the early 2000s, the final decision was made to redevelop the railway zone in Delft. A significant part of the project involved putting 2.3km of railway track and the train station underground. The main construction period on the railway track took place in 2014 and 2015. Figure 1 shows the railway station before and after the redevelopment. Much of the debate leading up to the commissioning of the project was about whether the improvement in the visual environment in the railway zone actually would exceed the considerable construction costs. Because of the significant financial burden the project turned out to be, even today, the project's benefits are debated.



Figure 1: Delft Railway station before (left) and after (right) redevelopment

To collect street-view images in the Delft railway zone, we created a grid of points with 25-metre spacing in an 800m circumference around the city centre. We retrieved the nearest street-view image id for each point on the grid using Google's street-view API. If multiple years were available, we collected images of all available years. Each street-view image id corresponds to a 360-degree panorama view at the street level. Finally, from each panorama, we generated two image urls with 90-degree angles to the direction of the street. This latter ensures the images are side-views (e.g., as opposed to views parallel to the driving direction of the Google car). All street-view images are (temporally) stored using png format with 900 x 600 pixels and 8 bits per colour channel (implying 16.7m colour values per pixel). For each image, the geo-location, year and month are stored. Figure 2 shows the collected number of street-view urls per year.

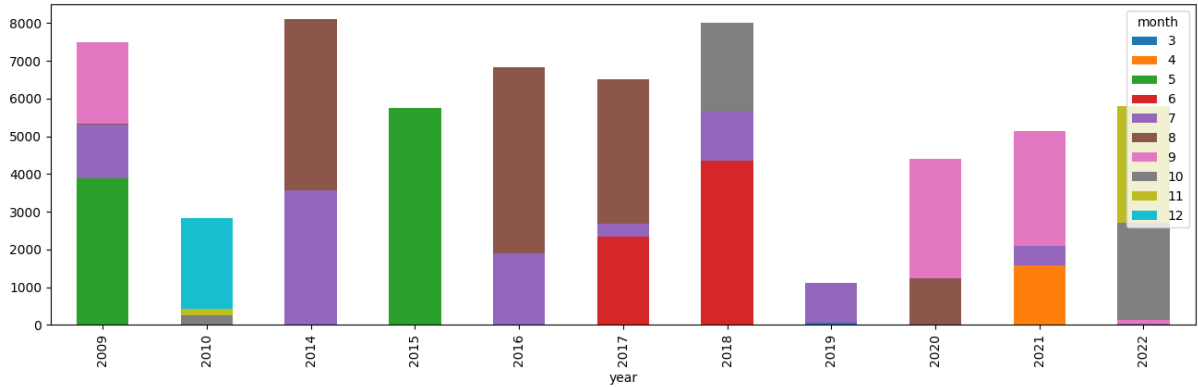


Figure 2: Number of street-view images per year

2.2. Computer vision-enriched discrete choice models

To obtain the utility levels of each image, we apply the CV-enriched DCM trained by Van Cranenburgh and Garrido-Valenzuela (2023) on residential location choice data. This model assumes decision-makers, denoted n , make decisions based on Random Utility Maximising (RUM) principles (McFadden 1974). Equation 1 shows the utility function of this model we use. As can be seen, in this model, the utility, U_j , is derived from the numeric attributes X_j and attributes embedded in the street-view images S_j , which were presented as part of the alternatives. Furthermore, the account for the possibility that images taken, e.g. in spring look, on average, more attractive than images taken in winter, a constant per month is included in the utility function (the second term from the left), where I_{sj} is a binary vector with value one if the image is taken in month mo , and zero otherwise.

Figure 3 shows a screenshot of the stated choice data on which the CV-enriched DCM is trained. In this experiment, respondents had to make trade-offs between street views (visual appearance of the neighbourhood) and two numeric attributes: monthly housing cost and commute travel time. The street-view images shown to respondents in choice tasks were randomly drawn from an extensive database of street-view image ids. Based on these data, the CV-enriched DCM could learn the preferences over elements embedded in the street-view images, such as compactness, openness, street topology, parking facilities, etc. For more details about the data collection and the CV-enriched DCM, see Van Cranenburgh and Garrido-Valenzuela (2023).

Suppose, you have to relocate to a different neighbourhood. Your house stays the same; only the neighbourhood changes. You have two options.

Which option would you choose?

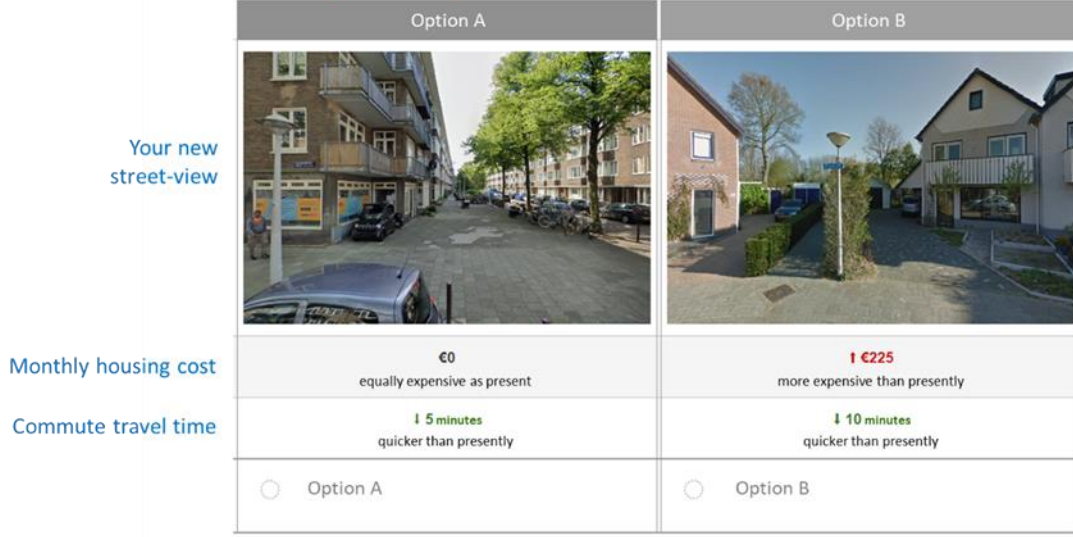


Figure 3: Screenshot of stated choice experiment used to train the CV-enriched DCM

Importantly, in Equation 1, ϕ is a mapping function – performed by the computer vision part of the model: DeiT base (Touvron et al. 2021) – which maps image S_j onto a lower dimensional feature space, denoted Z_j (which has a dimensionality of $1 \times K$). The feature map of an image embeds the relevant information from that image that generates (dis)utility and, in turn, maps linearly map onto the utility. w_k denotes the weight associated with the k^{th} feature of Z_j ; β_m denotes the marginal utility associated with attribute m , and x_{jmn} denotes the attribute level of numeric attribute m of alternative j , as faced by decision-maker n .

$$U_{jn} = \underbrace{\sum_m \beta_m x_{jmn}}_{\text{Utility derived from numeric attributes}} + \underbrace{\sum_{mo} \beta_{mo} I_{S_j}}_{\text{Utility derived from the month of the year}} + \underbrace{\sum_k w_k z_{jkn}}_{\text{Utility derived from image feature map}} + \varepsilon_{jn} \quad \text{where } Z_{jn} = \phi(S_{jn} | w_r) \quad \text{Equation 1}$$

In this application of the model, we are solely interested in the utility it produces from images. In other words, to deploy this model, we first apply the mapping function ϕ to images to obtain feature maps Z . In turn, we take the inner product with w to get the utility of the image. Finally, we 'correct' the utility of each image for the month of the year in which the image was taken (the second term of Equation 1). Figure 4 kernel density plots of the utility levels computed from the images (before and after the renewal project). The left-hand side plot shows utility levels uncorrected for the month of the year; the right-hand side plot shows utility levels corrected for the month of the year.

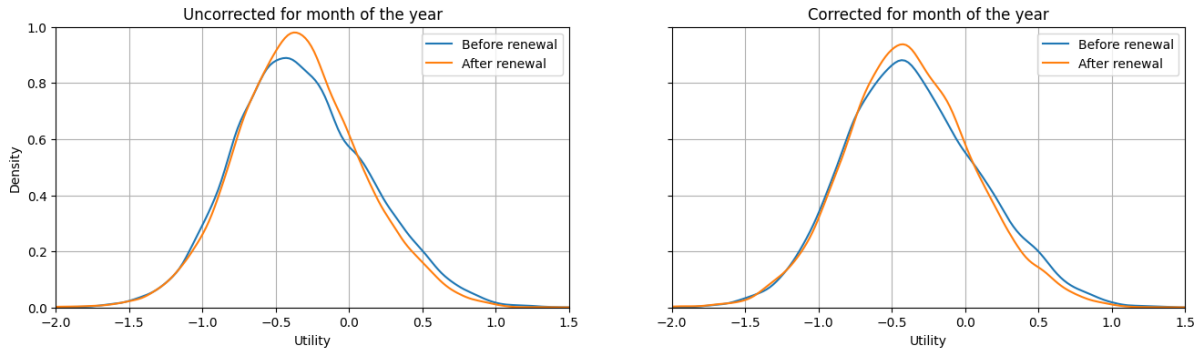


Figure 4: Kernel density plots of utility levels before and after the renewal project.

2.3. Aggregation

To investigate the potential changes in utility levels arising from the changes in the visual environment arising from the infrastructure project, we must define a spatial unit of analysis. For this purpose, we use regular hexagonal cells with 25-metre sides to tessellate the entire study area. Accordingly, on average, each hexagon contains 20 images from the period before and 15 images from the period after the redevelopment.

3. Preliminary results

Figure 5 shows the main results of this study. The hexagon's colour depicts how the utility level changed between the periods before and after the redevelopment of the railway zone. A green colour indicates a positive change in the average utility derived from the images within the hexagon; a red colour indicates a negative change in the average utility. The map shows an area of approximately 2 x 2 km. Delft has two landmark churches, which are depicted to ease navigation.

Based on Figure 5, we make a number of observations. Firstly, the visual appearance of the redeveloped area has considerably improved. Almost all hexagons within the red encircled area are greenish, implying a positive change in utility. This is in line with behavioural intuition. It also suggests that CV-enriched DCMs can indeed be used to evaluate changes in the visual environment. Secondly, the visual appearance of the area west of the redeveloped railway zone has also improved. Again, we see predominantly green hexagons. A possible explanation for this observation is that the redeveloped railway zone radiated positively in this direction and led to positive changes to the visual environment. Thirdly, the change in the visual appearance of the inner city (located Nord-East of the train station) is mixed. Some streets seem to have deteriorated (coloured orange and red), while others show positive changes in their visual appearance (coloured green). Presumably, local explanations can be found explaining these changes.

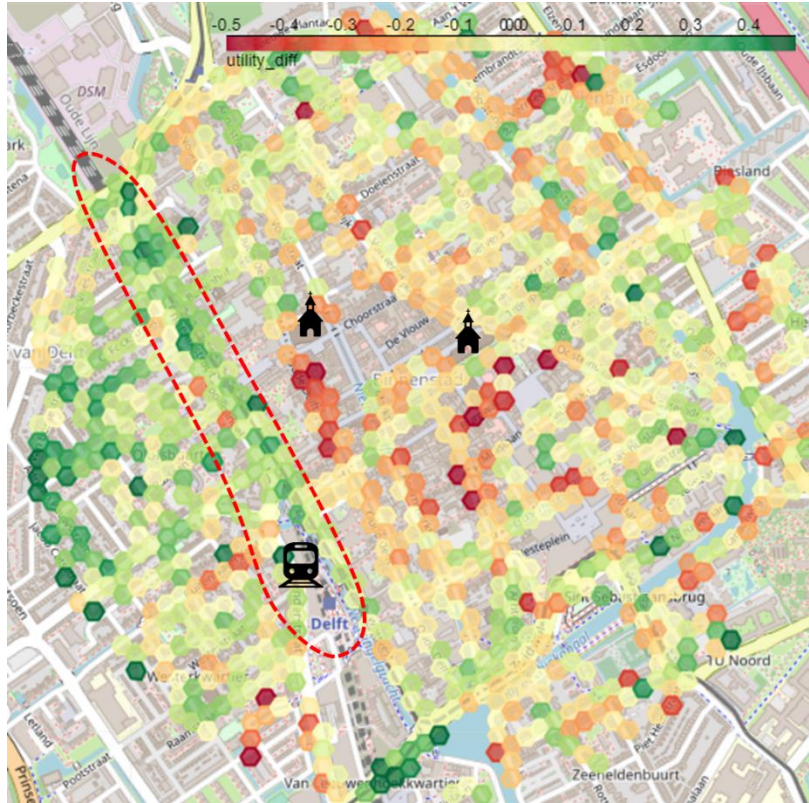


Figure 5: Changes in utility levels Delft. The railway zone is encircled in red.

4. Conclusion and discussion

This research has employed new computer vision-enriched discrete choice models to investigate changes in the visual environment arising from the redevelopment of the Delft railway station zone. The intuitively correct results we obtained from our case study suggest that CV-enriched DCMs can indeed be used to evaluate changes in the visual environment.

A key limitation is this research is that the utilities extracted from the CV-enriched DCM reflect the attractiveness as a residential location. But, the function of the railway station zone is mixed. Its functions include housing, transfer, gathering, working, eating, etc. In this research, we only looked at the visual environment from the lens of residential location.

Next steps

We envision taking the following steps in the coming months. Firstly, we would like to expand our study areas. We want to apply the approach to other areas that undergo renewal to establish its robustness and further applicability. In addition, we aim to develop a better grasp of optimal hexagon size in combination with data availability (Wong 2004). Finally, we aim to show what the trained CV-enriched DCM has learned. We want to understand what causes exceptionally high or low utility predictions. Such model explanations may help to better inform urban planners and policymakers on future transport infrastructure projects.

Acknowledgement

This work is supported by the TU Delft AI Labs programme.

References

- Garrido-Valenzuela, F., van Cranenburgh, S. & Cats, O. (2022). Enriching geospatial data with computer vision to identify urban environment determinants of social interactions. *AGILE: GIScience Series*, 3, 72.
- Kouwenhoven, M., de Jong, G. C., Koster, P., van den Berg, V. A. C., Verhoef, E. T., Bates, J. & Warffemius, P. M. J. (2014). New values of time and reliability in passenger transport in The Netherlands. *Research in Transportation Economics*, 47, 37-49.
- Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C. & Ren, F. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*, 110, 103086.
- McFadden, D. (1974). *The measurement of urban travel demand*. (Berkeley: Institute of Urban & Regional Development, University of California).
- Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L. & Hidalgo, C. A. (2017). Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences*, 114(29), 7571-7576.
- Ramírez, T., Hurtubia, R., Lobel, H. & Rossetti, T. (2021). Measuring heterogeneous perception of urban space with massive data and machine learning: An application to safety. *Landscape and Urban Planning*, 208, 104002.
- Rossetti, T., Lobel, H., Rocco, V. & Hurtubia, R. (2019). Explaining subjective perceptions of public spaces as a function of the built environment: A massive data approach. *Landscape and Urban Planning*, 181, 169-178.
- Small, K. A. (2012). Valuation of travel time. *Economics of Transportation*, 1(1), 2-14.
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A. & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. International Conference on Machine Learning, PMLR.
- Van Cranenburgh, S. & Garrido-Valenzuela, F. (2023) Computer vision-enriched discrete choice models to explore the importance of street-level factors and commute travel times in the residential location choice behaviour,
- Wong, D. W. (2004). The modifiable areal unit problem (MAUP). *WorldMinds: geographical perspectives on 100 problems: commemorating the 100th anniversary of the association of American geographers 1904–2004*, 571-575.

Combine and conquer: model averaging for out-of-distribution forecasting

Stephane Hess^{*1,2} and Sander van Cranenburgh²

¹University of Leeds, * s.hess@leeds.ac.uk

²Delft University of Technology

SHORT SUMMARY

Travel behaviour modellers are increasingly interested in using models from outside the traditional choice modelling area, first incorporating ideas from behavioural economics, such as in regret modelling, before looking at mathematical psychology and machine learning. A key question arises as to how well these different models perform in prediction, especially when predicting trips of different characteristics from those used in estimation. This paper first compares the elasticities and model fit of different models, bringing together models as diverse as logit, random regret, decision field theory and neural networks. We highlight differences in elasticities and also note that the prediction performance deteriorates at different rates for different models when moving further away from the estimation data. We then develop a model averaging approach that allows us to make the most of the entire collection of models and estimate weights for different models as a function of distance away from the estimation sample. **Keywords:** choice modelling; forecasting; machine learning; mathematical psychology; mode choice; model averaging

1 INTRODUCTION

The travel behaviour modelling literature has focussed extensively on two sorts of models, namely models for inference (henceforth inference models) and models for forecasting (henceforth forecasting models). Inference models aim to understand current travel behaviour (e.g. to recover the value of travel time), while forecasting models aim to forecast future travel behaviour in new settings (e.g. due to transport policies, such as toll roads and fuel levies). In other words, forecasting models are developed to generalise out-of-distribution.

Inference models and forecasting models are evaluated differently by analysts. When building an inference model, an analyst is keen that the model generates behaviourally plausible insights into causal factors and their relative impacts. Additionally, for inference models, it is well recognised that they should be able to replicate the behaviour in the empirical data as well as possible. In general, analysts perceive a high model fit as a proxy for a good model. The rationale is that the higher the likelihood of the empirical data given the model, the more reliable the results must be. Given this model evaluation approach, it should come as no surprise that researchers in the travel behaviour field are increasingly attracted by the comparatively good prediction performance of machine learning approaches (cf. Hagenauer & Helbich, 2017).

For forecasting models, the focus during evaluation is typically on elasticities. A widely held view is that forecasting models must produce behaviourally plausible elasticities, i.e. changes in demand in response to changes in journey characteristics. Furthermore, forecasting models are evaluated on their behavioural soundness. The dominant - although not necessarily evidenced-based - view is that forecasting models with a solid behavioural underpinning, such as Random Utility Maximisation (RUM) based discrete choice models, are better equipped to forecast behaviour under new settings than are models with a weak or no behaviour underpinning, such as e.g. machine learning models (cf. van Cranenburgh et al., 2022).

However, what is currently less well understood is how to value and incorporate model performance, i.e. in the model fit on the empirical data, when developing forecasting models. What is clear is that good model performance is not sufficient to establish that a model will make good out-of-distribution predictions. The fact that model parameters are estimated by maximising the model performance on the empirical data, i.e. how well they replicate current choices, is somewhat at odds with the aim of forecasting: to generalise out-of-distribution. The choice modelling literature has at times recognised that more advanced models that offer a better fit on the empirical data do not necessarily lead to better forecasts (see e.g. Fox et al., 2014). But the question is still

open on how to develop forecasting models considering the model’s performance and behavioural underpinning. After all, it is intuitive that the performance of a model on the empirical data still pertains to relevant information on its ability to generalise out-of-distribution.

The aim of this paper is twofold. First, we aim to quantify the deterioration of prediction performance as a function of the “distance” between the training data and forecasting scenarios for models with varying levels of behavioural underpinning. Second, we aim to develop a model averaging-based approach that reduces the bias in forecasts by assigning different weights to different models depending on this “distance”. Using this model, we aim to not only develop a flexible tool for combining models, but to craft rules-of-thumb for the conditions under which what sort of models perform best in terms of out-of-distribution forecasting.

The key hypotheses of the present paper are as follows:

1. Good absolute in-sample prediction performance does not necessarily translate into accuracy of elasticities.
2. Prediction performance for all models deteriorates as a function of the “distance” between the training data and scenarios for which a prediction is made.
3. Models with a solid behavioural underpinning will increasingly perform better than models without behavioural underpinning at increasing “distance”.
4. Flexible models and models with a weak or no behaviour underpinning will suffer from a considerable spread in out-of-distribution generalisation performance.

2 METHODOLOGY

Datasets

We use two different revealed preference datasets in this study, both focussing on mode choice. The first dataset comes from a large-scale survey conducted as part of the DECISIONS project carried out by the Choice Modelling Centre at the University of Leeds (Calastri et al., 2020). The data used for this work corresponds to the observed mode choice behaviour where after extensive data cleaning and data enrichment (Tsoleridis et al., 2022), 12,524 trips made by 540 individuals remained. For each trip, individuals travelled by one of six modes: car, bus, rail, taxi, cycling or walking. Attributes of the alternatives used in the models include in vehicle travel time, out of vehicle travel time, and travel cost.

The second dataset is the London mode choice data compiled by Hillel et al. (2018). This dataset contains four alternatives: walking, cycling, public transport (grouping together bus and rail) and driving. We use a sample of 81,086 trips. Attributes of the alternatives used in the models include in vehicle travel time, out of vehicle travel time, interchanges, and travel cost.

Model types

The following individual models were used in our analysis, combining models from traditional choice modelling, mathematical psychology and machine learning:

Logit models: Standard multinomial Logit models using three different specifications, namely a) linear in attributes; b) log-linear in attributes; and c) linear plus log-linear in attributes.

Nested logit models: Nested Logit using three different specifications, namely a) linear in attributes; b) log-linear in attributes; and c) linear plus log-linear in attributes. In terms of nesting structure, the DECISIONS models grouped together public transport options in a nest, while the London models grouped together motorised modes *vs* active modes.

Random regret minimisation (RRM): RRM models were used, treating attributes as linear, and with different constants depending on choice set size given varying mode availability in the DECISIONS data.

Decision field theory (DFT): DFT is a dynamic, stochastic model, introduced by Busemeyer & Townsend (1993). The key idea of the DFT model is that the preferences for different alternatives update over time whilst the decision-maker considers the different alternatives and their attributes. We use the implementation of T. O. Hancock et al. (2019).

MultiLayer Perceptron (MLP): This model comprises an input layer with input nodes, one or more hidden layers with hidden nodes, and an output layer with output nodes. In this model, signals propagate forward through the links connecting the nodes. The links have a numeric weight w , which is learned from the data. At each node, the weights are multiplied with the input value from the previous nodes and summed. Then the signal is propagated to the next layer using an activation function. We use tanh activation functions. In the output layer, a Softmax function (i.e. a logit) is applied to produce choice probabilities for each alternative.

XGBoost (XGB): The XGB model comprises a series of sequentially applied decision trees. A decision tree is a sequence of simple IF-THEN rules, optimised to classify data accurately. In the XGB model, each decision tree in the sequence ‘corrects’ the mispredictions of the models before it. This process is referred to as ‘boosting’. The word ‘gradient’ relates to the notion that each subsequent decision tree is fitted on residuals of the trees before it.

Identifying the role of “distance”

To study the impact of distance from the estimation data, we first divided the samples into 10 subsets by distance, e.g. using the 10 percent of shortest trips for the first segment.

Distance segments 1 and 10 were excluded from the model fitting work to retain them for later out-of-distribution validation. Five separate models were then estimated for each model type on rolling subsets of the data combining 4 distance segments, e.g. segments 2-5 for the first model, 3-6 for the second, etc. Let us define $M_{m,g}$ to be the model of type m estimated on distance grouping g , where, e.g. with $g = 1$, we would use distance segments 2-5.

Finally, each of the estimated models was used to make a prediction of each trip in the data, independently of the distance segment for that trip, so e.g. also using models estimated on segments 2-5 only to predict mode choice for trips in distance segment 9.

Model estimation relied on 80% of the sample, with the remaining 20% kept for later out-of-sample and out-of-distribution prediction.

Model averaging approach

The final step of the work uses a model averaging approach, as outlined for example by S. H. A. D. Hancock Thomas O. & Fox (2020). Model averaging relies on a sequential latent class approach, where M different models have been estimated on the data, with model m giving a likelihood $P_{n,m}$ for the choice in observation n (working either at the person or observation level).

The model averaging log-likelihood is then given by

$$LL = \sum_{n=1}^N \log \sum_{m=1}^M \pi_{n,m} P_{n,m}, \quad (1)$$

where $\pi_{n,m}$ is the estimated weight given to model m for observation/person n , where this is given by:

$$\pi_{n,m} = \frac{e^{\gamma'_m z_n}}{\sum_{l=1}^M e^{\gamma'_l z_n}}, \quad (2)$$

with an appropriate normalisation, where z_n are characteristics of person/observation n .

Model averaging typically estimates each model on the entire data and then computes the weights by considering how well each model fits for each of the data points.

We use a different approach. Specifically, we rely not only on the observations on which the models were estimated, given that each model uses only 4 distance segments, but include the prediction performance of every model on each of the trips. This allows us to study how the weight assigned by model averaging is a function of both the model type and how far, if at all, out-of-distribution, the model is. In other words, we expect that the weight given to a model decreases as a function of how far away the distance of the trip for each a mode choice is predicted is from the average distance of trips on which the model was estimated.

With this in mind, let D_g be the average distance of trips in distance grouping g , and let d_n be the distance of the specific trip for which we make a prediction, where we work at the level of individual trips.

We then have that our log-likelihood is given by:

$$LL = \sum_{n=1}^N \sum_{g=1}^G \log \sum_{m=1}^M \pi_{n,m,g} P_{n,m,g}, \quad (3)$$

where $G = 5$.

Each observation in the model averaging process thus uses the predictions for all the different model types for a given distance grouping. This approach produced better results than looking jointly at predictions from all GM models in a single row.

The model averaging weights are now specified as

$$\pi_{n,m,g} = \frac{e^{\delta_m + (\gamma_{d,m}(d_n < D_g) + \gamma_{i,m}(d_n > D_g)) \log |D_g - d_n|}}{\sum_{l=1}^M e^{\delta_l + (\gamma_{d,l}(d_n < D_g) + \gamma_{i,l}(d_n > D_g)) \log |D_g - d_n|}}, \quad (4)$$

where δ_m is a constant for model type m , and $\gamma_{d,m}$ and $\gamma_{i,m}$ are parameters capturing the influence on class allocation when the distance of a trip is below, respectively above the average distance of trips on which the model was estimated, where a non-linear transform was used.

3 RESULTS AND DISCUSSION

The work has produced a wealth of results of which we only focus on a subset in this brief paper.

Model fit comparisons

Figure 1 compares the model fit (using ρ^2) in prediction for MLP and linear logit, as a function of the estimation and prediction segment. A positive differences indicates better fit for MLP than logit. We note that the expected patterns emerge, with MLP overall predicting better in-distribution than logit, but losing out when moving out-of-distribution, though with some exceptions.

	Forecast_segment_1	Forecast_segment_2	Forecast_segment_3	Forecast_segment_4	Forecast_segment_5	Forecast_segment_6	Forecast_segment_7	Forecast_segment_8	Forecast_segment_9	Forecast_segment_10
Model_segments_2_5	-0.02105495	0.015811827	0.027874172	0.03394731	0.001921904	-0.01158405	-0.008914669	-0.09583537	0.32564889	0.07702223
Model_segments_3_6	-0.0160828	0.014424376	0.0236738	0.02070766	0.01798398	0.01204426	-0.001856788	-0.03780298	0.038396733	-0.00791557
Model_segments_4_7	-0.03849625	-0.004529412	0.007082065	0.03310145	0.027727421	0.04041223	0.030607856	-0.03929117	-0.133061398	-0.13602774
Model_segments_5_8	-0.17554066	-0.067432785	-0.035770125	-0.03242807	0.02202978	0.03880158	0.027496214	0.0434717	-0.034876721	-0.06171994
Model_segments_6_9	1.0476279	-0.408353156	-0.180015921	-0.11351349	-0.100038998	-0.02941908	-0.02603404	-0.02049614	-0.001933407	-0.03624159

Figure 1: Model prediction comparison between MLP and linear logit: DECISIONS data

Elasticities

Figure 2 compares the car cost elasticities for a subset of models estimated on the DECISIONS data. We see differences as a function of which segment model was estimated on. For most models, we see decrease in elasticities for models estimated on longer distance trips, but not for MLP and especially XGB. This is an initial indication of differences in prediction results for different models.

Model averaging results

Figure 3 and 4 show the weights for different models generated by model averaging as a function of the difference in the distance of the trip under question and the average trip distance used in model estimation. Figure 3 shows the results for the DECISIONS data set; Figure 4 shows the results for the London data set.

The results show a diverse patterns. Firstly, for both data sets we see that the MLP and XGB models outperform the other models best close to the estimation distance, before tailing off with increasing distance. Second, the behavioural models take over for larger distances. Especially notable is how well DFT performs for trips much longer than the estimation data on the DECISIONS data. But, in the London data this patterns is not visible. Third, for trips that are (much) shorter than the estimation data, the log-linear logit model performs well, on both data sets.

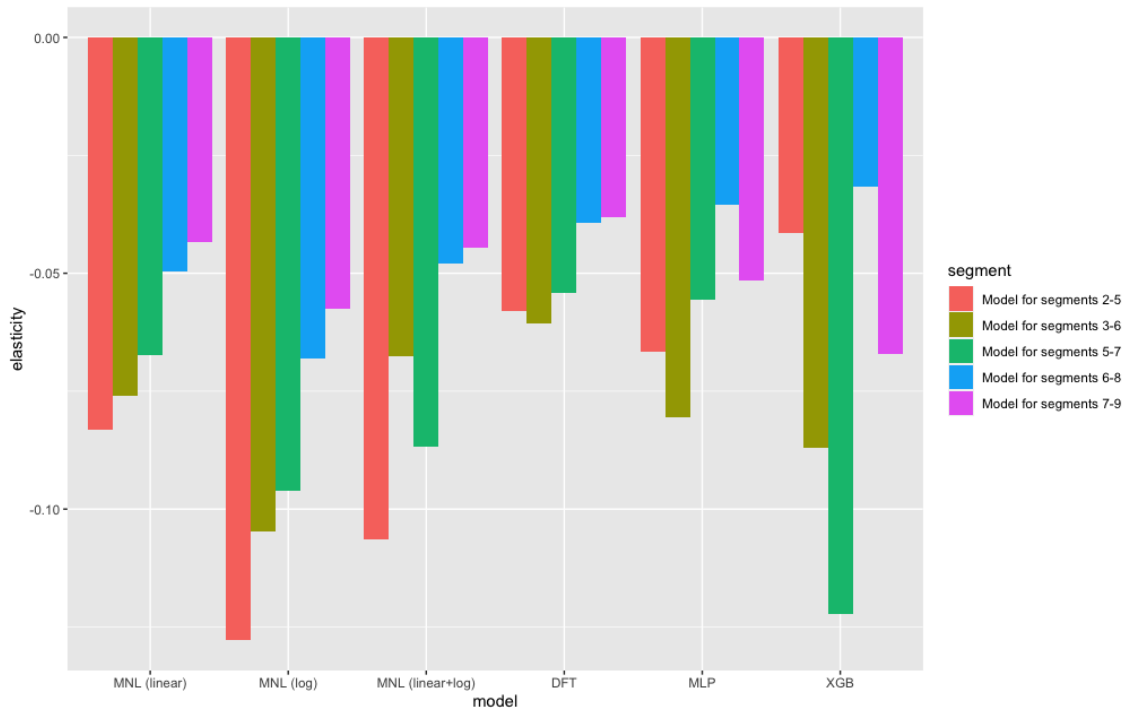


Figure 2: Car cost elasticity comparisons for select models: DECISIONS data

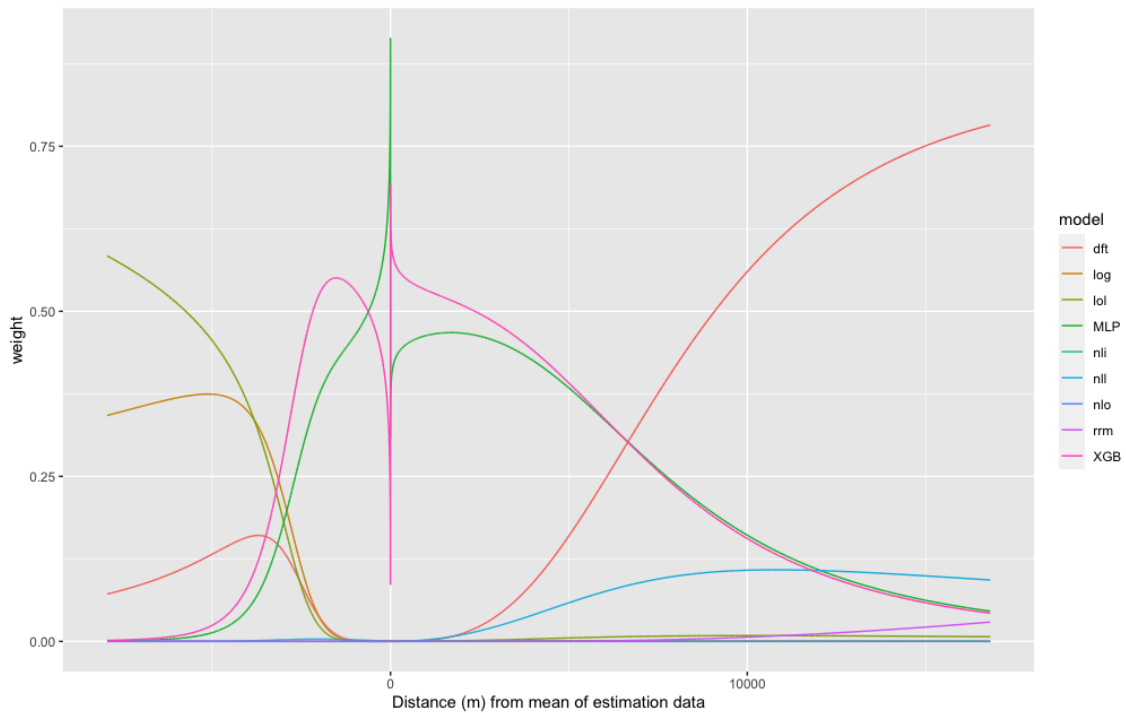


Figure 3: Model averaging weights as a function of distance from estimation data: DECISIONS data (log=log-linear logit, lol=linear plus log-linear logit, nli=linear nested logit, nll=linear plus log-linear nested logit, nlo=log-linear nested logit, all other acronyms as in main text)

4 CONCLUSIONS

This work has taken an important step forward in combining insights from different modelling approaches for travel demand forecasting. Specifically, we have shown that different models predict choices differently well depending on how far away from the estimation sample the prediction takes

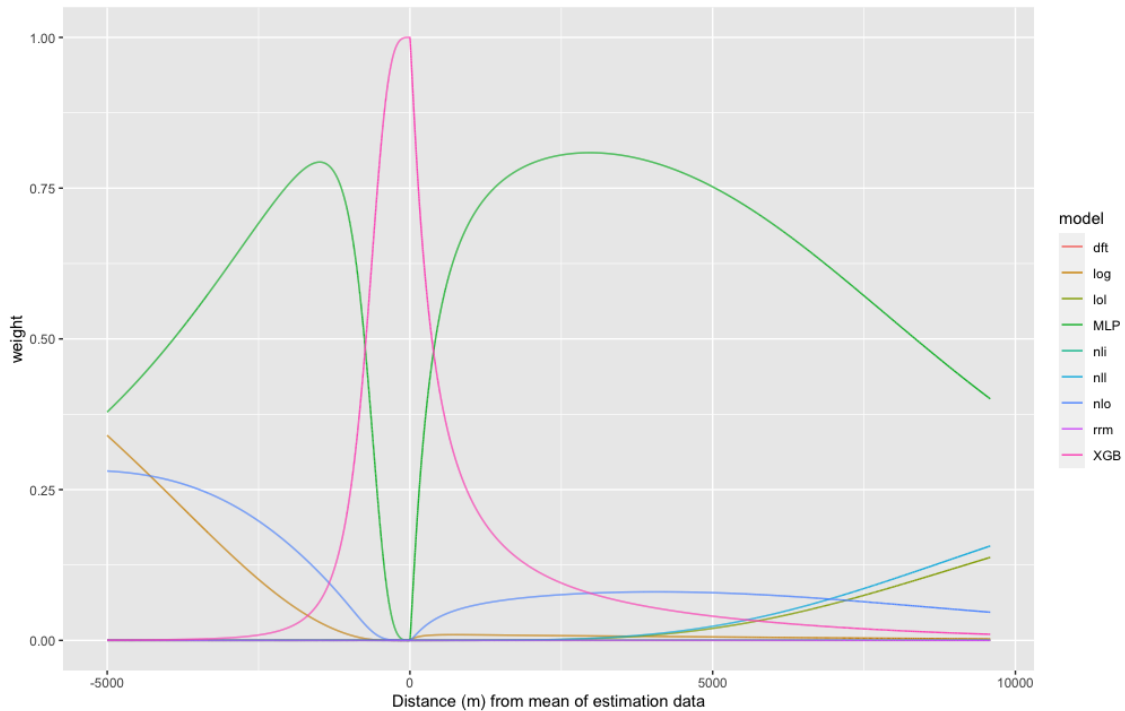


Figure 4: Model averaging weights as a function of distance from estimation data: DECISIONS data (log=log-linear logit, lol=linear plus log-linear logit, nli=linear nested logit, nll=linear plus log-linear nested logit, nlo=log-linear nested logit, all other acronyms as in main text)

place. This result is not surprising in itself, but is quantified by our work.

Building forth on these insights, we contribute by developing a model averaging-based approach that estimates weights for different models as a function of the distance away from the estimation data.

In the full paper, we present more detailed results, including on model fit, prediction out of sample, prediction completely out of distribution (i.e. distance segments 1 and 10), and full estimation results for the models. We also hypothesise about possible other attributes that could be used to measure ‘distance’ from the estimation data, going beyond trip distance alone.

ACKNOWLEDGEMENTS

Collate acknowledgements in this separate section at the end of the text, before the part of references. List here those individuals who provided help during the research (e.g., funding the project, providing language help, writing assistance or proof reading the article, etc.).

REFERENCES

- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100(3), 432.
- Calastri, C., dit Sourd, R. C., & Hess, S. (2020). We want it all: experiences from a survey seeking to capture social network structures, lifetime events and short-term travel and activity planning. *Transportation*, 47(1), 175–201.
- Fox, J., Daly, A., Hess, S., & Miller, E. (2014). Temporal transferability of models of mode-destination choice for the greater toronto and hamilton area. *Journal of Transport and Land Use*, 7(2), 41–62.
- Hagenauer, J., & Helbich, M. (2017). A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*, 78, 273–282.

- Hancock, S. H. A. D., Thomas O., & Fox, J. (2020). Using a sequential latent class approach for model averaging: Benefits in forecasting and behavioural insights. *Transportation Research Part A: Policy and Practice*, 139, 429-454.
- Hancock, T. O., Hess, S., & Choudhury, C. F. (2019). An accumulation of preference: two alternative dynamic models for understanding transport choices. *Submitted*.
- Hillel, T., Elshafie, M. Z. E. B., & Jin, Y. (2018). Recreating passenger mode choice-sets for transport simulation: A case study of london, uk. *Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction*, 171(1), 29-42.
- Tsoleridis, P., Choudhury, C. F., & Hess, S. (2022). Deriving transport appraisal values from emerging revealed preference data. *Transportation Research Part A: Policy and Practice*, 165, 225-245.
- van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., & Walker, J. (2022). Choice modelling in the age of machine learning - discussion paper. *Journal of Choice Modelling*, 42, 100340.

Modeling the ecological and economic footprint of last-mile parcel deliveries using open data: A case study for Lyon

Sebastian Hörl*¹, Jakob Puchinger²

¹ IRT SystemX, France

² EM Normandie, METIS, France

SHORT SUMMARY

The amount of parcels delivered in the urban space is steadily increasing and is often expected to double by 2030. At the same time rising energy prices and policies towards sustainable development affect the business models and distribution schemes in the sector. The present study uses open data to approximate today's parcel volumes for the specific case of Lyon and estimated how those parcels are delivered in terms of used vehicles, covered distances and ecological impacts. The first part describes our data collection process which hypothesizes market shares and cost structures of the parcel operators. In the second part, we solve Heterogeneous Vehicle Routing Problems to uncover the likely distribution schemes. This way, the study provides rough estimates on the total daily emissions and energy used for parcel deliveries outlined pathways for future modeling efforts and data collection.

Keywords: parcels, urban, last-mile, logistics, optimization, VRP

1 INTRODUCTION

The amount of parcels delivered in the urban space is expected to increase strongly in the coming years. Today, cities already reflect upon strategies to regulate urban logistics, understanding the complex interplay between its economic, ecological and social impacts becomes ever more important. While ideas and research efforts on sustainable urban logistics policies are gaining traction (Mucowska, 2021; Neghabadi et al., 2019; Patella et al., 2021). Recent advances in transport simulation aim to model urban logistics on a systemic level (de Bok et al., 2022; Sakai et al., 2020; Toilier et al., 2018), but reliable data remains scarce. The present short paper is an attempt to model one specific sector of urban logistics - home parcel deliveries - solely based on open data for a use case of Lyon.

2 METHODOLOGY

Our approach follows various steps from generating the parcel demand for a territory and defining the supply in terms of operators and distributions center. We then define cost structures to obtain the used vehicles and driven distances to deliver all parcels based on a cost-minimization and vehicle-routing approach. The individual steps are described below.

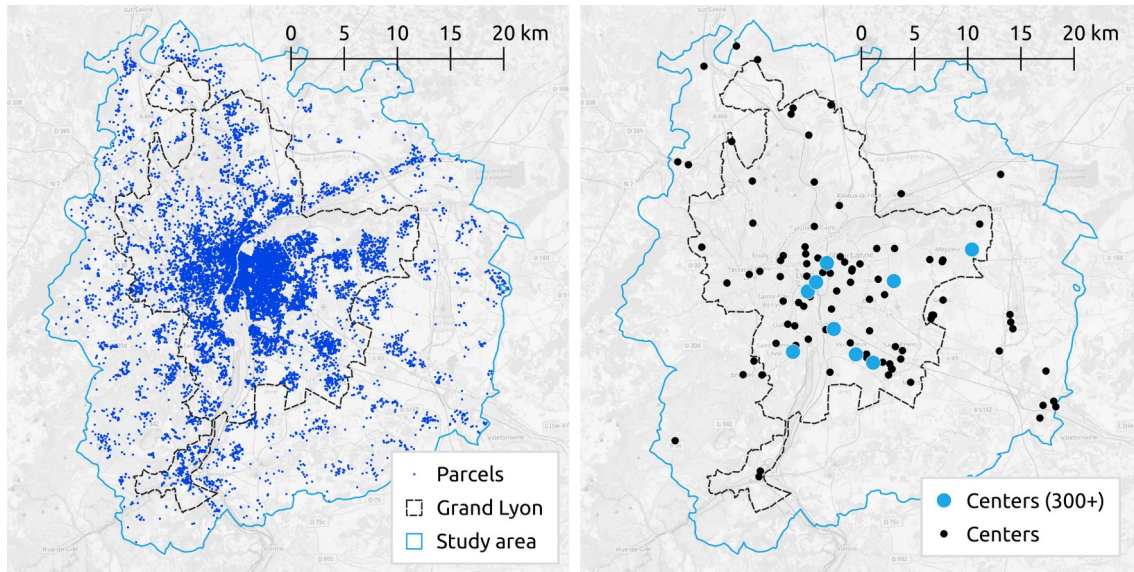


Figure 1: Map of the study area, generated parcels, and distribution centers (Background: OpenStreetMap)

Demand data

Our demand generation process is based on a synthetic population for the Rhône-Alpes region around Lyon. Such a synthetic population, which is a digital representation of households and persons in a region, along with their socio-demographic attributes can be generated based on open data in France. We make use of a replicable data processing pipeline that can be applied anywhere in France (Hörl & Balac, 2021). For the present study, we only consider households in our study area, which comprises the city of Lyon, the Grand Lyon metropolitan region and bordering municipalities with relevant logistics infrastructure (Figure 1). For this perimeter, the population synthesis pipeline generates 1.6 million persons in about 795,000 households.

We fuse the synthetic population data with surveys on the purchasing behavior of the local population. Specifically, (Gardrat, 2019) provides statistics on the annual number of orders made per household based on various socio-demographic characteristics. In (Hörl & Puchinger, 2022) we have proposed a method to make use of this information to generate the probable daily parcel demand for the synthetic households using Iterative Proportional Fitting. Applying the model to our study area yields 16,252 geolocated parcels to be delivered during an average day (Figure 1).

Operator model

The goal of our methodology is to let operators minimize their cost by choosing relevant vehicle types for delivering their assigned parcels and optimizing the vehicle routes. Unfortunately, information on the cost structures of parcel operators is scarce. However, we can assume that the main cost components for offering their service are salaries, vehicle maintenance and investment costs, and per-distance costs. A substantial part of our research was to collect information from gray literature on these cost components. While a detailed analysis of our sources and aggregation procedures exceed the scope of this paper, they will be detailed in an extended publication.

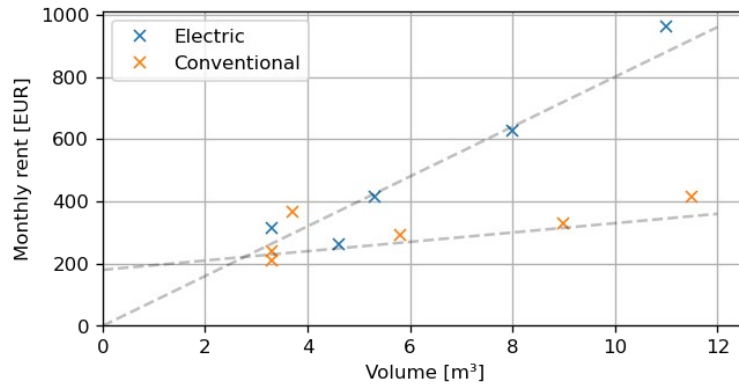


Figure 2: Monthly rent versus transport volume

For personnel costs, we assume an effective net salary of 1,300 EUR per month per driver, leading to an approximate gross salary of 1,700 EUR, and to monthly costs of about 3,400 EUR per full-time employee and month for the operator. Divided by 25 operating days, we arrive at daily salary costs of 136 EUR.

In terms of per-vehicle costs, we have examined the long-term rental offers of the major French vehicle manufacturers, along with the characteristics of the advertised vehicles. This analysis has yielded distinct vehicle classes (of about 3.3m³, 5m³, 10m³) for which costs increase linearly with the transport volume. This is true for thermal and electric vehicles while the slope of the latter is higher (Figure 2). We document the daily unit costs per prototypical vehicle that are used in our model in Table 1 ranging from 210 EUR for a small thermal vehicle up to 800 EUR for large electric truck. In the final optimization we divide these cost by 25 active days per month.

The per-distance costs depend strongly on the consumption of the individual vehicle types. Based on our analysis of manufacturer offers, we have attached representative values for thermal vehicles (in L/100km) and electric vehicles (in Wh/km) to our prototypical vehicle types in Table 1. The per-distance costs are calculated by multiplying the driven distance per vehicle type with the respective consumption factor and the price for fuel (in EUR/L) and electricity (in ct/kWh), respectively.

Additionally, we have noted down representative CO₂ equivalent emissions rates (in g_{CO₂eq}/km) for each vehicle type. The rates for thermal vehicles are based on our manufacturer analysis, while the rates for electric vehicles are based on the French average of 90g_{CO₂eq}/kWh for electricity production¹.

Finally, Table 1 shows the values for a prototypical cargo-bike (Be) based on current rental offers in France and typical consumption rates.

¹<https://www.rte-france.com/eco2mix>

Table 1: Operator model

Vehicle type	St	Mt	Lt	Se	Me	Le	Be
Size	S	M	L	S	M	L	S
Propulsion	T	T	T	E	E	E	E
Capacity	33	50	100	33	50	100	14
Fuel consumption [L/100km]	5	6	8	-	-	-	-
Electricity consumption [Wh/km]	-	-	-	160	200	300	42
Unit cost [EUR/month]	210	260	370	260	400	800	160
Distance cost* [EUR/100km]	304.5.0 0	377.00	522.00	14.00	18.00	27.00	3.80
Emissions** [g _{CO2eq} /km]	130	160	215	14.4	18	27	3.8

Be: Cargo-bike; Size: S – Small, M – Medium, L – Large; Propulsion: T – Thermal, E – Electric;

*Indicative distance costs based on 1.45 EUR/km and 9ct/kWh; **Electric vehicle emissions based on 90g_{CO2eq}/kWh

Operator assignment

To link the demand and the operators, we need to assign an operator to each generated parcel in the synthetic population. For that, we perform weighted random draws from the set of operators based on their market shares. Those market shares have been elaborated from gray literature and a dedicated model. These steps cannot be covered in detail but will be explained in an extended publication. Table 2 shows the resulting parcels assigned to each operator.

Table 2: Operator statistics

Operator	Distribution centers	Market share [%]	Parcels
La Poste (Colissimo)	72	40.08	6,384
Chronopost	6	14.98	2,430
UPS	2	13.55	2,210
DPD	3	9.94	1,632
DHL	8	8.95	1,477
GLS	2	6.93	1,169
Colis privé	2	5.36	917
Fedex	9	0.21	33
Total	104	100	16,252

Distribution centers

To know from where parcels are dispatched, we make use of the SIRENE database², which lists all enterprises and their facilities in France, along with their address and the number of employees. From this database we have extracted all facilities belonging to any of the parcel distributors listed in Table 2. The resulting distribution centers are shown in Figure 1 and Table 2 indicates the number of centers per operator.

For each parcel, we select the distribution center of the respective operator that is the closest in terms of road distance. These road distances are calculated using a network extracted from

²<https://www.sirene.fr>

OpenStreetMap data³ and the osmnx library (Boeing, 2017). The process results in nine distribution centers out of 104 with more than 300 assigned parcels (Figure 1).

Heterogeneous Vehicle Routing Problem

Based on the inputs above, we define a Heterogeneous Vehicle Routing Problem (HVRP) *per distribution center* with the following characteristics:

- The goal is to minimize the overall cost which is the sum of the unit costs and a daily salary per chosen vehicle and the total distance-based cost of the vehicle trajectories.
- The operator can vary the number of vehicles of each of the seven types (Table 1) and the individual vehicle routes.
- Vehicles start at the distribution center and must return before the end of the day. Their total active time cannot exceed a daily duration of 10h. It consists of the travel times between parcels and depot; service times of 120s at delivery; and service times of 60s per pick-up.
- Vehicles cannot carry more parcels than their capacity allows. We assume 10 parcels per m³ in Table 1. We allow multiple tours per vehicle during one day.

For each distribution center, we obtain a distance matrix and a travel time matrix from our extracted OpenStreetMap network using the osmnx library (Boeing, 2017). Since osmnx calculates travel times based on the speed limits of the road segments, we further inflate these values using averaged factors from the TomTom Congestion Index (Cohn et al., 2012) factors of Lyon to arrive at approximately congested travel times.

Finally, we solve the resulting Heterogeneous Vehicle Routing Problems using the open-source VRP solver VROOM⁴.

3 RESULTS AND DISCUSSION

We define three individual scenarios:

- **Baseline:** The scenario is based on our synthetic population for 2022. The prices are chosen such that they reflect the long-term cost structures of the operators that have given rise to the distribution schemes that we see today.
- **Today:** The scenario considers recent increases in energy prices beginning of 2023 with fuel prices of about 1.90 EUR/L⁵ and 14 ct/kWh⁶. It hence shows how the distribution system could develop in case prices stay at this level in the long term.
- **Future:** The scenario is a future scenario in which we consider an updated synthetic population that considers population growth⁷ and a general increase of parcels per capita by a factor of two. We assume that prices have increased by +80% for fuel and +60% for electricity.

³<https://download.geofabrik.de>

⁴<https://github.com/VROOM-Project/vroom>

⁵Diesel, France, 09/01/2023, <https://www.tolls.eu/fuel-prices>

⁶EUROSTAT, Non-household, S2 2022, https://ec.europa.eu/eurostat/cache/infographs/energy_prices/enprices.html

⁷Based on INSEE prediction scenarios <https://www.insee.fr/fr/information/6536990>

The results are shown in Table 3. For the **Baseline** case, we obtain 139 thermal vehicles being used, but only eight electric vehicles and 24 cargo bikes. This reflects today's reality where electric vehicles do not have a large share in the transport system. In terms of vehicle sizes, large vehicles (133) dominate, followed by smaller ones (24) of which the majority are cargo-bikes. Medium-sized vehicles are rarely used. The total distance driven for last-mile deliveries is almost 10,000 km per day for thermal vehicles, the distance for electric vehicles and cargo-bikes is ten times smaller. Only 7% of all parcels are delivered by electric vehicles or cargo-bikes. In terms of consumption, 780 liters of fuel are needed and 236 kWh of electricity. This consumption translates into about 2100 kg of CO₂ equivalents emitted for the last-mile deliveries during one day which makes 131g per parcel. To calculate the total consumed energy we assume a conversion rate of 10 kWh/L and arrive at a total of 8000 kWh per day with 497 Wh per parcel.

For the **Today** scenario with adjusted prices (increase of 30% for fuel and 55% for electricity), we see a slight shift of electric vehicles from eight to 14. Still, this shift represents a doubling of the driven distance of electric vehicles and a doubling in parcels delivered by electric vehicles while their overall percentage remains low with 7% for electric transporters and 12% for cargo-bikes. Accordingly, electricity use doubles while fuel consumption drops by 10%. These shifts lead to a reduction in emissions by 8% in total to 120g per parcel. Total energy use is also reduced by 5% while no large shifts in the cost structures can be observed. Despite electricity prices having increased stronger than fuel prices, the observed shifts can be explained by the different ratios of capital expenses versus operational expenses between thermal and electric vehicles. The latter have higher vehicle prices with lower per-distance costs. In the **Baseline** case, the break-even daily distance at which a single electric vehicle becomes cheaper in total is at about 34 km, while the point shifts to 26 km in the **Today** scenario (see Figure 3).

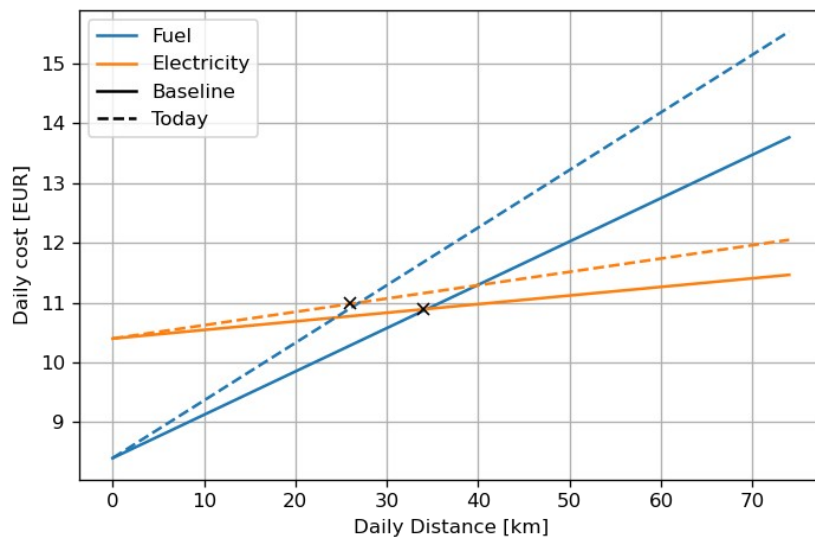


Figure 3: Break-even points for a small electric vehicle

In the **Future** scenario, both the numbers of thermal and electric vehicles increase because of the higher demand. However, electric vehicles increase strongly from 8 to 101. In terms of vehicle size especially large vehicles double in count. At the same time the distance for thermal vehicles goes down and the distance for electric vehicles increases tenfold. Interestingly, the total distance is not doubled, but increases by only 50%, which shows that there are scale effects with

Table 3: Optimization results

	Baseline	Today	Future
Scenario			
Population	2022	2022	2030
Demand factor	1.0	1.0	2.0
Fuel price [EUR/L]	1.45	1.90	3.40
Electricity price [EUR/kWh]	0.09	0.14	0.23
Vehicles by type			
Thermic	139	133	164
Electric	8	14	101
Cargo-bike	24	24	21
Vehicles by size			
Small (S)	32	33	32
Medium (M)	6	6	8
Large (L)	133	132	246
Distances [km]			
Thermic	9,835	8,916	6,185
Electric	1,194	2,294	11,438
Cargo-bike	750	785	861
<i>Total</i>	<i>11,778</i>	<i>11,995</i>	<i>18,484</i>
Parcels			
Thermic	15,001	14,493	21,926
Electric	500	1,009	11,236
Cargo-bike	751	750	761
<i>Total</i>	<i>16,252</i>	<i>16,252</i>	<i>33,923</i>
Consumption			
Fuel [L]	783	710	494
Electricity [kWh]	236	546	3,261
Environment			
Emissions [kg _{CO2eq}]	2,127	1,956	1,622
<i>Per parcel [g_{CO2eq}]</i>	<i>131</i>	<i>120</i>	<i>48</i>
Energy [kWh]	8,071	7,642	8,205
<i>Per parcel [Wh]</i>	<i>497</i>	<i>470</i>	<i>242</i>
Cost [EUR]			
Salaries	23,256	23,256	38,896
Vehicles	2,217	2,302	5,416
Distance	1,159	1,426	2,427
<i>Total</i>	<i>26,632</i>	<i>26,985</i>	<i>46,740</i>
<i>Per Parcel</i>	<i>1.64</i>	<i>1.66</i>	<i>1.38</i>

respect to the transported parcel volumes. While parcels delivered by electric vehicles are rare **Today** they make up 50% of all flows in the **Future** scenario. In the latter, fuel consumption goes down by 37% while electricity use increases by a factor of 13 and total used energy by a factor of 10. On the contrary, total emissions decrease, but only by 24% despite a reduction of 63% *per parcel*. This effect is due to the generally increased demand. Total costs increase by 75% but not equally on all cost components (67% on salaries, 144% on vehicles, 109% on distance), which puts a higher influence on operational costs on the overall costs. Per parcel, there is a margin of 26ct per parcel between the **Today** and **Future** scenario.

In all scenarios, we see that cargo-bikes are rarely used because of their limited capacity.

4 CONCLUSION

In this paper we have documented a model on the economic and ecological characteristics of last-mile parcel deliveries in a city. While the model makes use of a multitude of assumptions, its main value lies in the comparison of scenarios. In the future, more detailed and distinct scenarios should be evaluated. In terms of validation, system-level reference data is not likely to emerge in the near future. Hence, we are engaging actively in discussing our operational assumptions with domain experts and practitioners to compile a comprehensive list of limitations and future improvements, which will be detailed in an extended publication on the model.

ACKNOWLEDGEMENTS

This paper presents work developed in the scope of the project LEAD, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 861598. The content of this paper does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in this paper lies entirely with the authors.

REFERENCES

- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>
- Cohn, N., Kools, E., & Mieth, P. (2012). The TomTom Congestion Index. *19th ITS World Congress*. 19th ITS World Congress. <https://trid.trb.org/view/1280290>
- de Bok, M., Tavasszy, L., & Sebastiaan Thoen. (2022). Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in The Netherlands. *Transport Policy*, 124, 119–127. <https://doi.org/10.1016/j.tranpol.2020.07.010>
- Gardrat, M. (2019). *Méthodologie d'enquête: Le découplage de l'achat et de la récupération des marchandises par les ménages* (p. 115) [Research report]. LAET.

- Hörl, S., & Balac, M. (2021). Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transportation Research Part C: Emerging Technologies*, 130, 103291. <https://doi.org/10.1016/j.trc.2021.103291>
- Hörl, S., & Puchinger, J. (2022). *From synthetic population to parcel demand: Modeling pipeline and case study for last-mile deliveries in Lyon*. Transport Research Arena (TRA) 2022, Lisbon.
- Mucowska, M. (2021). Trends of Environmentally Sustainable Solutions of Urban Last-Mile Deliveries on the E-Commerce Market—A Literature Review. *Sustainability*, 13(11), 11. <https://doi.org/10.3390/su13115894>
- Neghabadi, P. D., Samuel, K. E., & Espinouse, M.-L. (2019). Systematic literature review on city logistics: Overview, classification and analysis. *International Journal of Production Research*, 57(3), 865–887. <https://doi.org/10.1080/00207543.2018.1489153>
- Patella, S. M., Grazieschi, G., Gatta, V., Marcucci, E., & Carrese, S. (2021). The Adoption of Green Vehicles in Last Mile Logistics: A Systematic Review. *Sustainability*, 13(1), 1. <https://doi.org/10.3390/su13010006>
- Sakai, T., Romano Alho, A., Bhavathrathan, B. K., Chiara, G. D., Gopalakrishnan, R., Jing, P., Hyodo, T., Cheah, L., & Ben-Akiva, M. (2020). SimMobility Freight: An agent-based urban freight simulator for evaluating logistics solutions. *Transportation Research Part E: Logistics and Transportation Review*, 141, 102017. <https://doi.org/10.1016/j.tre.2020.102017>
- Toilier, F., Gardrat, M., Routhier, J. L., & Bonnafous, A. (2018). Freight transport modelling in urban areas: The French case of the FRETURB model. *Case Studies on Transport Policy*, 6(4), 753–764. <https://doi.org/10.1016/j.cstp.2018.09.009>

Evaluating real-time information systems on public transport disturbances

Alessio Daniele Marra*¹, Francesco Corman²

¹ Dr., Institute for Transport Planning and Systems, ETH Zurich, Switzerland

² Prof. Dr., Institute for Transport Planning and Systems, ETH Zurich, Switzerland

SHORT SUMMARY

Real-Time Information (RTI) systems are a key component of the management process of public transport disturbances. Smartphone applications, in particular, are becoming a popular means of disseminating information to passengers. Despite the widespread usage of RTI systems, little is known on how accurate those systems are, which information they provide, or which disturbances are not reported.

This work proposes a methodological framework and a set of metrics to evaluate a text-based RTI system, comparing the alerts sent to passengers with the actual disturbances in the network, described by Automatic Vehicle Location data (AVL). A case study is conducted on the RTI system of Zurich to evaluate its performance. The results show high precision in providing correct information, despite only a small percentage of disturbances are reported. Finally, this work proposes recommendations on improving the RTI system analyzed.

Keywords: Real Time Information; Disruptions; Public Transport; AVL data

1. INTRODUCTION

Real time information (RTI) on public transport disturbances can significantly improve the travel experience of passengers. In a literature review, (Brakewood & Watkins, 2019) show the primary positive effects are the decreased waiting time, overall travel time, change in route choice and increased satisfaction with the transport system. RTI can be provided to passengers with different means, such as: a screen at a bus stop, voice alerts, or mobile applications. (Harmony & Gayah, 2017) identified that the preferred option for receiving RTI are mobile applications. Nevertheless, there is little knowledge on the effectiveness of mobile applications and social media to share RTI (Hu et al., 2018; Rahman et al., 2019).

Most of the works in literature exploits surveys or simulation to study how RTI influences passengers, in terms of travel behavior, route choice and waiting time (Akhla et al., 2022; Leng & Corman, 2020; Paulsen et al., 2021). However, according to (Papangelis et al., 2016), there is almost no evidence on RTI requirements for passengers during disruptions. In this sense, they recommend the information should be accurate, timely, and directed to passengers' needs, rather than generic.

Evaluating an RTI system is significantly beneficial for a service provider, to identify the main flaws of the system, and its effectiveness compared to other systems. However, despite the acknowledged importance of RTI systems, their performance in practice remains poorly understood, including their level of accuracy, which disturbances are notified, and which ones are ignored. In fact, a valuable RTI system should not only notify large disruptions, but also small disturbances, which may have a large impact on passengers (Marra & Corman, 2020), if they are not correctly informed (Marra & Corman, 2023).

This work proposes a methodological framework to analyze any text-based RTI system, and a set of metrics to evaluate it, in terms of correctness, amount of information provided and timeliness.

The core idea behind is the comparison of alerts sent to passengers with Automatic Vehicle Location data (AVL), describing all disturbances occurred in the network. The proposed methodology can be used to evaluate the performance of an RTI system, quantifying numerically some of its strengths and weaknesses. We apply the framework in a case study, the RTI system of Zurich, informing passengers on a smartphone application. Finally, we offer suggestions for improving the performance of the examined RTI system, which can also be applied to similar systems.

2. METHODS AND DATA

We analyze an RTI system, comparing the alerts sent to passengers with the disturbances occurred in the network. The comparison shows which disturbances are notified and how accurate is the information provided. Figure 1 shows the methodology to analyze the RTI system of Zurich. The same framework can be used to study any RTI system, adapting it to the information provided in the alert. The data and each step are described in the following sections.

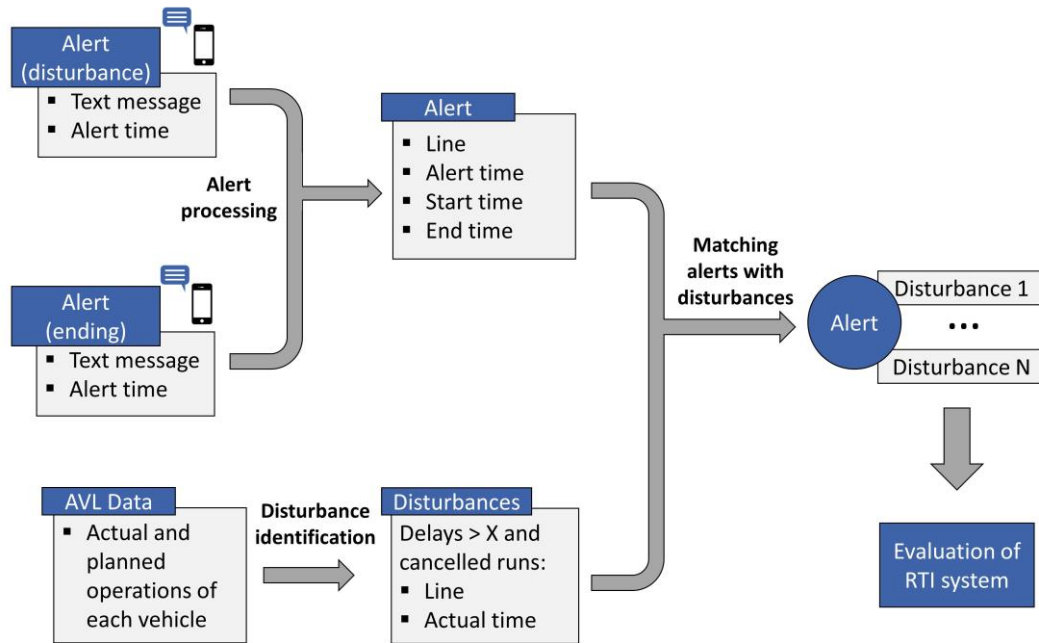


Figure 1: Methodology to analyze Zurich RTI system

Description of the alerts

We study the RTI system of the public transport network of Zurich, Switzerland, named “VBZ-alerto”. The system sends automatic alerts (text messages) on a smartphone app, Telegram, describing disturbances occurred in the city.

We collected 430 alerts between 11.11.2022 and 22.02.2023. Each alert is formed by a text and an “alert-time”, i.e. the time the alert is sent on the app. The alerts can be classified into three types: disturbance alerts (55%), notifying a disturbance; ending alerts (38%), notifying the end of a disturbance, previously notified; other alerts (6%), with unique text.

For a disturbance alert, the text follows the following structure:

“**Lines A, B, C:** description of disruption (unstructured)

dd.mm.yyyy hh:mm – indefinite

detailed description (unstructured)”.

Therefore, each alert contains information on the disturbed lines and the “starting-time”, i.e. the time the disruption started according to the operator. The “ending-time” is indefinite for most of the alerts (90%). However, an ending alert starting with “Resolved” notifies the “ending-time” of a previous disturbance alert. 71% of disturbance alerts have an ending alert. Additional information on the disturbance, like stops involved, are not provided in a standard format, and thus cannot be derived automatically from the text.

In the following analyses, we consider only alerts with the ending specified, in the text or by a following alert. An alert involving more than one line is considered as a different alert for each line. In total, 332 alerts are analysed.

Actual disturbances

We identify all disturbances occurred (and not only the ones notified), from long-term AVL data of Zurich public transport. The AVL data contain the actual and planned times of each vehicle at each stop in Zurich. Therefore, they describe what actually occurred in the network. According to (Zhang et al., 2022), AVL data are the preferred source to detect disturbances, compared to passenger data, incident logs and social media.

The AVL data describe the actual disturbances occurred each day, in the form of delays and cancelled runs. Larger disturbances, such as an entire line cancelled, are described by a set of smaller disturbances, such as a series of cancelled runs. Therefore, in this work, we consider as a disturbance any delay $>D$ and any cancelled run. We choose $D=8$ min, as it is the average headway in Zurich. Each disturbance has an “actual-time”, indicating the time the disturbance occurred. For each run, we consider only the first disturbance. Namely, if a vehicle is delayed at several stops, the first delay is considered as the begin of the disturbance.

Finally, we also analyze lines not functioning for longer time, i.e. at least N consecutive cancelled runs ($N=3$ in our experiments).

Matching alerts with actual disturbances

The quality of an RTI system can be evaluated, comparing the set of alerts sent, with the actual disturbances (described in the AVL data). In fact, matching the alerts with the disturbances occurred in the network, allows to identify which disturbances are notified and which not.

We consider an alert matching a disturbance, if the following conditions hold:

- The disturbance and the alert concern the same line;
- The actual-time of the disturbance lies between the starting-time and ending-time of the alert, including a buffer B (10 min);

Each alert may match multiple disturbances, since, for instance, an alert notifying a cancelled line for an hour corresponds to multiple cancelled runs.

Evaluating the RTI system

We evaluate an RTI system, based on how good the alerts inform about the disturbances. We propose a set of metrics to evaluate the correctness, amount of information provided, and timeliness of the RTI system. The correctness can be measured as for a classification model, identifying which operations are considered disturbances, and therefore should be alerted. The performance can be evaluated as in Table 1. The disturbances matched by an alert are indeed informed to the passengers, and represent the True Positives of the system. All disturbances not matched by an alert are the False Negatives. Disturbances alerted but not available in the data (i.e. disturbances not actually occurred, or errors in the AVL data) are the False Positives. All normal operations (without disturbances and not alerted) are the True Negatives.

Table 1: Correctness of an alert system

	Alert sent	No alert sent
Disturbance in the AVL data	Actual disturbance notified [True Positive]	Actual disturbance not notified [False Negative]
No disturbance in the AVL data	Disturbance notified but not in the data [False Positive]	Normal operations [True Negative]

We measure the performance of the system in terms of precision and recall, as follows:

$$\mathbf{Precision} = \text{True Positive} / (\text{True Positive} + \text{False Positive}) \quad (1)$$

$$\mathbf{Recall} = \text{True Positive} / (\text{True Positive} + \text{False Negative}) \quad (2)$$

Precision and recall measure respectively the correctness of an RTI system and the percentage of disturbances notified to passengers. High precision means the alerts provide correct information to passengers, while low precision means the alerts provide wrong information. High recall indicates most of the disturbances are alerted, while low recall indicates the passengers are not informed about most of the disturbances. We remark each alert matches multiple disturbances; therefore, the precision should be computed as number of correct alerts, while the recall as number of alerted disturbances.

We evaluate the performance also in terms of timeliness of the alerts. For each alert, three time-related information are available on the start of the disturbance: “alert-time”, when the alert is sent; “starting-time”, when the alert says the disturbance started; “actual-time”, when the disturbance actually started, according to the data, i.e. the time of the first disturbance matching the alert. We evaluate the timeliness, based on three metrics:

$$\mathbf{Promptness} = \text{alert time} - \text{starting time} \quad (3)$$

$$\mathbf{Latency} = \text{actual time} - \text{starting time} \quad (4)$$

$$\mathbf{Reactivity} = \text{alert time} - \text{actual time} \quad (5)$$

The promptness represents how late the alert is sent to passengers, compared to the beginning of the disturbance, according to the operator. The latency represents how late the disturbance started, compared to when the alert says it started. The reactivity represents how late the alert is sent, compared to the actual beginning of the disturbance.

3. RESULTS AND DISCUSSION

During the study period, we analyzed 332 alerts, for an average of 2.29 lines per day alerted. In contrast, the actual disturbances are much more, with a total of 52097 disturbed runs, and an average per day of 401 disturbed runs and 48.5 disturbed lines. This large difference is expected, since we are considering both small and large disturbances.

Regarding the correctness of the RTI system of Zurich, we observed a precision of 98% and a recall of 12%. The very high precision shows that the alerts provide correct information, and that when an alert is sent to a passenger, the disturbance can be observed in the AVL data and therefore occurred in reality (if there are no errors in the AVL data). Only 6 alerts did not find a correspondence in the AVL data, which may be due to a wrong alert-time in the text of the alert, a measurement error in the AVL data, or a disturbance with a delay shorter than 8 min. The low recall shows that passengers are informed only of 12% of disturbances (48 runs on average per day). This is

detrimental for passengers, since they are not informed of most of disturbances they encounter during their trip. However, we remark that it is unrealistic for an RTI system to have a very high recall, since it is not possible to inform passengers of all disturbances. In fact, sending just a single alert per disturbed line is equal to an average of 48.5 alerts per day, which may be overwhelming for a passenger.

Among the notified disturbances, 91% are cancelled runs, while 9% are delays. Instead, among the not notified disturbances, 78% are cancelled runs, while 22% are delays. Therefore, passengers are more informed when a line is not running, than when it is delayed.

We also analyzed long-term cancelled lines, i.e. when at least 3 consecutive runs are cancelled. In this case, the recall is higher (30%), showing long-term cancellations are more frequently notified than single cancellations (12%).

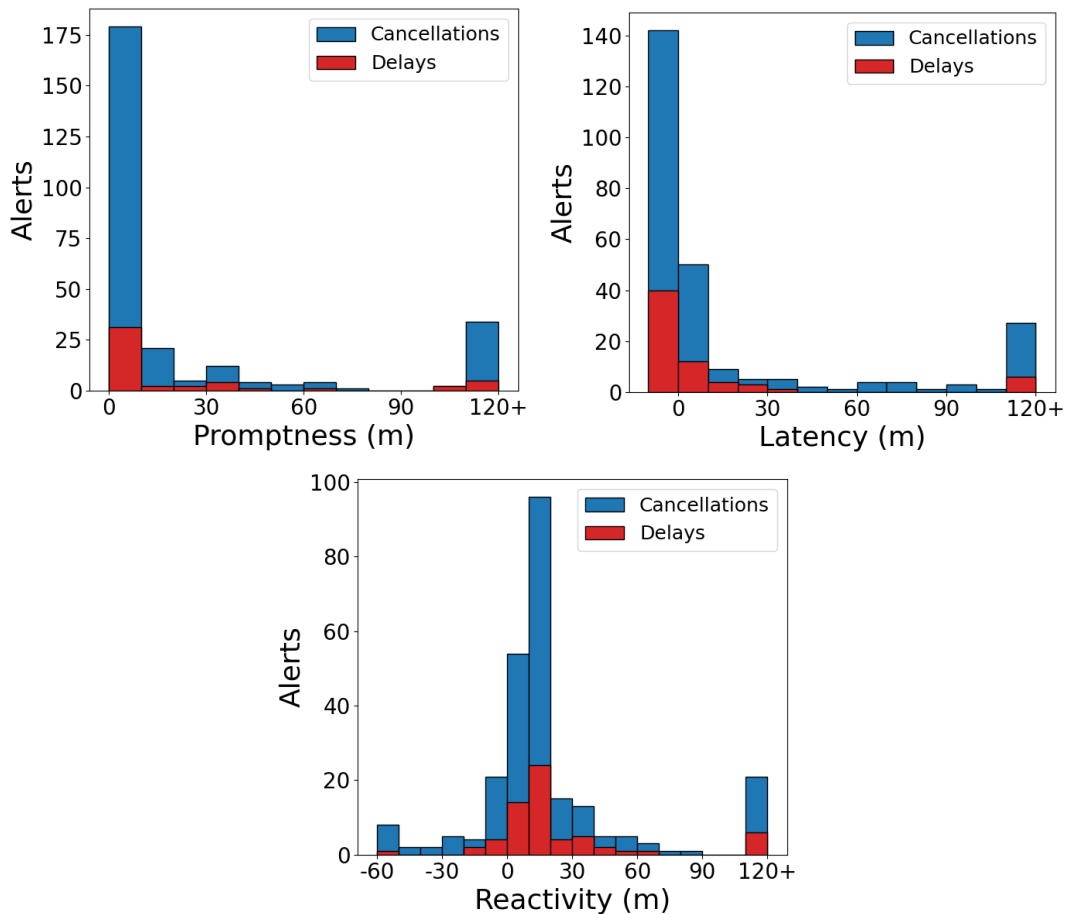


Figure 2: Promptness, Latency and Reactivity for the observed alerts

Figure 2 shows the timeliness of an RTI system based on three metrics, as defined in Section 3. For most of the alerts (73%) the promptness is below 10 minutes. This shows that most of the time the operators inform passengers in less than 10 minutes from the beginning of the disturbance, as written in the alert. This time may be adequate for long disturbances, while not for shorter ones (e.g. small delays), since informing passengers after 10 or more minutes may not be useful for them. For most of the alerts, the latency is between -10 and 10 minutes (77%), showing the information provided on the beginning of the disturbance is correct. However, for many alerts the latency is large (120+ minutes). This is often the case when the actual start is not known, thus it is reported in the early morning (e.g. 06:00). Finally, the reactivity has a wider range, showing the alerts are often sent much before or after the actual beginning of the disturbance.

The analyses and results shown suggest important recommendations to improve the quality of Zurich RTI system, which applies also to similar RTI systems. The disturbances occurring daily in the network are too many to be notified to all passengers indiscriminately. Therefore, personalized alerts are recommended to inform about all occurred disturbances (increasing the recall), without overwhelming the passengers with too many alerts. The Zurich RTI system provides correct information, in terms of disturbed line and time, for most of the alerts. However, further information on the involved stops, vehicles or type of disturbance are not always provided. This makes unclear to passengers if the disturbance is affecting them (maybe a passenger is at a non-disturbed stop). In this sense, enhancing the information quality can improve the passenger's reaction to disturbances and the overall travel experience. From the analysis of timeliness, we identified that the RTI system provides timely and accurate information on the starting time of a disturbance. However, the gap between the beginning of a disturbance and the notification to passengers can be reduced, especially for small disturbances, whose notification should be faster to be effective.

4. CONCLUSIONS

RTI systems are widely acknowledged as key contributors to passengers' travel experience, in case of public transport disturbances. Despite their recognized value, little research has investigated how to evaluate an RSI system, how accurate is the information provided and which disturbances are (or not) notified. This work answers these questions proposing a methodology to evaluate a text-based RTI system. The core idea is the comparison of alerts with AVL data, to identify which disturbances are notified to passengers and which not. Afterwards, a set of metrics is defined to assess the correctness and timeliness of the system. These metrics help to identify the drawbacks of an RTI system and the directions of improvement. We applied the proposed methods in a real test case in Zurich. The results identify that the alerts are highly precise and punctual, but they cover only a small fraction (12%) of all disturbances in the network.

For future work, the analyses can be extended in several directions. The disturbances can be divided into different categories (e.g. small delays, large disruptions). Analyzing how frequently those categories are notified may highlight which disturbances are prioritized and which are dismissed. Furthermore, the importance of different alerts can be estimated based on the effects of the disturbance on passengers, or the travel time saved thanks to the alert.

Finally, we remark this is an on-going study, and the dataset is increasing every day. Therefore, in the future we plan to study a much larger dataset, allowing more detailed analyses.

REFERENCES

- Akhla, A. A. N., Ling, T. C., Shibghatullah, A. S., Mon, C. S., Cherukuri, A. K., Yen, C. L., & Yi, L. C. (2022). Impact of Real-Time Information for Travellers: A Systematic Review. *Journal of Information & Knowledge Management*, 2250065.
- Brakewood, C., & Watkins, K. (2019). A literature review of the passenger benefits of real-time transit information. *Transport Reviews*, 39(3), 327–356.
- Harmony, X. J., & Gayah, V. V. (2017). Evaluation of Real-Time Transit Information Systems: An information demand and supply approach. *International Journal of Transportation Science and Technology*, 6(1), 86–98.
- Hu, M. W., Huang, W. K., & Chen, Y. (2018). Evaluating the Effects of Alleviating Urban Traffic Congestion Using Real-Time Traffic Information on Mobile Devices. In K. W. Chau, I. Y. S. Chan, W. Lu, & C. Webster (Eds.), *Proceedings of the 21st International Symposium on Advancement of Construction Management and Real Estate* (pp. 689–700). Springer.
- Leng, N., & Corman, F. (2020). The role of information availability to passengers in public transport disruptions: An agent-based simulation approach. *Transportation Research Part A: Policy and Practice*, 133, 214–236.
- Marra, A. D., & Corman, F. (2020). From Delay to Disruption: Impact of Service Degradation on Public Transport Networks. *Transportation Research Record*, 2674(10), 886–897.
- Marra, A. D., & Corman, F. (2023). How different network disturbances affect route choice of public transport passengers. A descriptive study based on tracking. *Expert Systems with Applications*, 213, 119083.
- Papangelis, K., Velaga, N. R., Ashmore, F., Sripada, S., Nelson, J. D., & Beecroft, M. (2016). Exploring the rural passenger experience, information needs and decision making during public transport disruption. *Research in Transportation Business & Management*, 18, 57–69.
- Paulsen, M., Rasmussen, T. K., & Nielsen, O. A. (2021). Impacts of real-time information levels in public transport: A large-scale case study using an adaptive passenger path choice model. *Transportation Research Part A: Policy and Practice*, 148, 155–182.
- Rahman, R., Roy, K. C., Abdel-Aty, M., & Hasan, S. (2019). Sharing Real-Time Traffic Information With Travelers Using Twitter: An Analysis of Effectiveness and Information Content. *Frontiers in Built Environment*, 5.
- Zhang, N., Graham, D. J., Bansal, P., & Hörcher, D. (2022). Detecting metro service disruptions via large-scale vehicle location data. *Transportation Research Part C: Emerging Technologies*, 144, 103880.

The Bus Rapid Transit Investment Problem

Rowan Hoogervorst^{*1}, Evelien van der Hurk¹, Philine Schiewe², Anita Schöbel^{3,4},
and Reena Urban³

¹DTU Management, Technical University of Denmark, Kongens Lyngby, 2800, Denmark

²Department of Mathematics and Systems Analysis, Aalto University, Espoo, 02150, Finland

³Department of Mathematics, RPTU Kaiserslautern-Landau, Kaiserslautern, 67663, Germany

⁴Fraunhofer Institute of Industrial Mathematics ITWM, Kaiserslautern, 67663, Germany

SHORT SUMMARY

Bus Rapid Transit (BRT) systems can be of great value to attract passengers towards public transport, as they offer an attractive service at relatively low investment costs. Often, BRT lines are created by giving the bus a dedicated right of way along segments of an existing bus line. This paper focuses on quantifying the trade-off between the number of attracted passengers and the available investment budget when upgrading a line. Motivated by the construction of a new BRT line around Copenhagen, we consider multiple municipalities that invest in the line. We additionally allow restrictions on the number of connected components to be upgraded to enforce connectedness. We suggest two passenger responses to determine the number of attracted passengers and propose an ϵ -constraint based algorithm to enumerate all non-dominated points. Moreover, we perform an extensive experimental evaluation on artificial instances and a case study for the BRT line around Copenhagen.

Keywords: Bus Rapid Transit, Network Design, Operations Research Applications, Public Transport

1 INTRODUCTION

Increasing the modal share of public transport is seen as one of the paths to reducing greenhouse gas emissions, even when considering the electrification of private cars (Messerli et al., 2019). Bus Rapid Transit (BRT) systems can contribute to this goal, as BRT lines provide a fast and reliable service to passengers due to having a dedicated right of way for buses along a large share of their route. However, while BRT investment costs are lower than for rail-based alternatives, these investments are still substantial for local authorities both in terms of cost and usage of city space that cannot be used for other purposes. Hence, careful planning is needed to decide which segments of the existing line should be upgraded to a BRT standard.

In this paper, we focus on quantifying the trade-off between the number of attracted passengers and the investment budget by determining the optimal sets of segment upgrades. Motivated by a new BRT line being built in the Greater Copenhagen area, we consider a problem setting in which multiple municipalities are responsible for different segments of the line and each municipality has a budget limit. To prevent frequent switching between upgraded and non-upgraded segments due to fragmented investments, which could reduce reliability and thus deter passengers, we additionally allow restricting the number of upgraded connected components on the line. We refer to this problem as the BRT investment problem.

The BRT investment problem relates to the well-studied network design problem for public transport (Laporte et al., 2000; Laporte & Mesa, 2019). While the network design problem generally focuses on constructing a network from scratch, numerous papers also look at the upgrading of existing public transport networks. Particularly relevant for us are those papers looking at the allocation of dedicated bus lines within existing transport networks, many of which focus on the trade-off between the benefits for public transport passengers and the congestion on the road network (Khoo et al., 2014; Bayrak & Guler, 2018; Tsitsokas et al., 2021). Another addition to the standard network design problem that is relevant for our application in the Greater Copenhagen region is the inclusion of multiple investing parties, which was studied by Wang & Zhang (2017) within a game-theoretical setting. Moreover, the underlying mathematical structure of the BRT investment problem is similar to the more general network improvement problem, which consists

of choosing edges in a network to be upgraded while minimizing costs or satisfying a budget constraint (Krumke et al., 1998; Zhang et al., 2004; Murawski & Church, 2009). Compared to existing literature, our work distinguishes itself by studying the combination of a trade-off between the number of attracted passengers and the investment budget, the inclusion of multiple investing parties, and the inclusion of constraints ensuring connectedness of the upgraded segments.

This paper builds on the work of our earlier conference paper (Hoogervorst et al., 2022), in which we looked at the single-objective problem of upgrading bus line segments under a budget constraint and without a constraint on the number of upgraded components. In this paper, we instead propose a new bi-objective mixed-integer programming model to solve the BRT investment problem that allows us to construct the Pareto curve between the total investment budget and the number of attracted passengers. We do so under two possible passenger responses, one corresponding to a linear relation and the other to a threshold relation between segment upgrades and the attracted share of passengers. We show how the set of all non-dominated points can be found under these passenger responses and test our proposed algorithm in a numerical study for both artificial instances and a case study for the BRT line in Greater Copenhagen.

2 METHODOLOGY

In this section, we first formally define the problem and afterwards describe the used solution methods.

Problem Definition

We consider an existing bus line given by a linear graph (V, E) , where $V = \{1, \dots, n\}$ for $n \in \mathbb{N}_{\geq 1}$ denotes the set of stations and $E = \{e_i = \{i, i + 1\} : i \in \{1, \dots, n - 1\}\}$ the set of segments between the stations. For each edge $e \in E$, we know the cost $c_e \in \mathbb{R}_{>0}$ for upgrading the edge as well as the improvement, i.e., improvement in travel time, $u_e \in \mathbb{R}_{>0}$ that is realized when the edge is upgraded. Moreover, let $D \subseteq \{(i, j) : i, j \in V, i < j\}$ be the set of origin-destination (OD) pairs for the line, where OD-pair $d = (i, j) \in D$ has the unique path $W_d = \{e_k : k \in \{i, i + 1, \dots, j - 1\}\}$ along the line. For each OD-pair $d \in D$, we additionally know the number of potential passengers a_d that are attracted when all edges in the path W_d are upgraded.

The set of municipalities that are investing in the BRT line is given by M . For each municipality, we know the set of consecutive edges $E_m \subseteq E$ that lie within the municipality. We will assume that these sets of edges of the different municipalities are pairwise disjoint, i.e., $W_k \cap W_l = \emptyset$ for $k, l \in M, k \neq l$. Moreover, we know the budget share b_m that is allocated to each municipality, i.e., each municipality gets budget $b_m B$ when considering some total budget B . Lastly, to prevent buses from switching too often between upgraded and non-upgraded ones, we enforce a maximum number of BRT components of Z .

While the number of potential passengers attracted is given for each OD-pair $d \in D$ when all edges in W_d are upgraded, it is beforehand unclear how passengers react to partial upgrading of the edges in W_d . We consider two different passenger responses $p_d(F)$ to a set of upgrades $F \subseteq E$:

- The LINEAR response to upgrades

$$p_d(F) := \frac{\sum_{e \in F \cap W_d} u_e}{\sum_{e' \in W_d} u_{e'}} \cdot a_d,$$

in which the number of passengers scales linearly with the amount of improvement realized.

- The MINIMPROV response to upgrades

$$p_d(F) := \begin{cases} a_d & \text{if } L_d \leq \sum_{e \in F \cap W_d} u_e, \\ 0 & \text{otherwise,} \end{cases}$$

in which all the potential passengers are only attracted when a minimum improvement of L_d is achieved.

Note that the MINIMPROV response resembles a shortest path based route and mode choice, where passengers only switch to the BRT line in case the upgrade is large enough to make it their option with the shortest travel time.

The BRT investment problem then becomes to find all the non-dominated solutions (F, B) , with F the set of upgraded edges and B the total investment budget, that solve:

$$\max \sum_{d \in D} p_d(F) \quad (1)$$

$$\min B \quad (2)$$

$$s.t. \sum_{e \in E_m \cap F} c_e \leq b_m B \quad m \in M, \quad (3)$$

$$G[F] \text{ has at most } Z \text{ connected components,} \quad (4)$$

$$F \subseteq E, \quad (5)$$

$$B \in \mathbb{R}. \quad (6)$$

The first objective (1) maximizes the number of attracted passengers, while the second objective (2) minimizes the investment budget. Constraints (3) enforce the budget limit for each municipality. Moreover, constraints (4) enforce the maximum number of BRT components. Here, $G[F]$ is the graph induced by the set of edges F , i.e., the graph obtained after deleting all edges from G that are not contained in F .

Solution Methodology

Formulation (1) – (6) can be transformed into a bi-objective mixed-integer linear programming (MILP) model through introducing variables for all $e \in E$ that depict if a segment is upgraded and, in the case of the MINIMPROV response, variables y_d for all $d \in D$ that depict if the minimum improvement is realized for an OD-pair. For example, this leads to the following formulation for the MINIMPROV objective:

$$\max \sum_{d \in D} a_d y_d \quad (7)$$

$$\min B \quad (8)$$

$$s.t. L_d y_d \leq \sum_{e \in W_d} u_e x_e \quad d \in D, \quad (9)$$

$$\sum_{e \in E_m} c_e x_e \leq b_m B \quad m \in M, \quad (10)$$

$$x_{e_i} - x_{e_{i+1}} \leq z_i, \quad i \in \{1, \dots, n-2\}, \quad (11)$$

$$x_{e_{i+1}} - x_{e_i} \leq z_i, \quad i \in \{1, \dots, n-2\}, \quad (12)$$

$$x_{e_1} + \sum_{i=1}^{n-2} z_i + x_{e_{n-1}} \leq 2Z, \quad (13)$$

$$x_e \in \{0, 1\} \quad e \in E, \quad (14)$$

$$z_i \in \{0, 1\} \quad i \in \{1, \dots, n-2\}, \quad (15)$$

$$y_d \in \{0, 1\} \quad d \in D, \quad (16)$$

$$B \in \mathbb{R}. \quad (17)$$

The objectives (7) and (8) maximize the number of attracted passengers and minimize the investment budget, respectively. Constraints (9) determine if the minimum improvement for an OD-pair is realized. The budget limit is enforced for each municipality by constraints (10). Moreover, constraints (11) – (13) enforce the maximum number of BRT components through counting the number of switches on the line between upgraded and non-upgraded segments. A bi-objective MILP model can be obtained for the LINEAR passenger response in a similar way.

We use the ϵ -constraint method to find the set of non-dominated solutions, i.e., solutions on the Pareto curve, for the proposed bi-objective programming problems. The used algorithm is given in Algorithm 1, which is an adaptation of the algorithm proposed by Bérubé et al. (2009). The idea of the algorithm is to iteratively compute all non-dominated points by solving the single-objective version of the BRT investment problem for a fixed total budget B and to decrease B in each step by a value that is small enough not to cut-off any non-dominated solution. In particular, we can prove that this algorithm generates the set of all non-dominated points on the Pareto curve.

Algorithm 1 Computing the non-dominated points for the BRT investment problem

- 1: **Input:** instance I of the BRT investment problem.
 - 2: **Output:** set Γ of all non-dominated points.
 - 3: As start values set
 - 4: $\Gamma \leftarrow \emptyset$,
 - 5: $B \leftarrow \max_{m \in M} \left\{ \frac{1}{b_m} \cdot \sum_{e \in E_m} c_e \right\}$,
 - 6: $v^* \leftarrow \max_{m \in M} \left\{ \frac{1}{b_m} \cdot \sum_{e \in E_m} c_e \right\}$,
 - 7: $p^* \leftarrow \sum_{d \in D} a_d$.
 - 8: **while** $B \geq 0$ **do**
 - 9: Solve instance I with budget B . Let F be an optimal solution, \bar{p} be the optimal objective value.
 - 10: Compute the minimum budget \bar{v} such that F remains feasible.
 - 11: Compute step width δ .
 - 12: **if** $\bar{p} < p^*$ **then**
 - 13: Set $\Gamma \leftarrow \Gamma \cup \{(p^*, v^*)\}$.
 - 14: Set $p^* \leftarrow \bar{p}$.
 - 15: **end if**
 - 16: Set $v^* \leftarrow \bar{v}$.
 - 17: Set $B \leftarrow \bar{v} - \delta$.
 - 18: **end while**
 - 19: Set $\Gamma \leftarrow \Gamma \cup \{(p^*, v^*)\}$.
 - 20: **return** Γ
-

3 RESULTS AND DISCUSSION

We perform computational experiments for both a set of artificial instances, based on those introduced in Hoogervorst et al. (2022), and on instances from the proposed BRT line in Greater Copenhagen that motivated our study. The artificial instances differ with respect to the passenger demand over the OD-pairs and the upgrade costs of the segments. The instances for the Greater Copenhagen BRT line are instead based on five proposed line alternatives for the BRT, depicted in Figure 1, and consider different ways of distributing the total budget over the municipalities.

Results Artificial Instances

The obtained Pareto plots for the artificial instances are given in Figure 2 for the setting of a single municipality that can invest in all edges. The different columns in the figure indicate the different passenger demand distributions, where each OD-pair has equal demand (*EVEN*), passengers mostly travel to the closest large station (*CENTER*), and passengers mainly travel between the two end-stations (*END*), respectively. The rows instead indicate the different cost patterns, where all segments have equal upgrade cost (*UNIT*), edges towards the middle are most expensive to upgrade (*MIDDLE*), and edges towards the ends are most expensive to upgrade (*ENDS*), respectively.

The Pareto plots in Figure 2 show that there is a noticeably different trade-off between attracted passengers and investment budget for the two passenger responses. For passenger response *LINEAR*, we obtain a mostly concave shape for all demand and cost patterns, where the first investments generate the largest number of new passengers. Instead, the shape of the Pareto curves is more variable over the demand and cost patterns for the *MINIMPROV* response. In particular, we can see a clear jump in the Pareto plots for the *MINIMPROV* response for the *END* demand pattern, which can be explained by the minimum improvement threshold that needs to be reached for attracting the large number of passengers traveling over the whole line in this demand pattern.

The Pareto plots also allow us to obtain insight into the effect of restricting the number of connected components. Restricting the BRT line to consist of a single component leads to a clear reduction in the number of passengers attracted, especially for the *CENTER* and *END* demand patterns. The reduction is significantly smaller when allowing at least two components, where especially the Pareto plots for allowing three components lie close to the ones where no restriction on the number

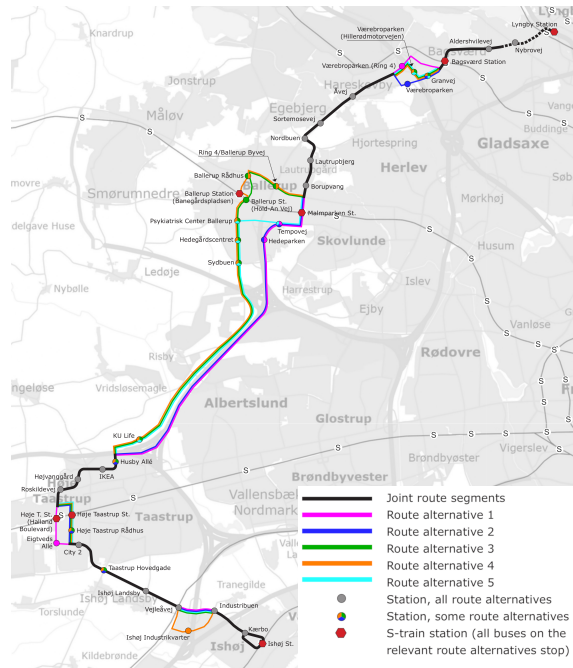


Figure 1: Route alternatives for a new BRT line in Greater Copenhagen. Adapted from Vejdirektoratet et al. (2022).

of components is enforced.

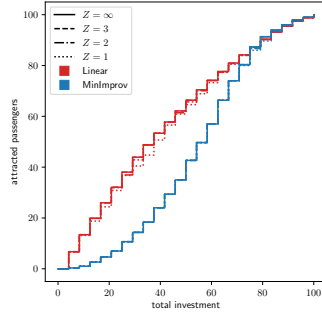
Results BRT Line Greater Copenhagen

Pareto plots for the instances based on the Greater Copenhagen BRT line are given in Figure 3. To evaluate the impact of multiple investing parties, these Pareto plots have been split into the case with multiple investing municipalities (*MIM*) and a case with a single investing party (*SOC*) that can spend the whole investment budget B . For the *MIM* case, the total budget is split both according to the number of passengers in a municipality (*pass*) and the costs of the edges in a municipality (*cost*). Note that no restriction is enforced on the number of BRT components.

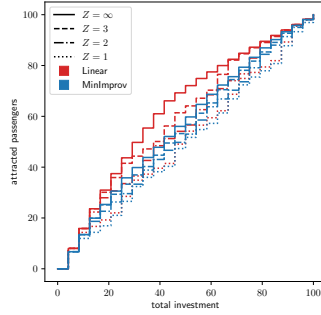
Figure 3 shows that there is not a universal ordering of the line alternatives but that the best alternative depends on the investment level. For the *SOC* case, line alternatives 4 and 5, e.g., lead to the highest number of passengers for higher investment levels under both passenger responses. On the other hand, line alternatives 1 and 2 perform well for low investment levels, in particular for the MINIMPROV response. When comparing the *SOC* and *MIM* cases, it can additionally be seen that the introduction of a budget per municipality leads to a clear reduction in the number of attracted passengers. This reduction seems to be strongest when passengers behave according to the MINIMPROV response, for which we again see a more convex shape of the Pareto curve when moving to the *MIM* case. Lastly, a comparison between the two budget assignments shows that the *cost* budget distribution often seems to lead to the highest number of passengers for high investment levels, while *pass* often performs well for lower investment level.

4 CONCLUSIONS

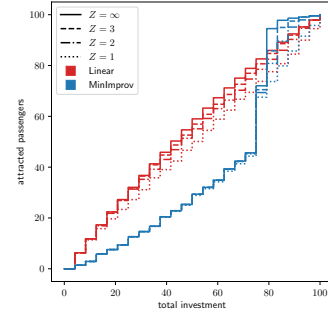
In this paper, we studied the BRT investment problem, which is focused on finding the trade-off between attracted passengers and investment budget when upgrading an existing bus line to a BRT line. We formulated the problem formally and suggested an ϵ -constraint based algorithm to enumerate the full set of non-dominated points. The algorithm was tested on both artificial instances and instances coming from a BRT line case study in Greater Copenhagen. Our artificial results give insight into the trade-off between the number of passengers and investment budget for different instance settings and show that the trade-off clearly depends on the assumed passenger response to upgrades. Moreover, they show that especially the limitation to a single BRT component leads to fewer passengers, while the impact is significantly lower if more components are allowed. Our results for the Greater Copenhagen case study show how the best line alternative can



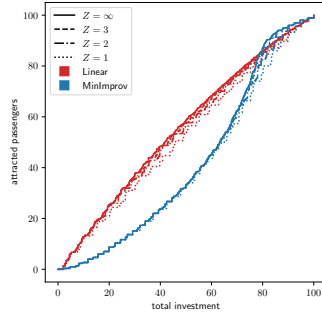
(a) UNIT, EVEN



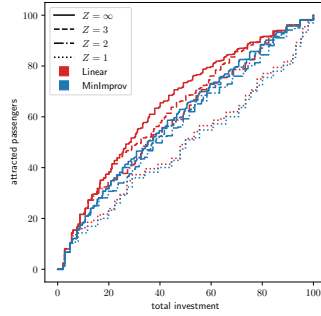
(b) UNIT, CENTER



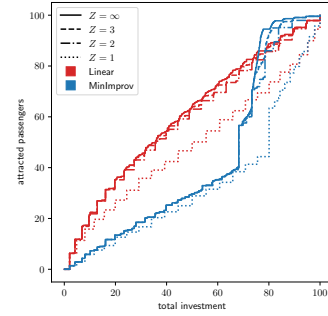
(c) UNIT, END



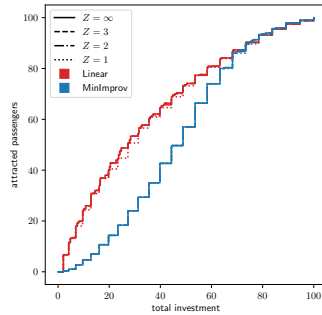
(d) MIDDLE, EVEN



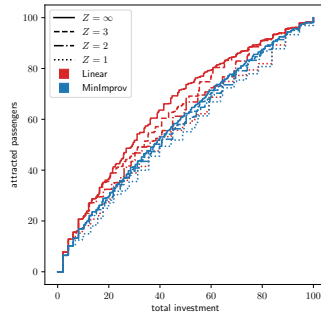
(e) MIDDLE, CENTER



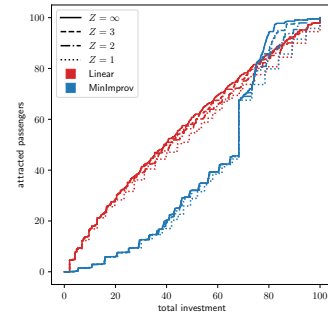
(f) MIDDLE, END



(g) ENDS, EVEN

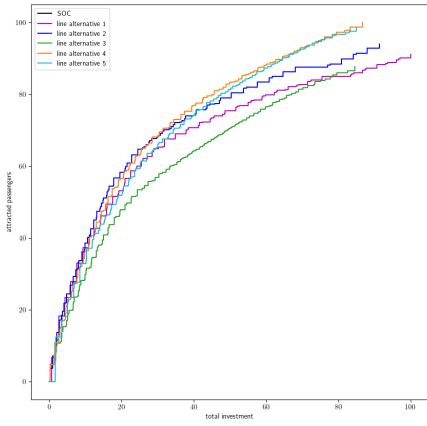


(h) ENDS, CENTER

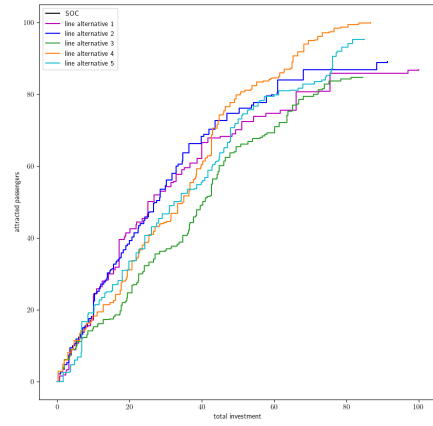


(i) ENDS, END

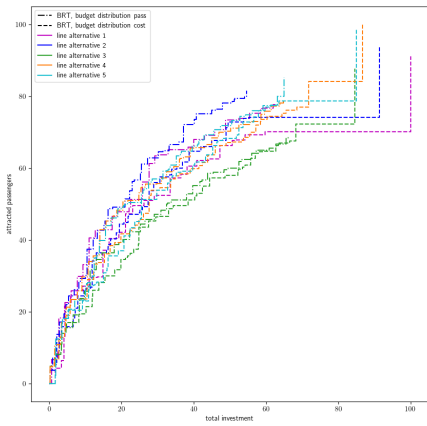
Figure 2: Non-dominated points for the LINEAR (red) and MINIMPROV (blue) passenger response. Solid lines represent the case $Z = \infty$, dashed lines $Z = 3$, dashed-dotted lines $Z = 2$ and dotted lines $Z = 1$. Attracted passengers and total investment are given as a percentage of the total number of potential passengers and costs for upgrading all segments, respectively.



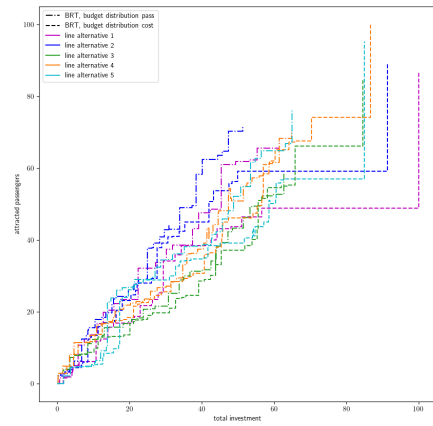
(a) SOC: LINEAR



(b) SOC: MINIMPROV



(c) MIM: LINEAR



(d) MIM: MINIMPROV

Figure 3: Comparing investment costs and attracted passengers for the different route alternatives for $Z = \infty$.

differ per investment level and show the impact that having multiple investing municipalities has on the number of attracted passengers. The latter is shown to depend on the passenger response, where the impact is strongest in the case of the threshold-based passenger response MINIMPROV.

ACKNOWLEDGEMENTS

This work was supported by the European Union’s Horizon 2020 research and innovation programme [Grant 875022], the Federal Ministry of Education and Research [Project 01UV2152B], and Innovationsfonden [0205-00002B] under the project sEAMless SustaInable EveRyday urban mobility (EASIER) as well as by DFG under SCHO 1140/8-2. Moreover, we would like to thank Region H [0205-00005B] and Movia for providing insight into the planning process of the BRT system and for the provision of data.

REFERENCES

- Bayrak, M., & Guler, S. I. (2018). Optimizing bus lane placement on networks while accounting for queue spillbacks. In *2018 21st international conference on intelligent transportation systems (itsc)* (p. 920-925). doi: 10.1109/ITSC.2018.8569375
- Bérubé, J.-F., Gendreau, M., & Potvin, J.-Y. (2009, April). An exact ϵ -constraint method for bi-objective combinatorial optimization problems: Application to the Traveling Salesman Problem with Profits. *European Journal of Operational Research*, *194*(1), 39–50. doi: 10.1016/j.ejor.2007.12.014
- Hoogervorst, R., van der Hurk, E., Schiewe, P., Schöbel, A., & Urban, R. (2022). The Edge Investment Problem: Upgrading Transit Line Segments with Multiple Investing Parties. In M. D’Emidio & N. Lindner (Eds.), *22nd Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2022)* (Vol. 106, pp. 9:1–9:19). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi: 10.4230/OASICS.ATMOS.2022.9
- Khoo, H. L., Teoh, L. E., & Meng, Q. (2014). A bi-objective optimization approach for exclusive bus lane selection and scheduling design. *Engineering Optimization*, *46*(7), 987 – 1007. doi: 10.1080/0305215X.2013.812728
- Krumke, S. O., Marathe, M. V., Noltemeier, H., Ravi, R., & Ravi, S. S. (1998, September). Approximation Algorithms for Certain Network Improvement Problems. *Journal of Combinatorial Optimization*, *2*(3), 257–288. doi: 10.1023/A:1009798010579
- Laporte, G., & Mesa, J. A. (2019). The Design of Rapid Transit Networks. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location science* (pp. 687–703). Cham: Springer International Publishing. doi: 10.1007/978-3-030-32177-2_24
- Laporte, G., Mesa, J. A., & Ortega, F. A. (2000, April). Optimization methods for the planning of rapid transit systems. *European Journal of Operational Research*, *122*(1), 1–10. doi: 10.1016/S0377-2217(99)00016-8
- Messerli, P., Murniningtyas, E., Eloundou-Enyegue, P., Foli, E. G., Furman, E., Glassman, A., ... others (2019). *Global sustainable development report 2019: the future is now—science for achieving sustainable development* (Tech. Rep.). New York: United Nations.
- Murawski, L., & Church, R. L. (2009, June). Improving accessibility to rural health services: The maximal covering network improvement problem. *Socio-Economic Planning Sciences*, *43*(2), 102–110. doi: 10.1016/j.seps.2008.02.012
- Tsitsokas, D., Kouvelas, A., & Geroliminis, N. (2021). Modeling and optimization of dedicated bus lanes space allocation in large networks with dynamic congestion. *Transportation Research Part C: Emerging Technologies*, *127*, 103082. doi: <https://doi.org/10.1016/j.trc.2021.103082>

Vejdirektoratet, Rambøll, & MoeTetraplan. (2022). *BRT i Ring 4-korridoren – forberedende analyse fra Ishøj station til kommunegrænsen til Lyngby-Taarbæk kommune* (Tech. Rep.). Vejdirektoratet. Retrieved from <https://dagsordener.gladsaxe.dk/vis/pdf/bilag/5398c71e-b824-48ac-9058-6e0fe75a2a5c> (In Danish)

Wang, H., & Zhang, X. (2017, February). Game theoretical transportation network design among multiple regions. *Annals of Operations Research*, *249*(1), 97–117. doi: 10.1007/s10479-014-1700-9

Zhang, J. Z., Yang, X. G., & Cai, M. C. (2004, March). A Network Improvement Problem Under Different Norms. *Computational Optimization and Applications*, *27*(3), 305–319. doi: 10.1023/B:COAP.0000013061.17529.79

Carpooling with Transfers and Travel Time Uncertainty

Patrick Stokkink*¹, André de Palma², and Nikolas Geroliminis¹

¹Ecole Polytechnique Fédérale de Lausanne (EPFL), Urban Transport Systems Laboratory (LUTS), Switzerland

²THEMA, CY Cergy Paris University, France

SHORT SUMMARY

Carpooling is known to have lower CO₂ emissions compared to driving individually. One of the limitations of carpooling is that matched drivers and passengers need to have similar itineraries, or their generalized costs will be high. By allowing a single transfer at a designated transfer hub, their itineraries need to be only partially similar. We allow for transfers within the carpooling system and between carpooling and public transport. Thereby, we include travel time uncertainty to evaluate its effect on carpooling with transfers. We model the ride-matching problem with transfers and travel time uncertainty as a two-stage stochastic programming problem. The results indicate that a single transfer hub can already reduce the average generalized cost of passengers by 15%. When travel times are uncertain, commuters tend to find a match that performs relatively well in every traffic situation, rather than one that performs well for only one scenario.

Keywords: Carpooling, Ride-Matching, Ride-Sharing, Transfers, Uncertain Travel Time

1 INTRODUCTION

Transport accounts for a large share of global CO₂ emissions. According to IEA (2022) cars and buses for passenger transport constitute 45.1% of transport emissions, including freight transport. Statistics gathered by the Center for Sustainable Systems, University of Michigan (2021) show that in 2019 the average car occupancy in the United States was 1.5 and in 2017 24% of the U.S. households had 3 or more vehicles. Carpooling as an alternative to traveling alone by car is known to reduce CO₂ emissions directly and a well-functioning carpooling system is expected to reduce car ownership in the long term.

One of the main limitations of direct carpooling is that a pairing of drivers and passengers needs to be found with similar itineraries (in space and time). By allowing transfers, a larger set of potential matches is available for drivers and passengers since the itineraries only need to be partially similar, as they spend only a part of their trip together. Thereby, passengers are allowed to transfer to and from public transport. A graphic illustration is displayed in Figure 1. Carpooling or ride-matching models with transfers have been considered before by, among others, Herbawi & Weber (2012); Masoud & Jayakrishnan (2017); Huang et al. (2018); Lu et al. (2020).

Despite their benefit, transfers may impose additional difficulties when travel time is uncertain. Passengers and drivers may carry on their first-leg delays to the second leg, thereby influencing their match, or they may fully miss their connection. In the presence of uncertainty, transfers can make carpooling with transfers less appealing due to their effect on tardiness and uncomfortable and unanticipated waiting times. Long et al. (2018) consider a bi-objective ride-sharing-matching model under travel-time uncertainty. They consider delay and schedule delay penalties that may change according to this uncertainty.

We consider a matching framework for carpooling with transfers and uncertainty in travel time. We allow for transfers within the carpooling system and between carpooling and public transport. By considering uncertainty in travel time, schedule delay penalties of potential matches can depend on this uncertainty, similar to Long et al. (2018). In our problem, potential matches may be infeasible for some uncertain scenarios which can therefore affect the optimal matching. We model the matching problem with transfers and public transport as a deterministic integer programming problem. We extend this model to a two-stage stochastic programming problem where travel time

uncertainty forms the division between first- and second-stage decisions. Matches on the first leg of the trip are made under uncertainty, whereas second-leg matches are made afterward when information on travel time is gathered.

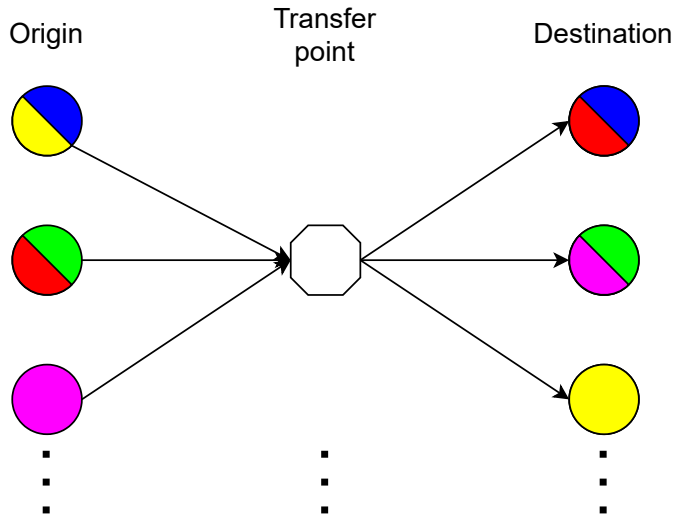


Figure 1: Graphic illustration of transfers

2 METHODOLOGY

We consider a matching problem of carpooling drivers and passengers. Every individual has a specific value of time, scheduling preferences, and desired arrival time. Unlike ride-hailing services, drivers are relatively inflexible in the system we consider. Drivers are only willing to perform pickups at their own origin or at a dedicated transfer hub along their route and are only willing to perform drop-offs at their own destination or at a dedicated transfer hub along their route. Thereby, drivers determine the departure time only based on their own schedule delay preferences and ignore those of the passengers. The reason for this is that in a complex system where drivers take multiple passengers and passengers take multiple drivers, coordinating jointly optimal departure times can be extremely difficult both theoretically and in practice. Contrary to the common carpooling approach where passengers spend their full trip with a single driver, we allow passengers to transfer at designated transfer hubs. These transfer hubs have connections to public transport services and allow for transfers between two carpooling drivers. We only allow one transfer to limit the discomfort of transfers and include the inconvenience of transferring and waiting in the generalized cost formulation.

Deterministic Formulation

The deterministic matching approach is based on a set of predefined passenger paths. Let I be the set of passengers, J the set of drivers, and H the set of transfer hubs. Given that the possible number of matches for passengers is polynomial, we can generate all possible paths in advance. We let the drivers set the departure times, such that the costs of passenger paths are independent of each other. Then we only need to consider that drivers may take multiple passengers at the same time, but only on the same route, and that they can pick up new passengers at the transfer point.

Every individual has an origin o_i , a destination d_i and a desired arrival time t_i^* . Let K be the set of passenger paths and let $e_{ik} = 1$ if passenger path k corresponds to passenger i , and 0 otherwise. The cost of passenger path k is denoted by c_k , which only contains the cost for passengers. By definition, drivers do not incur any scheduling delay costs nor make a detour. Therefore, we assume they are fully compensated for the inconvenience of sharing their car and their costs are not included in the objective function. Let decision variable $x_k = 1$ if passenger path k is chosen and 0 otherwise. Let q_j be the capacity of driver j , that is, the number of passengers driver j is able to transport at the same time. We distinguish between the following three kinds of trips:

- Direct trip: $a_{jk}^0 = 1$ if driver j contributes to passenger path k through a direct trip.

- First leg of indirect trip: $a_{jk}^{1h} = 1$ if driver j contributes to passenger path k through a first-leg trip to transfer hub h .
- Second leg of indirect trip: $a_{jk}^{2h} = 1$ if driver j contributes to passenger path k through a second-leg trip from transfer hub h .

We use decision variable y_{jh} to define through which transfer hub driver j is going. This allows to formulate the deterministic matching problem as follows:

$$\text{(P1) minimize } \sum_{k \in K} c_k x_k \quad (1a)$$

$$\sum_{k \in K} e_{ik} x_k = 1 \quad \forall i \in I \quad (1b)$$

$$\sum_{k \in K} a_{jk}^0 x_k \leq q_j \left(1 - \sum_{h \in H} y_{jh} \right) \quad \forall j \in J \quad (1c)$$

$$\sum_{k \in K} a_{jk}^{1h} x_k \leq q_j y_{jh} \quad \forall j \in J, h \in H \quad (1d)$$

$$\sum_{k \in K} a_{jk}^{2h} x_k \leq q_j y_{jh} \quad \forall j \in J, h \in H \quad (1e)$$

$$\sum_{h \in H} y_{jh} \leq 1 \quad \forall j \in J \quad (1f)$$

$$x_k \in B \quad \forall k \in K \quad (1g)$$

$$y_{jh} \in B \quad \forall j \in J, h \in H \quad (1h)$$

The objective (1a) is to minimize the cost of all matches. Every passenger needs to be matched to exactly one driver, which is enforced by Constraints (1b). Feasibility of the solution from the perspective of a driver is enforced through Constraints (1c) - (1e). The feasibility of the solution from the perspective of a passenger is enforced directly on the set of paths K . That is, the set K only contains paths that are feasible for a passenger. On every leg, a driver $j \in J$ may have at most q_j passengers in their car, which is enforced jointly by Constraints (1c), (1d) and (1e). A driver may either serve passengers directly from their origin to their destination or through a transfer point, but not both. Constraints (1f) ensure that a driver only makes a stop at one transfer point.

In the remainder of this section, we discuss in detail the three types of paths that we consider and the corresponding parameter values in **P1**. For this, we let α be the value of time spent in a car, β the penalty for every unit of time an individual is early, and γ the penalty for every unit of time an individual is late. Waiting time is penalized by α^{wait} and the value of time spent in public transport is defined as α^{pt} . Travel time between o and d is defined as $tt(o, d)$. For the sake of notation, these parameters are all homogeneous, but the formulation allows for heterogeneous parameter values.

Public Transport Paths

Every passenger $i \in I$ has the option to take public transport instead of carpooling. Public transport has a fixed cost per unit of time such that $c_k = \alpha^{pt} tt(o_i, d_i)$.

Direct Carpool Paths

For every passenger $i \in I$, a direct match can be found with a driver $j \in J$ if $o_i = o_j$ and $d_i = d_j$. As the driver selects the departure time to minimize her own cost, the arrival time at the final destination is equal to the desired arrival time of the driver, possibly imposing schedule delay costs on the passenger. For a match between $i \in I$ and $j \in J$, $e_{ik} = 1$, $a_{jk}^0 = 1$ and all other parameters are equal to 0. The cost of this direct match are as follows:

$$c_k = \alpha tt(o_i, d_i) + \beta (t_i^* - t_j^*)^+ + \gamma (t_j^* - t_i^*)^+ \quad (2)$$

Indirect Carpool Paths

For the sake of notation, we define $t_j^*(h)$ as the desired arrival time of driver j at transfer hub h if he travels through that hub. This is simply computed as $t_j^*(h) = t_j^* - tt(h, d_j)$.

We first consider the path where the passenger takes public transport on one of the two legs.

For a path where only the first leg is a carpooling leg, if passenger $i \in I$ and driver $j \in J$ are matched with a transfer at hub $h \in H$, they must share their origin $o_i = o_j$ and hub h must be on the path of driver j . For a path where only the second leg is a carpooling leg, if passenger $i \in I$ and driver $j \in J$ are matched with a transfer at hub $h \in H$, they must share their destination $d_i = d_j$ and hub h must be on the path of driver j . Given the fixed cost of public transport and the fact that we do not consider a schedule for public transport, the cost for first-leg or second-leg carpooling are highly similar and given as follows:

$$c_k = \alpha^{pt} tt(h, d_i) + \alpha tt(o_i, h) + \beta [t_i^* - t_j^*(h) - tt(h, d_i)]^+ + \gamma [t_j^*(h) + tt(h, d_i) - t_i^*]^+ \quad (3)$$

$$c_k = \alpha^{pt} tt(o_i, h) + \alpha tt(h, d_i) + \beta [t_i^* - t_j^*(h) - tt(h, d_i)]^+ + \gamma [t_j^*(h) + tt(h, d_i) - t_i^*]^+ \quad (4)$$

For paths that consist of two carpooling legs, we consider a passenger $i \in I$ and two drivers $j_1, j_2 \in J$ where j_1 takes i on the first leg and j_2 takes i on the second leg with a transfer at transfer hub h . Similar to before, this is only feasible if $o_i = o_{j_1}$, $d_i = d_{j_2}$ and h is both on the path of j_1 and j_2 . Thereby, $t_{j_1}^*(h) \leq t_{j_2}^*(h)$ to ensure that the passenger is dropped off at the transfer hub before the scheduled pickup. The cost for the passenger is then defined as follows:

$$c_k = \alpha [tt(o_i, h) + tt(h, d_i)] + \alpha^{wait} [t_{j_2}^*(h) - t_{j_1}^*(h)] + \beta [t_i^* - t_{j_2}^*(h) - tt(h, d_i)]^+ + \gamma [t_{j_2}^*(h) + tt(h, d_i) - t_i^*]^+ \quad (5)$$

Stochastic Formulation

To allow for uncertainty in travel times, we adapt our formulation to a two-stage stochastic programming problem. The matching is determined a-priori but may be adapted based on the observed state of the system (i.e. the travel times). The first leg of every driver and passenger is fixed and cannot be altered after observing the state. This can be seen as a contract between the driver and the passenger. The second leg, however, may be changed after observing the state of the system. We assume the state is observed after the first leg has been fixed (i.e., the contract has been negotiated) but before the second leg commenced. With modern technologies, commuters are aware of traffic conditions during or shortly before their trip. Let Ω be the uncertainty set and $\omega \in \Omega$ a realization of the uncertain travel times. We also refer to such a realization of the uncertain travel times as a scenario.

All variables and parameters are altered to be dependent on the scenario ω . This means that in stead of x_k we use $x_k(\omega)$ and in stead of y_{jh} we use $y_{jh}(\omega)$. In addition to this, the cost of path k also depends on the scenario as it influences travel time and may even make paths infeasible. Therefore, we change c_k to $c_k(\omega)$, where $c_k(\omega) = \infty$ if it path k is infeasible for scenario ω . This may happen, for example, when the passenger arrives at the transfer point after their driver has already departed because of a delay. We denote $p(\omega)$ the probability of scenario ω occurring, such that $p(\omega) \geq 0$, $\sum_{\omega \in \Omega} p(\omega) = 1$.

We enforce that all first-stage decisions are the same for all scenarios. Specifically, if driver j is involved in path k for scenario ω that is a direct match from origin to destination, he must commit to the same direct path in any other scenario. Therefore, as direct paths only have one leg, the chosen paths are identical for every scenario and are used to minimize the expected cost. This is enforced through

$$a_{jk}^0 x_k(\omega) = a_{jk}^0 x_k(\omega') \quad \forall j \in J, k \in K, \omega, \omega' \in \Omega \quad (6)$$

For an indirect path, only the first leg is fixed. In this case, the full path need not be the same, as long as the same driver goes to the same hub in both scenarios. In addition to this, we impose that both paths need to correspond to the same passenger. To enforce this, we use the following set of constraints. By enforcing the matched driver-passenger pair as well as the hub at which the passenger is dropped to be equal across scenarios, we guarantee that first-leg matches are fixed in advance.

$$\sum_{k \in K} e_{ik} a_{jk}^{1h} x_k(\omega) = \sum_{k \in K} e_{ik} a_{jk}^{1h} x_k(\omega') \quad \forall i \in I, j \in J, h \in H, \omega, \omega' \in \Omega \quad (7)$$

The full formulation of the stochastic programming problem, to which we refer as **P2**, is given as follows:

$$\text{(P2) minimize } \sum_{\omega \in \Omega} \sum_{k \in K} p(\omega) c_k(\omega) x_k(\omega) \quad (8a)$$

$$\sum_{k \in K} e_{ik} x_k(\omega) = 1 \quad \forall i \in I, \omega \in \Omega \quad (8b)$$

$$\sum_{k \in K} a_{jk}^0 x_k(\omega) \leq q_j \left(1 - \sum_{h \in H} y_{jh}(\omega) \right) \quad \forall j \in J, \omega \in \Omega \quad (8c)$$

$$\sum_{k \in K} a_{jk}^{1h} x_k(\omega) \leq q_j y_{jh}(\omega) \quad \forall j \in J, h \in H, \omega \in \Omega \quad (8d)$$

$$\sum_{k \in K} a_{jk}^{2h} x_k(\omega) \leq q_j y_{jh}(\omega) \quad \forall j \in J, h \in H, \omega \in \Omega \quad (8e)$$

$$\sum_{h \in H} y_{jh}(\omega) \leq 1 \quad \forall j \in J, \omega \in \Omega \quad (8f)$$

$$a_{jk}^0 x_k(\omega) = a_{jk}^0 x_k(\omega') \quad \forall j \in J, k \in K, \omega, \omega' \in \Omega \quad (8g)$$

$$\sum_{k \in K} e_{ik} a_{jk}^{1h} x_k(\omega) = \sum_{k \in K} e_{ik} a_{jk}^{1h} x_k(\omega') \quad \forall i \in I, j \in J, h \in H, \omega, \omega' \in \Omega \quad (8h)$$

$$x_k \in B \quad \forall k \in K \quad (8i)$$

$$y_{jh} \in B \quad \forall j \in J, h \in H \quad (8j)$$

The Objective (8a) is to minimize the expected costs, which is a linear function weighted by the probability of each scenario occurring. Constraints (8b) to (8f) are the same as in **(P1)**, but extended them with the scenario dependency ω . Constraints (8g) ensure that the same direct paths are chosen for every scenario. Constraints (8h) enforce the first leg of drivers to be the same on indirect paths and that they carry the same passenger.

3 RESULTS AND DISCUSSION

We evaluate our model on a circular city consisting of 33 nodes, as depicted in Figure 2. Every passenger and driver has an origin and destination at one of the 33 nodes. Origins are more likely to be in the suburbs (the outer rings) whereas destinations are more likely to be in the city center. Transfer hubs can be at any of the nodes in the network. Drivers can perform a pick-up or a drop-off at one of the transfer hubs, but only if the hub is on their shortest path. We consider 1000 drivers and 500 passengers. Desired arrival times are drawn from a truncated normal distribution with a mean at 8:00 and a standard deviation of 1 hour. The distribution is truncated such that we only allow desired arrival times between 6:30 and 9:30.

The parameter settings are homogeneous among the entire population and are defined as follows. The value of time spent in a car α is equal to 6.4[\$/h]. Earliness and lateness are penalized with β and γ equal to 3.9[\$/h] and 15.21[\$/h] respectively. The value of time in public transport α^{pt} is higher and is set equal to 11.0[\$/h]. In addition to this, public transport has a fixed cost of 2.0\$ per trip to compensate for waiting times. Waiting time is penalized by α^{wait} which is equal to 13.5[\$/h] such that $\beta < \alpha < \alpha^{pt} < \alpha^{wait} < \gamma$, consistent with the literature.

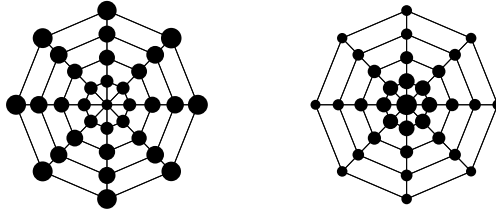


Figure 2: Circular city with the distribution of origins (left) and destinations (right)

Influence of transfers on cost and mode choice

We consider the influence transfers make on the average cost per individual in the deterministic system. The results are displayed in Figure 3 where the left-hand panel displays the modal split of passengers and the right-hand panel displays the composition of the average cost of passengers. Clearly, when there are no transfer hubs, the only possible mode choices are direct carpooling and public transport. By opening transfer hubs, a modal shift to the two modes that use transfers is observed. Especially the number of passengers carpooling on two separate legs increases. The reason for this is that by using a transfer, more options exist for matching to someone with the same destination and a similar desired arrival time, at the cost of waiting at the transfer point. The number of direct matches may be limited as the origin and destination of the passenger and driver need to be identical and the desired arrival time needs to be somewhat similar.

By using a single transfer hub in the center of the network, the average cost decreases from 14\$ to 12\$ ($\approx 15\%$). Increasing the number of transfer hubs allows a further decrease in the average cost, but not nearly as substantial as for the first hub in the center. When all 9 hubs are opened, the average cost decreases to 11\$ ($\approx 22\%$). We emphasize that carpooling benefits from economies of scale when the number of commuters increases. The reason for this is that the matching opportunities increase, which decreases the expected cost of a match.

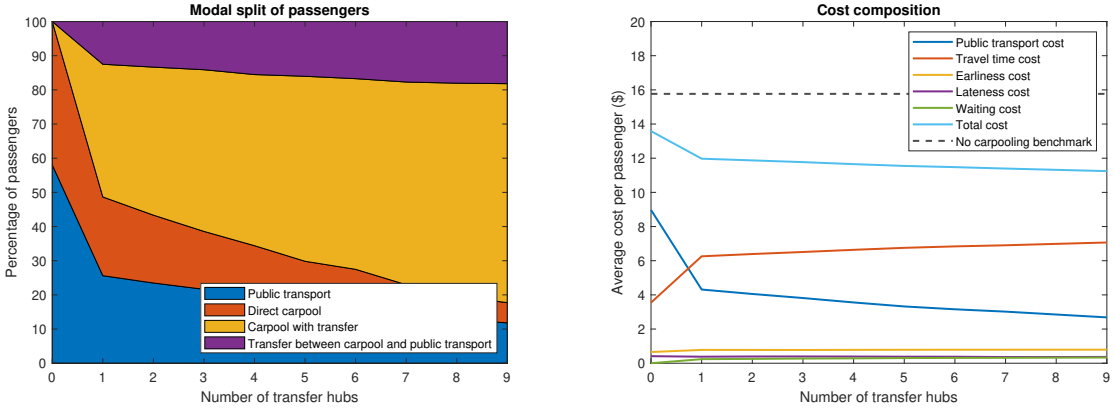


Figure 3: Statistics for a varying number of transfer hubs

Distribution of passengers by desired arrival time

We evaluate the distribution of passengers by desired arrival time and the mode they use to commute. We consider the deterministic system with only one hub in the center of the network. The results are displayed in Figure 4 where the left-hand panel displays the number of passengers using a mode and the right-hand panel displays the proportion of passengers with a specific desired arrival time using a mode.

The proportion of passengers using public transport is the highest in the tails. The reason for this is that the number of potential matches with identical origins and destinations and similar desired arrival times is low since the number of individuals here is rather low. This effect is more apparent for passengers with an early desired arrival time. When these passengers match to a driver, it is highly likely that the desired arrival time of the driver is later than that of the passenger, and therefore the passenger will suffer from lateness. As lateness is penalized heavier than earliness, the effect is more apparent at the start of the morning commute than it is at the end. At the peak of the rush hour, the number of carpoolers is the highest. We see a slight skewness towards later desired arrival times, which follows the same reasoning as stated before.

Stochastic Programming Results

We analyze the results of the stochastic programming problem, using three scenarios ($|\Omega| = 3$), with all travel times at 100%, 125%, and 150% of the free-flow travel times, respectively. A selection of 50 passengers is made to display the matching, where a bar represents a match to a driver. The color of the bar, as well as the index displayed on the bar, identifies the driver. The first bar

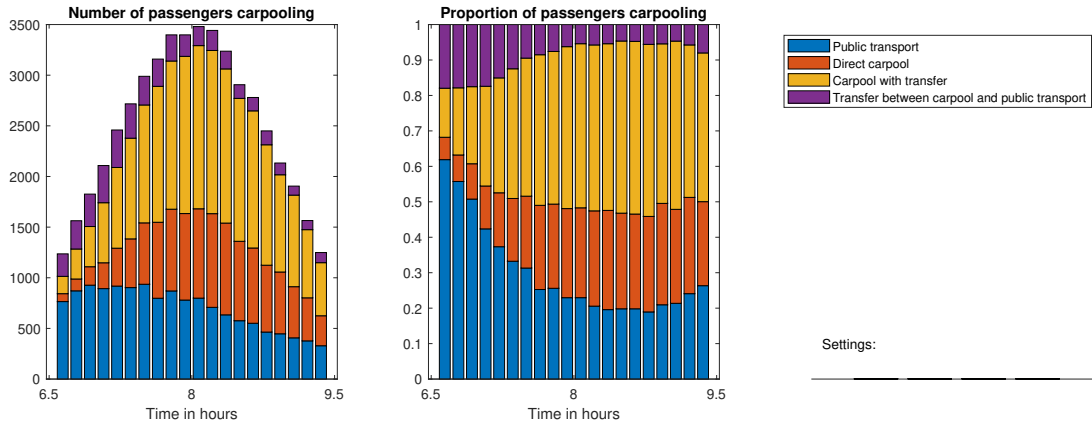


Figure 4: Distribution of passengers by desired arrival time and mode

displays a first-leg match, the second bar displays a second-leg match and the full bar represents a direct match. When no bar is given on a leg, the passenger uses public transport on this leg. The results are given in Figure 5.

We observe that in line with Constraints (8g) and (8h) the direct matches and first-leg matches are the same for all scenarios. Approximately 75% of the matches are identical across all three scenarios, whereas for the remaining 25% the passenger either changes mode or driver on the second part of the trip. We compare the result to the wait-and-see benchmark, where Constraints (8g) and (8h) are relaxed. This is displayed in Figure 6. For the wait-and-see benchmark only 50% of the matches are identical. The wait-and-see benchmark shows that the higher the travel times (for scenarios 2 and 3) the fewer matches are optimal. For the recourse problem, however, a middle ground needs to be found since the matching cannot be fully adapted to the exact traffic situation. Therefore, we observe slightly fewer matches in scenario 1 for the recourse problem, whereas we observe significantly more matches for scenario 3 compared to the wait-and-see problem.

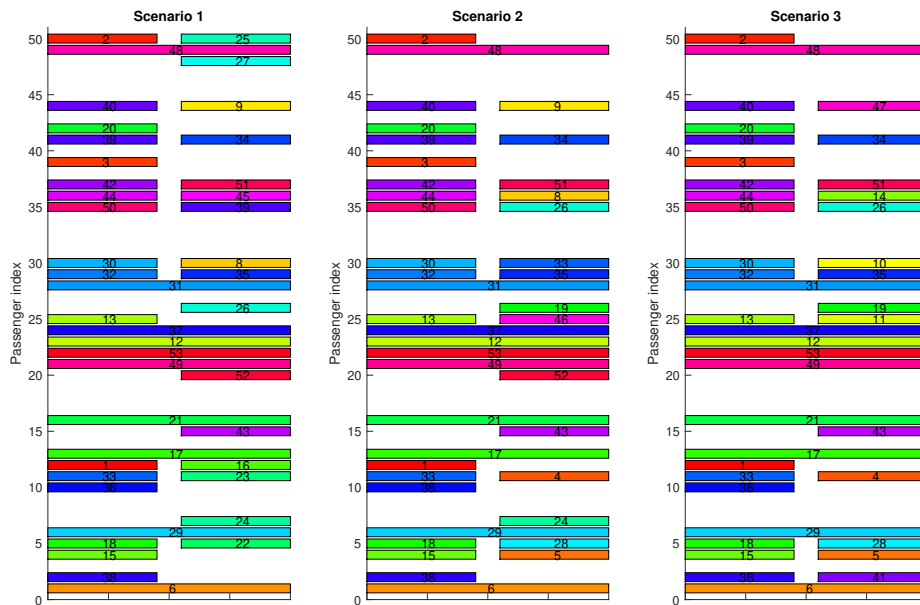


Figure 5: Gantt chart of matching for the recourse problem.

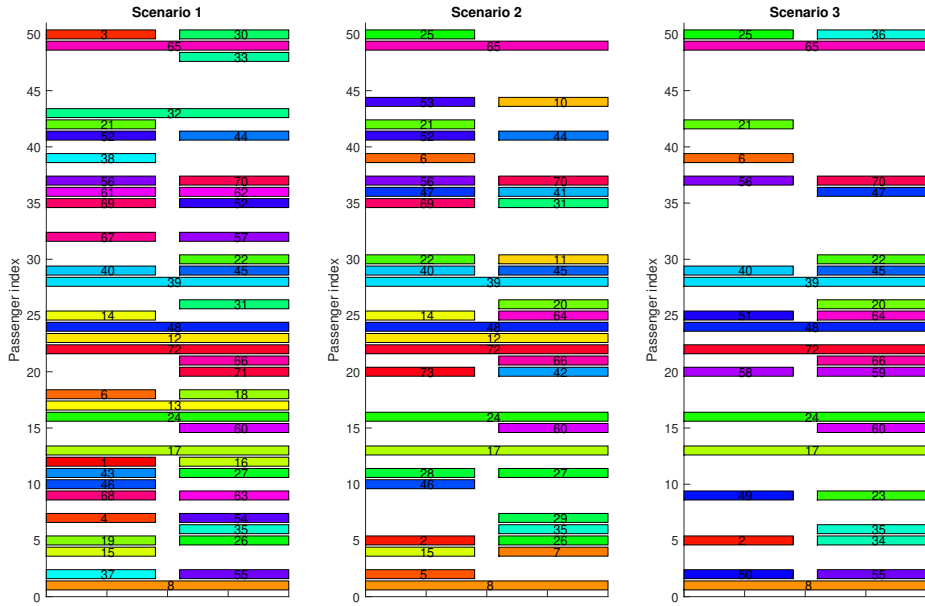


Figure 6: Gantt chart of matching for the wait-and-see benchmark problem.

4 CONCLUSIONS

We conclude that transfers can significantly reduce the average costs of commuters. Carpooling with transfers partially replaces direct carpooling as well as public transport as a mode of transport. It is especially beneficial for commuters during the peak of the commute, whereas commuters in the tails are usually better off taking public transport. In the case of stochastic travel times, we observe a large share of the commuters stick to the same mode and match for each scenario. However, 25% of the commuters may change their mode or match. We also observe that due to uncertainty, the number of carpoolers generally decreases.

REFERENCES

- Center for Sustainable Systems, University of Michigan. (2021). *Center for Sustainable Systems, University of Michigan. 2021. "Personal Transportation Factsheet." Pub. No. CSS01-07.*
- Herbawi, W., & Weber, M. (2012). The ridematching problem with time windows in dynamic ridesharing: A model and a genetic algorithm. In *2012 IEEE Congress on Evolutionary Computation* (pp. 1–8).
- Huang, H., Bucher, D., Kissling, J., Weibel, R., & Raubal, M. (2018). Multimodal route planning with public transport and carpooling. *IEEE Transactions on Intelligent Transportation Systems*, 20(9), 3513–3525.
- IEA. (2022). *Transport sector CO2 emissions by mode in the Sustainable Development Scenario, 2000-2030, IEA, Paris* <https://www.iea.org/data-and-statistics/charts/transport-sector-co2-emissions-by-mode-in-the-sustainable-development-scenario-2000-2030>, IEA. Licence: CC BY 4.0.
- Long, J., Tan, W., Szeto, W., & Li, Y. (2018). Ride-sharing with travel time uncertainty. *Transportation Research Part B: Methodological*, 118, 143–171.
- Lu, W., Liu, L., Wang, F., Zhou, X., & Hu, G. (2020). Two-phase optimization model for ride-sharing with transfers in short-notice evacuations. *Transportation research part C: emerging technologies*, 114, 272–296.
- Masoud, N., & Jayakrishnan, R. (2017). A decomposition algorithm to solve the multi-hop peer-to-peer ride-matching problem. *Transportation Research Part B: Methodological*, 99, 1–29.

A Spatial Branch and Bound Algorithm for Continuous Pricing with Advanced Discrete Choice Demand Modeling

Tom Haering*¹ and Michel Bierlaire²

¹Doctoral assistant, TRANSP-OR, EPFL, Switzerland

²Professor, TRANSP-OR, EPFL, Switzerland

SHORT SUMMARY

In this paper, we present a spatial branch and bound algorithm to tackle the continuous pricing problem, where demand is captured by an advanced discrete choice model (DCM). Advanced DCMs, like mixed logit or latent class models, are capable of modeling demand on the level of individuals very accurately due to a focus on behavioral realism. The downside of such realistic models is that it is highly nontrivial to include the resulting demand probabilities into an optimization problem, as they usually do not have a convex or even closed-form expression when decision variables are part of the choice model. To this end, a simulation procedure proposed by Paneque et al. (2021) is applied to get a formulation as a mixed integer linear program (MILP). However, due to the large number of variables stemming from the simulation, this MILP is very hard to solve. We first propose to solve the problem as a non-convex quadratically constrained quadratic program (QCQP) instead, where total unimodularity guarantees the integrality of the solution. Isolating all non-convexity into a set of bilinear constraints leads to a formulation as a non-convex quadratically constrained linear program (QCLP) that proves computationally beneficial for general-purpose solvers. Lastly, we present a spatial branch and bound algorithm that employs the McCormick envelope to obtain relaxations and makes use of total unimodularity to generate feasible solutions and thus lower bounds for the maximization fast. We compare the proposed method to the fastest commercially available solver GUROBI, on a parking choice case study from Ibeas et al. (2014). The results show that the custom spatial branch and bound approach outspeeds GUROBI by a factor of at least 35x for the MILP formulation and at least 2.5x for the QCLP in single-price optimization, and a factor of at least 4.5x for the QCQP and 1.3x for the QCLP when optimizing multiple prices simultaneously. The ratio of the speedup further increases with the size of the instance.

Keywords: branch and bound, discrete choice, mixed multinomial logit, optimization, pricing, simulation

1 INTRODUCTION

Pricing optimization is essential when pricing decisions need to be made for one or multiple products, particularly when there are cross-effects between their demands (Talluri & Van Ryzin, 2004). This problem can arise in various areas, including revenue management for airlines, railways, and hotels, assortment pricing in retail, or product line pricing in consumer goods industries.

While previous research has utilized the price-dependent multinomial logit (MNL) model to optimize prices for firms offering multiple products (Dong et al., 2009; Song et al., 2021), advanced discrete choice models such as mixed logit or latent class models have not been commonly used.

DCMs can capture the heterogeneity of customer preferences and the complex interactions between product attributes, which are often lost when the demand is aggregated. Furthermore, individual-level data can be used to identify profitable customer segments and to develop targeted pricing strategies. However, modeling demand at an individual level requires more data and computational resources compared to modeling the demand on an aggregate level.

In product assortment (PA) optimization, where a seller must make discrete decisions about the selection of products and their prices, the mixed multinomial logit (MMNL) has become increasingly popular (see e.g. Feldman et al., 2022). MMNL is regarded as a potent tool that captures the cross-effects in demand and can approximate any random utility choice model arbitrarily closely (Train, 2009). Since the PA problem under the MMNL choice model (or any other advanced choice

model that leads to non-convex probability formulas) is NP-hard (G. Li et al., 2015; Désir et al., 2015), much work has been focused on deriving upper bounds and efficient approximations, with the recent exception of Sen et al. (2017) who propose an exact conic MIP approach. Despite its theoretical and practical relevance, the MMNL model and its incorporation into revenue maximization have received little attention in the dynamic pricing literature (e.g. Keskin, 2014), with researchers often sacrificing behavioral realism for tractable (concave) formulations and therefore considering MNL (Dong et al., 2009; Keller et al., 2014) or nested logit H. Li & Huh (2011) instead. A general implementation approach for integrating any advanced choice model into an optimization problem has been proposed in Paneque et al. (2021), where Monte Carlo simulation is used to generate a deterministic problem at the cost of an increase in complexity since the resulting mixed integer linear problem (MILP) involves finding the best price over a large number of scenarios, generated by taking draws from the stochastic components of the formulation. With a sufficiently large number of draws, the MILP formulation guarantees convergence to globally optimal solutions. However, since the complexity of the MILP scales exponentially with the number of draws, the approach can currently only be applied to solving small-scale instances, i.e., with few individuals and alternatives.

In this work, we extend the MILP approach in Paneque et al. (2021) by first restating it as a non-convex quadratically constrained quadratic program (QCQP), and then as a non-convex quadratically constrained linear program (QCLP), for which we develop a spatial branch and bound algorithm that efficiently solves the problem for large numbers of draws. We compare the MILP, the QCQP, and QCLP formulations (all solved using the mathematical solver GUROBI) to our spatial branch and bound approach by application to a parking choice case study by Ibeas et al. (2014).

2 METHODOLOGY

We first present the original MILP formulation of the CPP that results when applying the approach of Paneque et al. (2021) directly:

MILP formulation

$$\max_{p, \omega, U, H} \frac{1}{R} \sum_r \sum_n \sum_i p_i \omega_{inr}$$

s. t.

$$\sum_i \omega_{inr} = 1 \quad (\mu_{nr})$$

$$H_{nr} = \sum_i U_{inr} \omega_{inr} \quad (\zeta_{nr})$$

$$H_{nr} \geq U_{inr} \quad (\alpha_{inr})$$

$$U_{inr} = \sum_{k \neq p} \beta_k x_{ink} + \beta_p p_i + \varepsilon_{inr} \quad (\kappa_{inr})$$

$$\omega \in \{0, 1\}$$

$$p, U, H \in \mathbb{R}$$

Consider a set of $n = \{1, \dots, N\}$ individuals choosing exactly one alternative among a set of $i = \{1, \dots, I\}$ alternatives. Assume that in every scenario r , each individual n selects the alternative i corresponding to the maximal utility U_{inr} . The utility function depends on K parameters β_k which are exogenous and estimated using an advanced discrete choice model where all prices were fixed. These parameters are multiplied by individual or alternative specific attributes x_{ink} , which are also exogenous, just like the added error term draws ε_{inr} . For the CPP the price p_i of alternative i becomes a decision variable that is to be optimized in order to maximize profit. Note that the only assumption we make on the utilities is that

Formulation 1 – CPP as a MILP

they are linear in the price variables. Denote by ω_{inr} the binary decision variable that indicates whether individual n chooses alternative i in scenario r . The choice probabilities are then approximated by $P_n(i) \approx \frac{1}{R} \sum_r \omega_{inr}$ and are guaranteed to converge to the real probabilities with a sufficiently large number of scenarios R , see Paneque et al. (2021). The objective function is equal to the profit and is thus defined as the average number of times that individual n chooses alternative i over all scenarios r (i.e. its choice probability) multiplied by the alternative's price p_i . The constraints define the individual choices: Constraints (μ_{nr}) guarantee that only one alternative can be chosen per individual and scenario. Constraints (κ_{inr}) model the utility U_{inr} of each alternative i for individual n in scenario r . Constraints (ζ_{nr}) and constraints (α_{inr}) ensure that the choice being made corresponds to the one with the highest utility. Note that both the objective and the constraints ζ_{nr} contain the product $p_i \omega_{inr}$, which can be linearized using a big-M approach.

Continuous reformulation

$$\begin{aligned}
& \max_{p, \omega, U, H} \frac{1}{R} \sum_r \sum_n \sum_i p_i \omega_{inr} \\
& \text{s.t.} \\
& \sum_i \omega_{inr} = 1 \quad (\mu_{nr}) \\
& H_{nr} = \sum_i U_{inr} \omega_{inr} \quad (\zeta_{nr}) \\
& H_{nr} \geq U_{inr} \quad (\alpha_{inr}) \\
& U_{inr} = \sum_{k \neq p} \beta_k x_{ink} + \beta_p p_i + \varepsilon_{inr} \quad (\kappa_{inr}) \\
& \omega \in [0, 1] \\
& p, U, H \in \mathbb{R}
\end{aligned}$$

Formulation 2 – CPP as a QCQP

$$\begin{aligned}
& \max_{p, \omega, \eta, U, H} \frac{1}{R} \sum_r \sum_n \sum_i \eta_{inr} \\
& \text{s.t.} \\
& \sum_i \omega_{inr} = 1 \quad (\mu_{nr}) \\
& H_{nr} = \sum_i \left(\sum_{k \neq p} \beta_k x_{ink} + \varepsilon_{inr} \right) \omega_{inr} + \beta_p \eta_{inr} \\
& H_{nr} \geq U_{inr} \quad (\alpha_{inr}) \\
& U_{inr} = \sum_{k \neq p} \beta_k x_{ink} + \beta_p p_i + \varepsilon_{inr} \quad (\kappa_{inr}) \\
& \eta_{inr} = p_i \omega_{inr} \quad (\lambda_{inr}) \\
& \omega \in [0, 1] \\
& p, \eta, U, H \in \mathbb{R}
\end{aligned}$$

Formulation 3 – CPP as a QCLP

for the prices p_i we have to assume that its possible to define a reasonable range for each price, $p_i \in [p_i^L, p_i^U]$, which is usually the case in practice. To go from solving a relaxation to an approximation of the optimal solution of the original problem, we employ a so-called spatial branch and bound algorithm: We start by solving the relaxation with the initial bounds $p \in [p^L, p^U]$. If it is infeasible, the original problem is also infeasible and we are done. If it is feasible and the optimal solution found is also feasible for the original problem, we are done as well. There are multiple ways to define feasibility for the original problem, one way would be to check if all ω are close enough to integer values, another is to check how strongly the relaxed bilinear constraints are violated. If we find an optimal solution for the relaxation that is infeasible for the original, we store the objective value of that solution as an upper bound for the objective value of all subpolyhedra and we start the branching: This means that instead of looking at the entire space $p \in [p^L, p^U]$, we choose an alternative i and split the domain $[p_i^L, p_i^U]$ of p_i into two smaller intervals $[p_i^L, \frac{p_i^L + p_i^U}{2}]$ and $[\frac{p_i^L + p_i^U}{2}, p_i^U]$. Using these two new sets of bounds, we create two new (sub)polyhedra where we solve the relaxation again and iterate the procedure. After a branching, we always proceed on the branch which has the highest upper bound on its objective value (this is also called a *best-first-search*). Furthermore, after solving each relaxation, we can use the optimal value of the p_i variable that we get from the relaxation to compute an integer solution to the original problem. This can be done very efficiently as for fixed prices, all individuals and scenarios become completely independent, and finding the optimal values of the ω variables reduces to assigning 1 to the alternative with the highest utility and 0 to all others.

Our first reformulation (Formulation 2) defines the CPP as a non-convex quadratically constrained quadratic problem (QCQP), with a quadratic objective and quadratic equality constraints, making them non-convex. The formulation is equivalent to the MILP in Formulation 1, except that the variables ω_{inr} are no longer constrained to be binary and instead are relaxed to be in the interval $[0, 1]$. Integrality still holds, since for any price p_i the problem of choosing the alternative with the highest utility is a knapsack problem, which is totally unimodular.

Our second reformulation (Formulation 3) isolates all non-convexity into a set of bilinear constraints λ_{inr} which define the product $p_i \omega_{inr}$, turning the problem into a non-convex quadratically constrained linear program (QCLP).

Spatial Branch and Bound algorithm

We start from the QCLP formulation of the CPP shown in Formulation 3. We then construct a linear relaxation (Formulation 4) by replacing the bilinear constraints λ_{inr} by a set of inequalities (λ_{inr}^1 to λ_{inr}^4) that define the McCormick envelope, see McCormick (1976). This relaxation is a commonly used device to tackle problems with bilinear constraints. For the McCormick envelope, we need to provide bounds for both variables in the product $p_i \omega_{inr}$. For ω_{inr} this is straightforward, as we can simply set the lower and upper bound to 0 and 1 respectively, whereas

$$\begin{aligned}
& \max_{p, \omega, \eta, U, H} \frac{1}{R} \sum_r \sum_n \sum_i \eta_{inr} \\
& \text{s.t.} \\
& \sum_i \omega_{inr} = 1 \quad (\mu_{nr}) \\
& H_{nr} = \sum_i \left(\sum_{k \neq p} \beta_k x_{ink} + \varepsilon_{inr} \right) \omega_{inr} + \beta_p \eta_{inr} \\
& H_{nr} \geq U_{inr} \quad (\alpha_{inr}) \\
& U_{inr} = \sum_{k \neq p} \beta_k x_{ink} + \beta_p p_i + \varepsilon_{inr} \quad (\kappa_{inr}) \\
& \eta_{inr} \geq p_i^L \omega_{inr} \quad (\lambda_{inr}^1) \\
& \eta_{inr} \geq p_i^U \omega_{inr} + p_i - p_i^U \quad (\lambda_{inr}^2) \\
& \eta_{inr} \leq p_i^L \omega_{inr} + p_i - p_i^L \quad (\lambda_{inr}^3) \\
& \eta_{inr} \leq p_i^U \omega_{inr} \quad (\lambda_{inr}^4) \\
& \omega \in [0, 1] \\
& p, \eta, U, H \in \mathbb{R}
\end{aligned}$$

Formulation 4 – Linear relaxation of the CPP using the McCormick envelope

where we branch along the asset which displays the largest maximum violation of the constraints $\eta_{inr} = p_i \omega_{inr}$. Algorithm 1 provides the pseudo-code for the described procedure. It is worth noting that this branch and bound algorithm will always terminate with a 0% gap in a finite number of steps since we do not actually need to find the exact optimal price, but rather the bounds for the price such that the optimal choices are generated. It then follows that, within those sufficiently optimal bounds, the obtained price will be optimal as well.

This integer solution is feasible for the original problem and thus provides us with a global lower bound for the branching tree, meaning we can delete all branches whose upper bound is less or equal to the best (highest) known lower bound. If we ever find a solution during the branching which is also feasible for the original, we can use it as a lower bound as well. This process continues until the highest upper bound u_{max} from all active branches is at most a certain tolerance percentage perc_{tol} away from the best lower bound l_{max} , i.e.:

$$\frac{l_{max} - u_{max}}{l_{max}} \cdot 100 \leq \text{perc}_{\text{tol}}.$$

If we deal with more than one price that we want to optimize, an important decision we have to make at each branching is which alternative's price to branch on. The traditional approach of spatial branch and bound is to always select the alternative i where the interval $[p_i^L, p_i^U]$ is the largest (*longest-edge-branching*). This is comparable to a strategic exhaustive search. In our case we utilize a custom branching rule

Algorithm 1: A spatial Branch & Bound algorithm to solve the CCP

Result: perc_{tol} -optimal solution (p^*, ω^*, η^*) for Formulation 3.

Initialization: Set $j := 0$, $\Delta^j := [p_1^L, p_1^U] \times \dots \times [p_j^L, p_j^U]$, $o^* := -\infty$, $\hat{o}_j := \infty$,

$\Omega := \{\{\Delta^j, \hat{o}_j\}\}$

while $\frac{o^* - \max_j \{\hat{o}_j\}}{o^*} \cdot 100 \leq \text{perc}_{\text{tol}}$ **and** $\Omega \neq \emptyset$ **do**

 let $j := \text{argmax}\{\hat{o}_j \mid \{\Delta^j, \hat{o}_j\} \in \Omega\}$. Remove $\{\Delta^j, \hat{o}_j\}$ from Ω and solve Formulation 4 with bounds Δ^j .

if Formulation 4 is feasible **then**

 denote its optimal solution by (p_j, ω_j, η_j) and its optimal objective value by o_j as well as its integer optimal value \bar{o}_j .

if $\bar{o}_j > o^*$ **then**

 compute $\bar{\omega}_j, \bar{\eta}_j$ from p_j and set $o^* = o_j$, $(p^*, \omega^*, \eta^*) := (p_j, \bar{\omega}_j, \bar{\eta}_j)$, delete from Ω all instances $\{\Delta^j, \hat{o}_j\}$ where $\hat{o}_j \leq o^*$.

end

if $o_j > o^*$ **then**

if (p_j, ω_j, η_j) is feasible for Formulation 3 **then**

$o^* = o_j$, $(p^*, \omega^*, \eta^*) := (p_j, \omega_j, \eta_j)$, delete from Ω all instances $\{\Delta^j, \hat{o}_j\}$ where $\hat{o}_j \leq o^*$.

else

 let $i = \text{argmax}\{\max_{nr} |\eta_{inr} - p_i \omega_{inr}| \mid i \in J\}$ and divide the interval $[p_i^L, p_i^U]$ into two new intervals $[p_i^L, \frac{p_i^L + p_i^U}{2}]$ and $[\frac{p_i^L + p_i^U}{2}, p_i^U]$. Construct the two new subpolyhedra Δ' and Δ'' . Define $\hat{o}' = \hat{o}'' := o_j$ and augment $\Omega = \Omega \cup \{\Delta', \hat{o}'\} \cup \{\Delta'', \hat{o}''\}$.

end

end

end

end

3 RESULTS AND DISCUSSION

To test the presented methodology we rely on a case study of a parking services operator, which is motivated by a published disaggregate demand model for parking choice by Ibeas et al. (2014). The choice set consists of three services: paid on-street parking (PSP), paid parking in an underground car park (PUP), and free on-street parking (FSP). Since the latter does not provide any revenue to the operator, it represents the opt-out option. We assume that all customers must pay the same price for the same service. The explanatory variables considered in the discrete choice model estimated by Ibeas et al. (2014) include the following socioeconomic characteristics: trip origin (if outside town, it affects the utility of free street parking), age of the vehicle (if less than three years old, it affects the utility of paid underground parking), the income of the driver (if low, it affects the utility of paid alternatives), area of residency of the driver (if in town, it affects the utility of paid alternatives). Additionally, the following attributes of the alternatives are considered: access time to destination, access time to parking and parking fee. For the latter two continuous variables, the corresponding coefficients are normally distributed in the utility function, making the choice model a mixed multinomial logit (MMNL). Table 1 illustrates the parameters of the discrete choice model.

Table 1 – Utility parameters derived from Ibeas et al. (2014)

Parameter	Value
ASC_{FSP}	0.0
ASC_{PSP}	32.0
ASC_{PUP}	34.0
Fee (€)	$\sim \mathcal{N}(-32.328, 14.168)$
Fee PSP - low income (€)	-10.995
Fee PUP - low income (€)	-13.729
Fee PSP - resident (€)	-11.440
Fee PUP - resident (€)	-10.668
Access time to parking (min)	$\sim \mathcal{N}(-0.788, 1.06)$
Access time to destination (min)	-0.612
Age of vehicle (1/0)	4.037
Origin (1/0)	-5.762

We run two series of tests: in the first, we fix the price of PSP to be 0.6€ and only optimize the price of PUP. This reduces the complexity enough to make meaningful comparisons to the computationally heavy MILP model. We consider a random subset of 100 customers and stepwise increase the number of random draws from 100 to 1000. For the second series of tests, we optimize both the price of PSP and PUP, but we only consider a set of 50 customers, with the same range of random draws. All experiments are performed using GUROBI 10.0.0 (Gurobi Optimization, LLC, 2021) on a 2.6 GHz 6-Core Intel Core i7 processor with 16 GB of RAM, on a single thread and with a two-hour time limit per instance. Tables 2 and 3 show the solve time with achieved optimality gap and the objective values with computed prices respectively for optimizing only the price of the PUP alternative, whereas Tables 4 and 5 depict the same outputs when optimizing PSP and PUP prices together. It is evident that computing the prices of two competing alternatives that influence each other’s demands simultaneously is much more computationally challenging than optimizing a single price only. For optimizing only the PUP price, our algorithm outspeeds the MILP by a factor of at least 35x, the QCQP by at least 4x, and the QCLP by at least 2.5x, with the ratio increasing with the number of draws. When optimizing both the PSP and PUP prices, the MILP solver never terminates but for the QCQP we note a speedup of at least 4.5x and for the QCLP of at least 1.3x. Again, increasing the number of draws also increases the ratio of the speedup. For more than 600 draws, the QCLP solver is not able to generate any feasible solutions in the two-hour time window, whereas our spatial branch and bound approach finds feasible solutions with an objective value up to 2.5x higher than the one found by the MILP or the QCQP.

Table 2 – Solve time (seconds) for optimizing PUP price only

N	R	MILP		QCQP		QCLP		B&B	
		Time	Gap (%)	Time	Gap (%)	Time	Gap (%)	Time	Gap (%)
100	100	3360	0	291	0	182	0	95	0
100	200	7200	22.89	1256	0	794	0	398	0
100	300	7200	134.28	3307	0	2584	0	976	0
100	400	7200	128.56	6522	0	4275	0	1593	0
100	500	7200	141.54	7200	0.74	7093	0	2661	0
100	600	7200	118.14	7200	26.19	7200	0.57	3620	0
100	700	7200	128.45	7200	36.83	7200	0.75	5283	0
100	800	7200	113.12	7200	-	7200	5.27	7200	0.37
100	900	7200	142.49	7200	-	7200	-	7200	1.94
100	1000	7200	149.03	7200	-	7200	-	7200	10.89

Table 3 – Objective value and optimal solution for optimizing PUP price only

N	R	MILP		QCQP		QCLP		B&B	
		Obj.	Price	Obj.	Price	Obj.	Price	Obj.	Price
100	100	54.13	[0.6, 0.66]	54.13	[0.6, 0.66]	54.13	[0.6, 0.66]	54.13	[0.6, 0.66]
100	200	54.54	[0.6, 0.65]	54.6	[0.6, 0.66]	54.6	[0.6, 0.66]	54.6	[0.6, 0.66]
100	300	54.38	[0.6, 0.64]	54.48	[0.6, 0.67]	54.48	[0.6, 0.67]	54.48	[0.6, 0.67]
100	400	54.15	[0.6, 0.63]	54.39	[0.6, 0.66]	54.39	[0.6, 0.66]	54.39	[0.6, 0.67]
100	500	54.27	[0.6, 0.66]	54.23	[0.6, 0.65]	54.29	[0.6, 0.67]	54.29	[0.6, 0.67]
100	600	54.15	[0.6, 0.64]	49.13	[0.6, 0.97]	54.25	[0.6, 0.65]	54.29	[0.6, 0.67]
100	700	54.14	[0.6, 0.63]	49.18	[0.6, 0.97]	54.37	[0.6, 0.65]	54.39	[0.6, 0.66]
100	800	54.32	[0.6, 0.66]	-	-	53.82	[0.6, 0.61]	54.32	[0.6, 0.65]
100	900	54.43	[0.6, 0.67]	-	-	-	-	54.44	[0.6, 0.67]
100	1000	54.42	[0.6, 0.66]	-	-	-	-	54.35	[0.6, 0.68]

Table 4 – Solve time (seconds) for optimizing PSP and PUP prices together

N	R	MILP		QCQP		QCLP		B&B	
		Time	Gap (%)	Time	Gap (%)	Time	Gap (%)	Time	Gap (%)
50	100	7200	33.76	4283	0	1231	0	918	0
50	200	7200	70.03	6719	0	5187	0	3495	0
50	300	7200	125.65	7200	1.03	7200	1.53	7200	0.28
50	400	7200	186.21	7200	9.55	7200	20.7	7200	2.88
50	500	7200	272.78	7200	33.07	7200	40.88	7200	7.7
50	600	7200	379.53	7200	42.8	7200	41.99	7200	10.59
50	700	7200	440.25	7200	260.91	7200	-	7200	16.82
50	800	7200	495.39	7200	260.28	7200	-	7200	22.91
50	900	7200	493.85	7200	260.97	7200	-	7200	25
50	1000	7200	-	7200	260.67	7200	-	7200	29.16

Table 5 – Objective value and optimal solution for optimizing PSP and PUP prices together

N	R	MILP		QCQP		QCLP		B&B	
		Obj.	Price	Obj.	Price	Obj.	Price	Obj.	Price
50	100	27.37	[0.58, 0.72]	27.55	[0.61, 0.7]	27.55	[0.61, 0.7]	27.55	[0.61, 0.7]
50	200	23.92	[0.66, 0.95]	27	[0.55, 0.68]	27	[0.55, 0.68]	26.99	[0.55, 0.68]
50	300	16.76	[1.02, 1.1]	27.12	[0.56, 0.67]	27.09	[0.56, 0.68]	27.12	[0.56, 0.67]
50	400	15.83	[1.01, 1.17]	27.07	[0.59, 0.69]	26.29	[1.09, 0.67]	27.15	[0.56, 0.66]
50	500	12.18	[1.3, 1.34]	26.4	[0.59, 0.8]	26.42	[0.59, 0.8]	27.13	[0.57, 0.68]
50	600	9.45	[1.32, 1.66]	26.36	[0.59, 0.8]	26.37	[0.59, 0.8]	27.23	[0.56, 0.69]
50	700	8.43	[1.33, 1.85]	11.73	[1.2, 1.39]	-	-	26.87	[0.62, 0.69]
50	800	7.65	[1.76, 1.76]	11.74	[1.2, 1.39]	-	-	26.37	[0.75, 0.62]
50	900	7.63	[1.74, 1.76]	11.7	[1.2, 1.39]	-	-	26.36	[0.75, 0.62]
50	1000	-	-	11.73	[1.2, 1.39]	-	-	26.35	[0.75, 0.62]

4 CONCLUSIONS

We propose a spatial branch and bound algorithm to tackle the continuous pricing problem, where demand is captured by an advanced discrete choice model (DCM). The stochasticity in the demand is dealt with using simulation, which leads to a large MILP formulation that is difficult to solve. We show that already reformulating the MILP as a non-convex QCQP improves computational speed significantly, even more so when formulated as a non-convex QCLP. The spatial branch and bound procedure solves the problem significantly faster than GUROBI on the tested instances, outspeeding the MILP by a factor of at least 35x and the QCLP by at least 2.5x for single price optimization, and outspeeding the QCQP by a factor of at least 4.5x and the QCLP by at least 1.3x for optimizing two prices simultaneously. Increasing the size of the instance also increases the ratio of the speedup. The methodology could be substantially advanced in future research by making use of the separability of each relaxed subproblem, since the price is the only complicating variable preventing the problem from being solved for each individual and scenario separately. Thus a Benders decomposition would be a suitable candidate for further improvement of the method. The authors would also like to perform a series of comparisons using only open-source software like for example SCIP and COUENNE, as not every business might be capable of purchasing a license for GUROBI. Last but not least, the code has not yet been optimized in terms of language or data structure usage, which could both have a strong impact on the performance.

REFERENCES

- Désir, A., Goyal, V., Segev, D., & Ye, C. (2015). Capacity constrained assortment optimization under the markov chain based choice model. *Available at SSRN 2626484*.
- Dong, L., Kouvelis, P., & Tian, Z. (2009). Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11(2), 317–339.
- Feldman, J., Wagner, L., Topaloglu, H., & Bai, Y. (2022). Assortment optimization under the multinomial logit model with utility-based rank cutoffs. *Available at SSRN*.
- Gurobi Optimization, LLC. (2021). *Gurobi Optimizer Reference Manual*. Retrieved from <https://www.gurobi.com>
- Ibeas, A., Dell’Olio, L., Bordagaray, M., & Ortúzar, J. d. D. (2014). Modelling parking choices considering user heterogeneity. *Transportation Research Part A: Policy and Practice*, 70, 41–49.
- Keller, P. W., Levi, R., & Perakis, G. (2014). Efficient formulations for pricing under attraction demand models. *Mathematical Programming*, 145, 223–261.
- Keskin, N. B. (2014). Optimal dynamic pricing with demand model uncertainty: A squared-coefficient-of-variation rule for learning and earning. *Available at SSRN 2487364*.

- Li, G., Rusmevichientong, P., & Topaloglu, H. (2015). The d-level nested logit model: Assortment and price optimization problems. *Operations Research*, 63(2), 325–342.
- Li, H., & Huh, W. T. (2011). Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Operations Management*, 13(4), 549–563.
- McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part i—convex underestimating problems. *Mathematical programming*, 10(1), 147–175.
- Paneque, M. P., Bierlaire, M., Gendron, B., & Azadeh, S. S. (2021). Integrating advanced discrete choice models in mixed integer linear optimization. *Transportation Research Part B: Methodological*, 146, 26–49.
- Sen, A., Atamturk, A., & Kaminsky, P. (2017). A conic integer programming approach to constrained assortment optimization under the mixed multinomial logit model. *arXiv preprint arXiv:1705.09040*.
- Song, J.-S. J., Song, Z. X., & Shen, X. (2021). Demand management and inventory control for substitutable products. *Available at SSRN 3866775*.
- Talluri, K., & Van Ryzin, G. (2004). Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1), 15–33.
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

The Impact of Weather Phenomena on Passenger Volumes for Commuter Trains

Ruben A. Kuipers*^{1,2}, Michelle Ochsner^{1,2}

¹ PhD student, Department of Technology and Society, Lund University, Lund, Sweden

² PhD student, K2 Swedish Knowledge Centre for Public Transport, Lund, Sweden

SHORT SUMMARY

Weather impacts several aspects of our daily life, of which the way we travel is one. The relationship between weather phenomena and ridership of trains has not received much attention previously. The study we present here aims to understand the impact of temperature, wind, and precipitation on passenger volumes for commuter trains. To do so we make use of automatic passenger count and weather data from over a million unique station stops, spanning two years, in the Southern region of Scania in Sweden. Our findings show that changes in the level of precipitation do not affect the volume of boarding passengers. Statistically significant effects are found for changes in temperature and wind speeds. These effects are most prominent for departures outside of peak hours. The results are useful for planning more accurate dwell times and rolling stock circulations and can serve as inputs during real-time rescheduling problems and demand modelling.

Keywords: Trains, weather, planning, passengers, railway

1. INTRODUCTION

The weather has an impact on many aspects of our daily life and the activities we conduct. One of these activities is how we travel. Spinney and Millward (2011), for example, found that cold weather and precipitation lead to more home-based activities. There is a large body of literature focusing on travel behaviour and mode choices in relation to weather conditions. For an extensive review, we refer to Böcker et al. (2013). However, studies focusing on the effects of weather on rail transport are scarce (Koetse & Rietveld, 2009). The study we present here focuses on changes in passenger volumes for commuter trains in relation to weather conditions to help fill this knowledge gap. Although not explicitly focusing on travel by railway, some indications of changes in passenger demand and weather for public transport can be identified in the literature. Strong winds and warmer weather conditions have been found to reduce the use of public transport for example. On the other hand, it was found that public transport usage increases under cold weather conditions and when precipitation increases (Sabir, 2011). In contrast to this, Cools et al. (2010) found that temperature has a lesser effect on travel behaviour than other weather conditions such as fog and wind. These studies thus indicate a potential relationship between the ridership of commuter trains and weather conditions.

It is relevant to study how different weather phenomena influence the number of people who travel by train since passenger demand has several important implications for railway operations and scheduling, as well as passenger demand modelling. The expected passenger demand affects the necessary rolling stock to be in use, for example. Having insufficient rolling stock in use can lead to high levels of onboard crowding, which is negative in terms of passenger comfort and experiences (Cox et al., 2006). Whereas having too much rolling stock in use results in increased operational costs. The volume of boarding passengers has also been found to influence dwell time

punctuality (Kuipers & Palmqvist, 2022; Li et al., 2016). Dwell time, in the case of passenger trains, refers to the time a train is stationary at a station to allow for the exchange of passengers. Dwell time delays reduce the overall punctuality and reliability of railways, and with this the attractiveness of railways as a mode of transport (Brons & Rietveld, 2008; van Loon et al., 2011).

Although it has been shown that weather influences how we travel, and that passenger load factors influence the operation of railways, the possible effect of weather on passenger demand for commuter trains has not received as much attention. The study we present here, therefore, focuses on how the volume of boarding passengers changes under different weather phenomena: wind speed, temperature, and precipitation. This knowledge can help planners to schedule more accurate rolling stock circulations and dwell times. Furthermore, it can serve as important input during real-time rescheduling problems and demand modelling. The latter is relevant since demand modelling is commonly based on data from travel surveys, aggregated over multiple days, which excludes nuances such as changes in weather conditions (Lepage & Morency, 2021).

2. METHOD

To study we present here makes use of automatic passenger count data collected on board commuter trains in the Southern region of Scania in Sweden during 2018 and 2019. A simplified line map of the region is shown in Figure 1. The commuter train sets consist of four carriages, with a total of 240 seats available and five doors on either side of the train. Individual trainsets can be combined to increase capacity, increasing the available seats and doors. The automatic passenger counters make use of infrared beams to count both the number of boarding and alighting passengers on a door-by-door level, with a minimum detection height of one meter. For the purpose of this study, we only make use of the volume of boarding passengers. The number of boarding passengers is a good reflection of the number of people who choose to use the train as a mode of transport. The weather data used in this study is collected by the Swedish Meteorological and Hydrological Institute on an hourly basis. The weather data has observations for the same period as the automatic passenger count data, from several different weather stations. The number of weather stations per weather variable is shown in Table 1, along with the aggregation steps for the weather data.

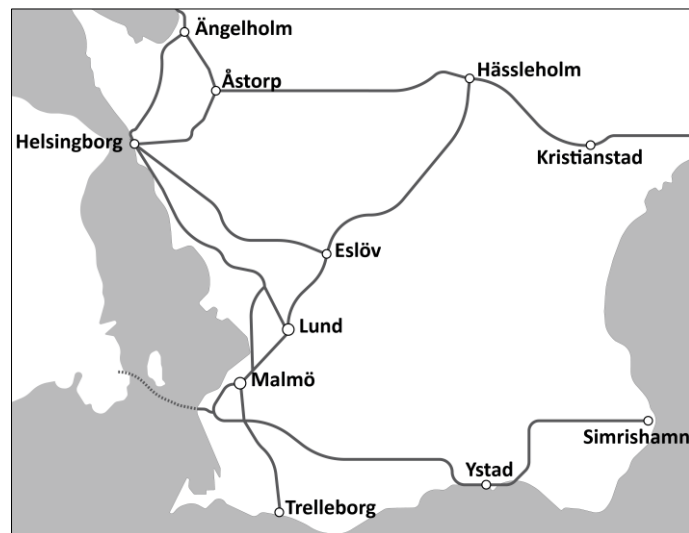


Figure 1: Simplified line map showing the railway network and some stations in Scania, Sweden.

Table 1: Overview of weather data.

Weather type	Number of weather stations	Scale	Aggregation
Temperature	12	Celsius	Bins of five degrees Celsius
Wind	14	Meters per second	Beaufort scale
Precipitation	15	Millimetre per hour	Dry (<i>0 mm/h</i>) Slight Precipitation (<i>0 to 0.4 mm/h</i>) Moderate Precipitation (<i>0.5 to 3.9 mm/h</i>) Heavy Precipitation (<i>>= 4 mm/h</i>)

To combine the automatic passenger count data and weather data, we matched the data from the weather stations to a train stop in space and time. The first step involved matching the railway stations to the nearest weather station based on their respective coordinates. The upper limit for the distance between a weather station and a railway station was set at 25 kilometres. Matches where this threshold was exceeded were excluded from the analysis. The second step consisted of matching the observations from the automatic passenger count system at each station with the hourly observations from the previously matched weather station. The process of matching the automatic passenger count data and weather data resulted in 1,296,576 data points having information on the volume of boarding passengers for unique station stops.

For this study, we are interested in changes in the frequency of passenger volumes rather than the exact volume. Knowing how often a certain number of people will board a train given a specific situation is more relevant for planning purposes compared to knowing the exact volume. Prior to the analyses, we split the data into peak hour and off-peak hour departures. Studies on trip and mode choices in relation to weather conditions show that it is important to account for the trip purpose, where less flexibility is found for trips made by commuters (Cools et al., 2010; Liu et al., 2014). A trip is considered to be made during peak hours when it departed on a weekday between 06:00 and 08:00 or between 15:30 and 17:30. These times are based on observed passenger volumes.

We make use of a series of pairwise Chi-square goodness of fit tests to determine whether changes in the frequency in the volume of boarding passengers under different weather conditions are statistically significant. To control for a potential familywise error rate for multiple comparisons we corrected the significance levels using the Bonferroni correction, as suggested by (McDonald, 2014). To perform the Chi-square goodness of fit tests we compared the difference between the frequency distribution under different weather conditions and the unconditional frequency distribution of boarding passenger volumes. When weather phenomena do not have an impact we expect the frequency distribution to be the same as the unconditional frequency distribution. When performing a Chi-square goodness of fit, the expected frequencies should at least be five. In order to ensure we fulfil this criterion we aggregated the passenger count data in steps of five passengers. The analyses are limited to a volume of 40 boarding passengers due to a lack of sufficient data points for larger volumes of passengers. This limitation left us with 1,155,266 data points on unique station stops to use for the analyses in this study.

3. RESULTS AND DISCUSSION

The frequency of passenger volumes for different precipitation levels is shown in Figure 2. Visual analysis indicates that the level of precipitation does not influence the number of people boarding a train, except for conditions with heavy rain during peak hours. The results from the Chi-square goodness of fit tests, Table 2, show that there is no statistically significant effect of precipitation on the frequency of passenger off-peak hours. We do find a statistically significant effect of conditions with moderate precipitation on the frequency of passenger volumes during peak hours. Although the visual analysis indicates a difference, this is thus not found to be statistically significant.

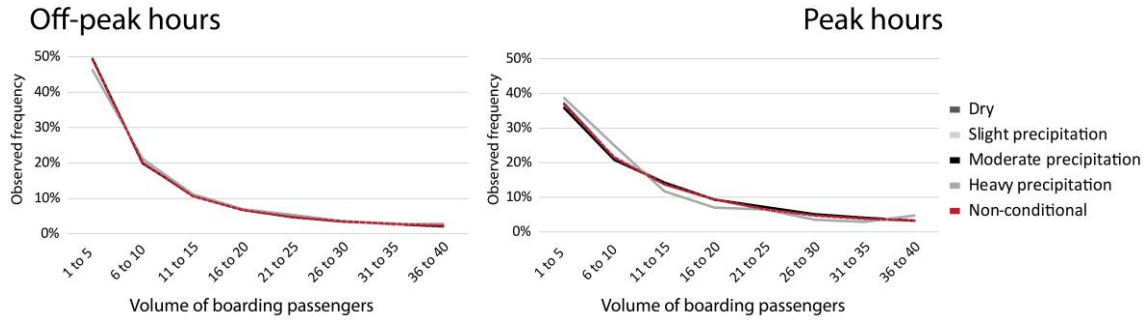


Figure 2: Frequency of passenger volumes under different precipitation levels.

Table 2: Chi-square goodness of fit test results for the difference in frequency of passenger volumes under different precipitation levels. Significant results indicated with an *, significance level with Bonferroni correction $p < 0.013$ ($0.05/4 = 0.013$).

Precipitation level	Off-peak hours		Peak hours	
	<i>Chi-square</i>	<i>P-value</i>	<i>Chi-square</i>	<i>P-value</i>
Dry	1	0.999	2	0.931
Slight precipitation	5	0.693	14	0.048
Moderate precipitation	8	0.321	18	0.012*
Heavy precipitation	5	0.637	9	0.224

Figure 3 shows the frequency of passenger volumes given different temperatures. The visual analysis shows relatively larger differences in the frequency of passengers during off-peak hours, where the frequency of larger passenger volumes increases as the temperature increases. The frequency of passengers during peak hours is found to be less affected by changes in temperature. The results of the Chi-square goodness of fit tests, Table 3, show that statistically significant effects of the temperature occur during peak hours, but not for all brackets under consideration. In contrast to this, we find that changes in the temperature have a statistically significant effect on the frequency of passenger volumes for all brackets under consideration during off-peak hours. The higher Chi-square values suggest that the effect is strongest when temperatures exceed 20 degrees Celsius.

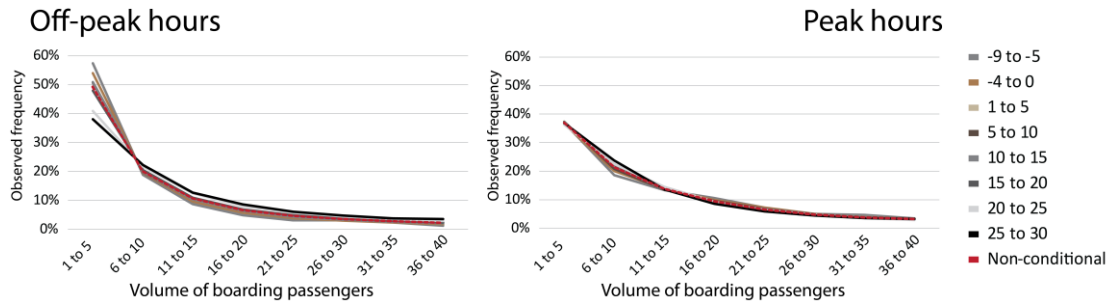


Figure 3: Frequency of passenger volumes under different temperature levels.

Table 3: Chi-square goodness of fit test results for the difference in frequency of passenger volumes under different temperature levels. Significant results indicated with an *, significance level with Bonferroni correction $p < 0.006$ ($0.05/9 = 0.006$).

Temperature (Celsius)	Off-peak hours		Peak hours	
	Chi-square	P-value	Chi-square	P-value
-9 to -5	221	0.000*	22	0.002*
-4 to 0	586	0.000*	34	0.000*
1-5	264	0.000*	16	0.028
5-10	178	0.000*	27	0.000*
10-15	224	0.000*	11	0.132
15-20	112	0.000*	23	0.002*
20-25	2653	0.000*	46	0.000*
25-30	1505	0.000*	31	0.000*

The frequency of passenger volumes in relation to the wind speed is shown in Figure 4. As with the changes in temperature, we find the largest effect of wind speed on the frequency of passengers to occur during off-peak hours. Larger passenger volumes are found to be somewhat more common as the wind speed increases, both during peak and off-peak hours. The results from the Chi-square goodness of fit test, Table 4, reveal that there are statistically significant effects of the wind speed on the frequency of passengers with the effect being stronger during off-peak hours.

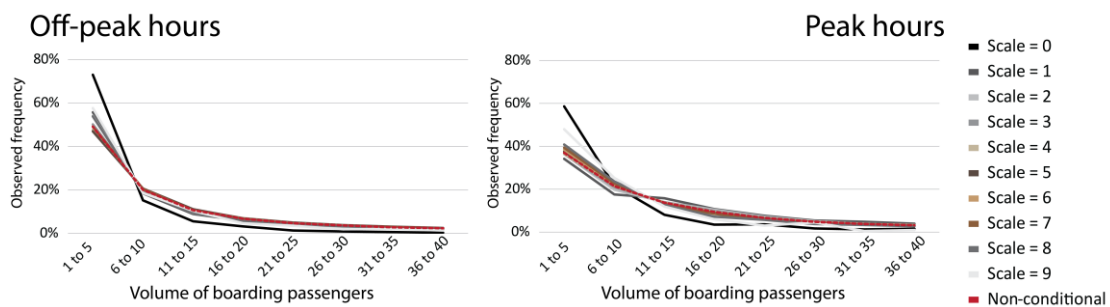


Figure 4: Frequency of passenger volumes under different wind speeds.

Table 4: Chi-square goodness of fit test results for the difference in frequency of passenger volumes under different wind speeds. Significant results indicated with an *, significance level with Bonferroni correction $p < 0.005$ ($0.05/10 = 0.005$).

Beaufort scale	Off-peak hours		Peak hours	
	<i>Chi-square</i>	<i>P-value</i>	<i>Chi-square</i>	<i>P-value</i>
Calm (0)	699	0.000*	117	0.000*
Light air (1)	391	0.000*	92	0.000*
Light breeze (2)	920	0.000*	77	0.000*
Gentle breeze (3)	76	0.000*	30	0.000*
Moderate breeze (4)	269	0.000*	19	0.009
Fresh breeze (5)	417	0.000*	33	0.000*
Strong breeze (6)	5	0.670	41	0.000*
High wind (7)	6	0.490	42	0.000*
Gale (8)	52	0.000*	34	0.000*
Strong gale (9)	30	0.000*	24	0.000*

The findings we present here show that weather phenomena have statistically significant effects on the volume of boarding passengers for commuter trains. Larger Chi-square values for the effects are found during off-peak compared to peak hours for both the changes in temperature and wind speed. Weather conditions thus have a stronger effect on passenger demand during off-peak hours, which is in line with previous remarks regarding the flexibility of trips made by Cools et al. (2010) and Liu et al. (2014). The lack of an effect of precipitation levels on the frequency of passenger volumes can be regarded as somewhat surprising. Previous studies have highlighted a shift from active, mostly open-air, modes towards covered modes such as public transport and private vehicles as a result of increased precipitation (Sabir, 2011). In terms of practical implications, our findings indicate that weather should be taken into account both during tactical and operational planning. The change in passenger volumes can be incorporated into the timetable by adapting dwell time during prolonged periods of warm weather during off-peak hours. A similar notion can be made concerning long-term rolling stock circulation plans. On the operational level, rolling stock circulations can be adapted based on weather predictions. To account for the increase of passengers during periods with higher wind speeds, for example. These changes can also be incorporated into demand modelling and real-time rescheduling problems.

The study we present here does come with some limitations. To account for different responses to weather variables between commuters and non-commuters we split our observations based on the departure time. Although this can be considered a good proxy for the type of trip, it does not capture a shift in travel times for commuters. More detailed information on the individual traveller is required to reveal such responses to changes in the weather. Another limitation is the distance between the weather stations and railway stations and the granularity of the observations in terms of time. As mentioned by Creemers et al. (2015) some caution is advised when making use of data from point sources such as weather stations when analysing changes in space and time. Although hourly weather data is relatively detailed, weather changes can be volatile and change within the hour. Nevertheless, we argue that the data used for this study are of sufficient quality to make meaningful inferences. The study we present here is limited to focusing on the volume of boarding passengers. Although one of the main factors affecting the duration of dwell times, other aspects such as the ratio and spread of passengers influence dwell times as well. Future studies could include these passenger flow characteristics, which can help improve real-time travel time predictions. In addition to this, station design characteristics such as the available roof coverage can be included in a future study. Doing so allows for a more in-depth understanding of the effect of weather conditions, and can highlight whether station characteristics play a role in the effects found here.

4. CONCLUSIONS

The study we present here focuses on changes in the volume of boarding passengers under different weather conditions. The relationship between weather phenomena and the ridership of trains has not received much attention previously. Our study aims to fill this research gap. To do so we make use of automatic passenger count and weather data on over a million unique station stops, spanning two years, in the Southern region of Scania in Sweden. We find that changes in the level of precipitation, in general, do not affect the volume of boarding passengers except for conditions with moderate precipitation during peak hours. Both changes in temperature and wind speeds are found to have a statistically significant effect on the frequency distribution of the volume of boarding passengers. These changes are different depending on whether the trip takes place during peak hours or off-peak hours. Passenger volumes for trips made during off-peak hours are found to be more susceptible to changes in both wind and temperature. The effects we find can serve as input during both tactical and operational planning, by guiding timetable principles as well as rolling stock circulation plans, and serve as input during real-time rescheduling problems.

REFERENCES

- Böcker, L., Dijst, M., & Prillwitz, J. (2013). Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review. *Transport Reviews*, 33(1), 71–91. <https://doi.org/10.1080/01441647.2012.747114>
- Brons, M. R. E., & Rietveld, P. (2008). *Rail mode, access mode and station choice: The impact of travel time unreliability* [Research Report for the project TRANSUMO BTK]. VU University.
- Cox, T., Houdmont, J., & Griffiths, A. (2006). Rail passenger crowding, stress, health and safety in Britain. *Transportation Research Part A: Policy and Practice*, 40(3), 244–258. <https://doi.org/10.1016/j.tra.2005.07.001>
- Creemers, L., Wets, G., & Cools, M. (2015). Meteorological variation in daily travel behaviour: Evidence from revealed preference data from the Netherlands. *Theoretical and Applied Climatology*, 120(1–2), 183–194. <https://doi.org/10.1007/s00704-014-1169-0>
- Koetse, M. J., & Rietveld, P. (2009). The impact of climate change and weather on transport: An overview of empirical findings. *Transportation Research Part D: Transport and Environment*, 14(3), 205–221. <https://doi.org/10.1016/j.trd.2008.12.004>
- Kuipers, R. A., & Palmqvist, C.-W. (2022). Passenger Volumes and Dwell Times for Commuter Trains: A Case Study Using Automatic Passenger Count Data in Stockholm. *Applied Sciences*, 12(12), 5983. <https://doi.org/10.3390/app12125983>
- Lepage, S., & Morency, C. (2021). Impact of Weather, Activities, and Service Disruptions on Transportation Demand. *Transportation Research Record: Journal of the Transportation Research Board*, 2675(1), 294–304. <https://doi.org/10.1177/0361198120966326>
- Li, D., Daamen, W., & Goverde, R. M. P. (2016). Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station. *Journal of Advanced Transportation*, 50(5), 877–896. <https://doi.org/10.1002/atr.1380>

- McDonald, J. H. (2014). *The Handbook of Biological Statistics*. <http://www.biostathandbook.com/>
- Sabir, M. (2011). *Weather and Travel Behaviour* [PhD thesis]. Vrije Universiteit Amsterdam.
- Spinney, J. E. L., & Millward, H. (2011). Weather impacts on leisure activities in Halifax, Nova Scotia. *International Journal of Biometeorology*, 55(2), 133–145. <https://doi.org/10.1007/s00484-010-0319-z>
- van Loon, R., Rietveld, P., & Brons, M. (2011). Travel-time reliability impacts on railway passenger demand: A revealed preference analysis. *Journal of Transport Geography*, 19(4), 917–925. <https://doi.org/10.1016/j.jtrangeo.2010.11.009>

Welfare impacts of remote and flexible working policies in the bottleneck model

Takara Sakai^{*1}, Takashi Akamatsu¹, and Koki Satsukawa¹

¹Graduate School of Information Sciences, Tohoku University, Miyagi, Japan

SHORT SUMMARY

This study investigates the effects of remote and flexible working styles on traffic congestion. We first formulate an integrated equilibrium model simultaneously considering the working style, official work start time, and departure time choice of workers via an extension of the bottleneck model. Subsequently, we derive the equivalent optimization problem of the equilibrium problem as linear programming (LP) and demonstrate that we can obtain an analytical solution to the LP. This analytical solution enables us to assess the effects of remote and flexible working on social surplus and queueing loss. By comparing various situations, we show that implementing remote and flexible work causes higher queueing loss with an equal social surplus than implementing only remote work. Finally, we propose an integrated road management scheme that includes dynamic pricing to prevent this paradoxical phenomenon and efficiently implements remote and flexible working.

Keywords: remote working, flexible working, working style choice, departure time choice, bottleneck model

1 INTRODUCTION

New working styles, which are different from the conventional style of commuting to an office at a designated time, have become widespread, mainly owing to the COVID-19 pandemic. These new working styles include teleworking, staggered work hours, and flexible work. Each working style not only decreases the number of opportunities available to workers to contact each other in the office but also reduces or disperses the commuting demand during peak-periods. Therefore, promoting these working styles could potentially reduce commute-related congestion. However, many companies and workers have adopted these ways of working independently of the aim of the urban transportation system, which is to relieve traffic congestion. Under these circumstances, whether remote work and staggered work hours contribute to reducing congestion and their effect on road congestion are not well understood.

Many studies examined the relationship between traffic congestion and working style based on Vickrey's bottleneck model. Mun & Yonekawa, 2006; Fosgerau & Small, 2017; Takayama, 2015 formulated peak-period congestion models based on the bottleneck model and developed models describing the choice of firms and workers to adopt fixed or flexible schedules. In addition, many analyses of bottleneck models that consider the heterogeneity of preferred arrival times can be interpreted as modeling flexible or staggered work hours (e.g., Hendrickson & Kocur, 1981; Lindsey, 2004; Lindsey et al., 2019). Zhang et al. (2005) studied the trade-off between teleworking and office working, and they considered the elastic travel demand by incorporating teleworking as an alternative. Gubins & Verhoef (2011) analyzed the welfare effects of teleworking on road traffic congestion in the context of Vickrey's model. These studies highlight the positive effects of flexible and remote working on traffic congestion and social welfare. However, these results are based on analyses that consider only one type of working style. To develop effective congestion reduction policies for the current situation where multiple new working styles are widespread, we need to investigate how the relationship between these working styles affects traffic congestion.

This study investigates the effects of remote and flexible work on traffic congestion. We first formulate an integrated equilibrium model that accounts for flexible and remote work simultaneously via an extension of the bottleneck model. Subsequently, we derive the equivalent optimization problem of the equilibrium problem as linear programming (LP) and demonstrate that we obtain an analytical solution to the LP. This analytical solution enables us to assess the impacts of remote

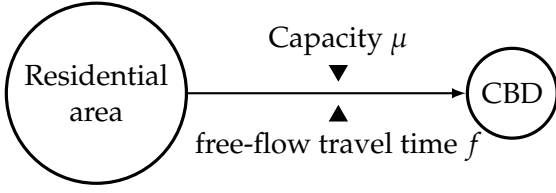


Figure 1: Network.

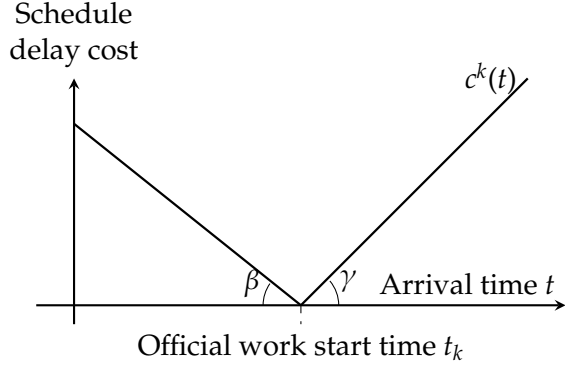


Figure 2: Schedule delay cost function.

and flexible work on social surplus and queueing loss. We present the following facts by comparing four scenarios: - no policies, remote working, flexible working, and both remote and flexible working:

- Implementing remote work causes lower queueing loss and higher social surplus than implementing remote work.
- Implementing flexible work causes lower queueing loss and higher social surplus than not implementing flexible work.
- Implementing remote and flexible work causes higher queueing loss with equal social surplus than implementing only remote work.

The third fact describes a paradoxical phenomenon in which the simultaneous introduction of flexible and remote working may increase queueing losses. This phenomenon may be interesting and important for road management.

The remainder of this paper is structured as follows: Section 2 introduces the integrated equilibrium model and derives its equivalent optimization problem. Section 3 presents an analytical approach to solving the problem and the effects of remote and flexible working arrangements on traffic congestion. Finally, Section 4 concludes the study and discusses future studies.

2 MODEL

Consider a city that consists of a central business district (CBD) and residential area connected by a freeway (Figure 1). This freeway has a single bottleneck with capacity μ , and its free-flow travel time is denoted by f . If the arrival rates of the workers at the bottleneck exceed bottleneck capacity, a queue develops. To model queueing congestion, we use first-in-first-out (FIFO) and a point queue in which vehicles have no physical length, as in standard bottleneck models.

All workers reside in the residential area. The workers are treated as a continuum, and the total mass Q is a given constant. The firm, located in the CBD, offers two working styles for workers: office and remote work. In addition, the firm allows K official work start (OWS) times $\{t^1, \dots, t^K\}$ for office workers. Each worker can choose between office and remote work. If they choose the former, they must commute and choose the OWS time and the actual departure/arrival time. If they choose the latter, they work from home without commuting to the office.

The trip costs for each commuter (i.e., office worker) are assumed additively separable into free-flow travel, queueing delay, and schedule delay costs. The schedule delay cost is defined as the difference between the actual arrival times and OWS time at the office. We assume piecewise linearity in the schedule delay cost function, which is expressed as follows (Figure 2):

$$c_k(t) \equiv \begin{cases} \beta(t_k - t) & \text{if } t_k < t \\ \gamma(t - t_k) & \text{if } t_k \geq t \end{cases} \quad \forall k \in \mathcal{K}, \quad \forall t \in \mathcal{T}. \quad (1)$$

The trip cost of a commuter whose destination arrival time is t and OWS time is t^k is defined as follows:

$$C_k(t) \equiv c_k(t) + \alpha(w(t) + f) \quad \forall k \in \mathcal{K}, \quad \forall t \in \mathcal{T}. \quad (2)$$

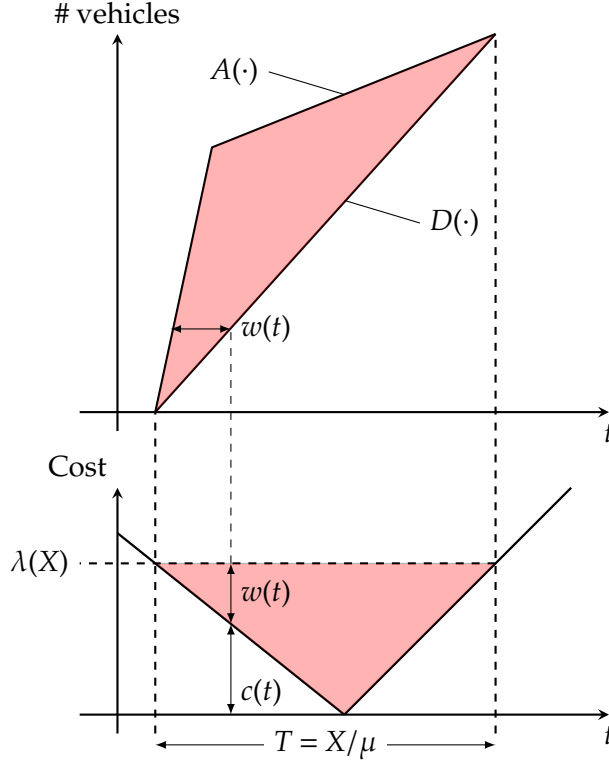


Figure 3: Commuting equilibrium with $K = 1$.

where $w(t)$ is the queuing delay experienced at the bottleneck by commuters with CBD arrival time t . Parameters f and α represent the free-flow travel time and value of time, respectively. This study assumes that $f = 0$ and $\alpha = 1$ for all commuters.

Workers choose their working styles to maximize their own utility. Utility of an office worker whose OWS time is t_k and actual CBD arrival time is t determine $\theta_O - C_k(t)$, where θ_O represents the wage parameter for office workers. In contrast, the utility of a remote worker is θ_R , where θ_R represents the wage parameter for remote workers. We assume that office workers have higher productivity and wages, i.e., $\theta_O > \theta_R$. In the equilibrium resulting from these choices of the workers, the following properties hold: no worker can reduce their utility by unilaterally changing their working style, and no commuter can reduce their commuting costs by unilaterally changing their destination arrival time.

The equilibrium problem can be formulated as a linear complementary problem comprising five equilibrium conditions. First, the worker conservation condition is expressed as follows:

$$\text{(Worker conservation)} \quad Q_O + Q_R = Q, \quad (3)$$

where Q_O and Q_R represent the numbers of office and remote workers, respectively. Second, the equilibrium condition for workers in terms of working style is expressed as

$$\text{(equilibrium condition for office workers)} \quad \begin{cases} \rho = \theta_O - \lambda & \text{if } Q_O > 0 \\ \rho \geq \theta_O - \lambda & \text{if } Q_O = 0 \end{cases} \quad (4)$$

$$\text{(equilibrium condition for remote workers)} \quad \begin{cases} \rho = \theta_R & \text{if } Q_R > 0 \\ \rho \geq \theta_R & \text{if } Q_R = 0 \end{cases} \quad (5)$$

where ρ represents the equilibrium utility, and λ represents the equilibrium commuting cost. Third, the office worker (commuter flow) conservation conditions for the commuting demands must satisfy

$$\text{(commuter conservation)} \quad \sum_{k \in \mathcal{K}} \int_{t \in \mathcal{T}} q_k(t) dt = Q_O \quad \forall k \in \mathcal{K}, \quad (6)$$

where $q_k(t) \geq 0$ is the arrival flow rate of commuters whose CBD arrival time is t and OWS time is

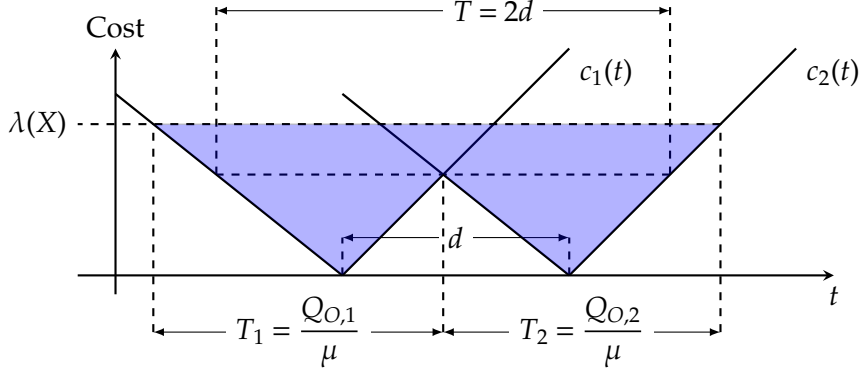


Figure 4: Commuting equilibrium with $K = 2$.

t^k . Fourth, the equilibrium condition for commuters is expressed as

$$\begin{aligned}
 \text{(departure time choice condition)} \quad & \begin{cases} \lambda = c_k(t) + \alpha(w(t) + f) & \text{if } q_k(t) > 0 \\ \lambda \leq c_k(t) + \alpha(w(t) + f) & \text{if } q_k(t) = 0 \end{cases} \\
 & \forall k \in \mathcal{K}, \quad \forall t \in \mathcal{T}. \quad (7)
 \end{aligned}$$

Fifth, the queuing condition at the bottleneck is expressed as follows (Akamatsu et al., 2021):

$$\begin{aligned}
 \text{(bottleneck queuing condition)} \quad & \begin{cases} \sum_{k \in \mathcal{K}} q_k(t) = \mu & \text{if } w(t) > 0 \\ \sum_{k \in \mathcal{K}} q_k(t) \leq \mu & \text{if } w(t) = 0 \end{cases} \\
 & \forall t \in \mathcal{T}. \quad (8)
 \end{aligned}$$

The equilibrium state represents the collection of variables $\{Q_O, Q_R, \lambda, \rho, w(t), q_k(t)\}$ that satisfy Eqs. (3) to (8).

Based on the methods reported by Iryo & Yoshii (2007); Akamatsu et al. (2021), an optimization problem can obtain the aforementioned equilibrium state. Specifically, we can derive the equilibrium number of office/remote workers and commuter arrival flow patterns as the optimal solution to the following linear programming. In addition, the equilibrium cost pattern, including the equilibrium utility, equilibrium commuting cost, and queuing delay pattern, can also be obtained as the optimal Lagrange variable of the problem.

$$\min_{Q_O, Q_R, \lambda, q_k(t) \geq 0} \cdot \sum_{k \in \mathcal{K}} \int_{t \in \mathcal{T}} c_k(t) q_k(t) dt - \theta_O Q_O - \theta_R Q_R \quad (9)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}} q_k(t) \leq \mu \quad \forall t \in \mathcal{T} \quad [w(t)], \quad (10)$$

$$\sum_{k \in \mathcal{K}} \int_{t \in \mathcal{T}} q_k(t) dt = Q_O \quad [\lambda], \quad (11)$$

$$Q_O + Q_R = Q \quad [\rho], \quad (12)$$

where the variables inside the square brackets represent the Lagrangian multipliers for each constraint. We refer to this problem as the equivalent optimization problem. Equivalency can be proven by comparing the first-order conditions with the equilibrium conditions.

3 WELFARE IMPACT OF REMOTE AND FLEXIBLE WORKING POLICIES

The model formulated in the previous section describes the equilibrium under various situations by fixing certain parameters. In this section, we first develop a general approach to obtaining equilibrium by solving the equivalent optimization problem. We then derive the equilibrium in the following four situations by fixing the appropriate parameters (This paper assumes the two OWS times differ by d in flexible working situation, i.e., $K = 2$ and $t_2 - t_1 = d$):

- Scenario (1): No policies ($\theta_R = -\infty$, $K = 1$)

- Scenario (2): Remote working ($K = 1$)
- Scenario (3): Flexible working ($\theta_R = -\infty, K = 2$)
- Scenario (4): Remote working and flexible working ($K = 2$)

Finally, by comparing equilibrium in scenarios (1)-(4), we analyze the effect of the interaction between remote and flexible working on traffic congestion.

To derive the equilibrium solutions systematically, let us decompose the equivalent optimization problem into a hierarchical optimization problem consisting of the master problem and sub-problem. The master problem determines the mass of office workers Q_O and remote workers Q_R . The sub-problem determines the commuting departure flow patterns $q_k(t)$, in which Q_O is the given parameter. We here introduce the hierarchical optimization problem below.

$$[\text{Master}] \quad \min_{Q_O, Q_R \geq 0} \cdot TC(Q_O) - \theta_O Q_O - \theta_R Q_R \quad (13)$$

$$\text{s.t.} \quad Q_O + Q_R = Q \quad [\rho]. \quad (14)$$

$$[\text{Sub}] \quad TC(Q_O) \equiv \min_{\{q_k(t)\} \geq 0} \cdot \sum_{k \in \mathcal{K}} \int_{t \in \mathcal{T}} c_k(t) q_k(t) dt \quad (15)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}} q_k(t) \leq \mu \quad \forall t \in \mathcal{T} \quad [w(t)], \quad (16)$$

$$\sum_{k \in \mathcal{K}} \int_{t \in \mathcal{T}} q_k(t) dt = Q_O \quad [\lambda]. \quad (17)$$

By solving the sub-problem, we first derive the equilibrium commuting cost as a function of the total commuting demand (the number of office workers) determined by the master problem. Subsequently, we solve the master problem using the equilibrium commuting cost.

The sub-problem has the same structure as the single bottleneck model, and we can analytically solve it by using the condition that all commuters incur the same commuting costs ($w(t) + c_k(t)$) in the equilibrium state, as shown in Figures 3 and 4. Figure 3 illustrates the equilibrium arrival/departure flow pattern at the bottleneck and the equilibrium cost pattern with $K = 1$ when the total commuting demand is X . Similarly, Figure 4 illustrates the equilibrium commuting cost pattern with $K = 2$ when the total commuting demand is X . In both figures, $\lambda(X)$ represents the equilibrium commuting cost derived as follows:

$$\lambda(X) = \begin{cases} \frac{X}{\mu} \delta & \text{if } K = 1 \\ \min \left\{ \frac{X}{2\mu} \delta, \frac{X}{\mu} \delta - d\delta \right\} & \text{if } K = 2 \end{cases}, \quad (18)$$

where $\delta = \beta\gamma / (\beta + \gamma)$.

Based on the equilibrium commuting cost $\lambda(X)$, we obtain the solutions to the master problem, i.e., Q_O and Q_R . Specifically, from the optimality condition of the master problem, we find that the following relationship holds in equilibrium:

$$Q_O = \min\{\lambda^{-1}(\theta_O - \theta_R), Q\}, \quad (19)$$

where $\lambda^{-1}(\cdot)$ is the inverse function of $\lambda(X)$. Thus, “the solution of equation $\theta_O - \theta_R = \lambda(X)$ ” or “ Q ”, whichever is smaller, is the total commuting demand Q_O . By combining Q_O and the conservation condition of the workers, we find the number of remote workers Q_R . This mathematical approach corresponds to a graphical approach determining the intersection between $\theta_O - \lambda(X)$ and θ_R . Figure 5 illustrates the intersection between θ_R and $\theta_O - \lambda(X | K)$. In Figure 5, intersections $I^{(1)}$, $I^{(2)}$, $I^{(3)}$, and $I^{(4)}$ represent equilibrium states for each scenario. By finding the coordinates of these

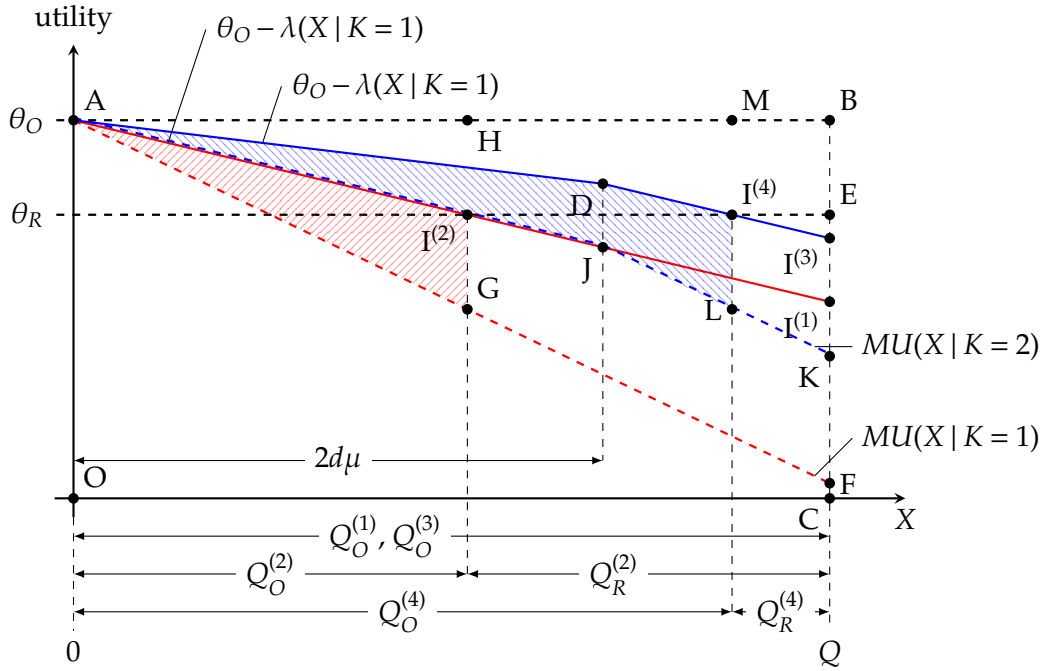


Figure 5: Equilibrium number of office/remote workers in scenarios (1)-(4).

intersections, the analytical solution for each scenario can be derived as follows¹:

$$\text{Scenario (1): } Q_O^{(1)} = Q, \quad Q_R^{(1)} = 0, \quad \lambda^{(1)} = \frac{Q}{\mu}\delta, \quad \rho^{(1)} = \theta_O - \lambda^{(1)}, \quad (20)$$

$$\text{Scenario (2): } Q_O^{(2)} = \frac{\mu}{\delta}(\theta_O - \theta_R), \quad Q_R^{(2)} = Q - Q_O^{(2)}, \quad \lambda^{(2)} = \theta_O - \theta_R, \quad \rho^{(2)} = \theta_R, \quad (21)$$

$$\text{Scenario (3): } Q_O^{(3)} = Q, \quad Q_R^{(3)} = 0, \quad \lambda^{(3)} = \frac{Q}{\mu}\delta - d\delta, \quad \rho^{(3)} = \theta_O - \lambda^{(3)}, \quad (22)$$

$$\text{Scenario (4): } Q_O^{(4)} = \frac{2\mu}{\delta}(\theta_O - \theta_R), \quad Q_R^{(4)} = Q - Q_O^{(4)}, \quad \lambda^{(4)} = \theta_O - \theta_R, \quad \rho^{(4)} = \theta_R. \quad (23)$$

Using Eqs. (20) to (23), we can calculate the social surplus and queuing loss in each scenario. The social surplus and queuing loss correspond to the areas of the regions in Figure 5, as shown in Table 1. In Figure 5, $MU(X | K = 1)$ and $MU(X | K = 2)$ are social marginal utility functions for $K = 1$ and $K = 2$, respectively.

Table 1: Comparison of the scenarios

	Social Surplus SS	Queuing Loss QL	Schedule Loss SL
Scenario(1)	$SS^{(1)} = AFKO$	$QL^{(1)} = AI^{(1)}F$	$SL^{(1)} = ABI^{(1)}$
Scenario(2)	$SS^{(2)} = AGI^{(2)}ECO$	$QL^{(2)} = AI^{(2)}G$	$SL^{(2)} = AHI^{(2)}$
Scenario(3)	$SS^{(3)} = AJKCO$	$QL^{(3)} = ADI^{(3)}KJ$	$SL^{(3)} = ABI^{(3)}D$
Scenario(4)	$SS^{(4)} = AJLI^{(4)}ECO$	$QL^{(4)} = ADI^{(4)}LJ$	$SL^{(4)} = AMI^{(4)}D$

By comparing these areas, we obtain the following theorem:

Theorem 3.1 (Remote work effect). Implementing remote work causes lower queuing loss and higher social surplus than not implementing remote work.

$$QL^{(2)} < QL^{(1)}, \quad SS^{(2)} > SS^{(1)} \quad (24)$$

¹Strictly speaking, the cases must be divided according to θ_R . Because of space limitations, this study shows only the most standard cases.

Theorem 3.2 (Flexible work effect). Implementing flexible work causes lower queuing loss and higher social surplus than not implementing flexible work.

$$QL^{(3)} < QL^{(1)}, \quad SS^{(3)} > SS^{(1)} \quad (25)$$

Theorem 3.3 (Remote work paradox). Implementing remote and flexible work causes higher queuing loss with equal social surplus than implementing only remote work.

$$QL^{(2)} < QL^{(4)}, \quad SS^{(4)} = SS^{(2)} \quad (26)$$

Theorem 3.1 and Theorem 3.2 state the positive effects of remote and flexible work, respectively. In contrast, Theorem 3.3 describes a paradoxical phenomenon in which the simultaneous implementation of flexible and remote working may cause higher queuing losses. This paradox is due to the induced demand created by decreasing commuting costs for flexible working. That is, even if the flexible working spreads out the OWS time, additional traffic congestion around each OWS time occurs because the utility of office workers is equal to (to balance) the utility of remote workers in the equilibrium state.

This paradox of the relationship between flexible and remote working styles can be prevented by implementing a dynamic pricing scheme. Specifically, the road manager imposes a time-varying congestion toll that mimics the queuing delay pattern in scenario (4) to the commuters. In equilibrium under this dynamic pricing, bottleneck congestion is completely eliminated, and the utility of all workers maintain the same as before the implementation of the pricing. Since this pricing scheme gives the road manager toll revenue equal to queuing loss, the utility of the workers increases if the toll revenue is appropriately returned to workers. These results imply that combining multiple policies can affect efficiency and highlight the importance of analyzing the combined effects of multiple policies.

4 CONCLUSION

This study investigated the impact of flexible and remote work on traffic congestion using the bottleneck model. We first formulated an integrated equilibrium model that simultaneously considered three worker choices: office/remote work, OWS time, and departure time choices. Furthermore, we elucidated that the equilibrium model had an equivalent optimization problem. We derived the equilibrium solution using a hierarchical decomposition approach and calculated the social surplus and queuing loss under various situations. Comparing these situations showed a paradoxical phenomenon in which queuing losses were higher while social surplus remained constant when flexible and remote work policies were considered simultaneously than when only remote work policies were considered. We conclude that the cause of this paradox is the demand induced by reduced commuting costs for flexible working.

Future studies may investigate a more general model and reveal the impact of the relationships between various working styles on traffic congestion. Specifically, we may extend the network structure to corridor networks with multiple residential areas. This extension analysis allows us to understand the impact of the relationships between flexible and remote working styles on the location choice of workers. We also must model the agglomeration economics of office work and investigate the policy effects more precisely, considering the trade-offs between the office and remote work.

ACKNOWLEDGEMENT

This study was supported by JSPS KAKENHI (Grant Numbers JP20J21744, JP21H01448, and JP20K14843).

REFERENCES

- Akamatsu, T., Wada, K., Iryo, T., & Hayashi, S. (2021, June). A new look at departure time choice equilibrium models with heterogeneous users. *Transportation Research Part B: Methodological*, 148, 152–182.
- Fosgerau, M., & Small, K. (2017, May). Endogenous scheduling preferences and congestion. *International economic review*, 58(2), 585–615.
- Gubins, S., & Verhoef, E. T. (2011, July). *Teleworking and congestion: A dynamic bottleneck analysis*.
- Hendrickson, C., & Kocur, G. (1981, February). Schedule delay and departure time decisions in a deterministic model. *Transportation Science*, 15(1), 62–77.
- Iryo, T., & Yoshii, T. (2007). Equivalent optimization problem for finding equilibrium in the bottleneck model with departure time choices. In *4th IMA international conference on mathematics in Transport Institute of mathematics and its applications*. trid.trb.org.
- Lindsey, R. (2004, August). Existence, uniqueness, and trip cost function properties of user equilibrium in the bottleneck model with multiple user classes. *Transportation Science*, 38(3), 293–314.
- Lindsey, R., de Palma, A., & Silva, H. E. (2019, June). Equilibrium in a dynamic model of congestion with large and small users. *Transportation Research Part B: Methodological*, 124, 82–107.
- Mun, S.-I., & Yonekawa, M. (2006). Flextime, traffic congestion and urban productivity. *Journal of Transport Economics and Policy (JTPE)*, 40(3), 329–358.
- Takayama, Y. (2015, November). Bottleneck congestion and distribution of work start times: The economics of staggered work hours revisited. *Transportation Research Part B: Methodological*, 81, 830–847.
- Zhang, X., Yang, H., Huang, H.-J., & Zhang, H. M. (2005, January). Integrated scheduling of daily work activities and morning–evening commutes with bottleneck congestion. *Transportation Research Part A: Policy and Practice*, 39(1), 41–60.

Willingness to wait with real-time crowding information in urban public transport – before vs. after COVID-19 pandemic

Arkadiusz **Drabicki***¹, Oded **Cats**², Rafał **Kucharski**³

¹ MSc, Department of Transportation Systems, Cracow University of Technology, Poland

² PhD, Department of Transport and Planning, Delft University of Technology, Netherlands

³ PhD, Group of Machine Learning Research, Jagiellonian University, Poland

SHORT SUMMARY

Passenger overcrowding is a major problem influencing travel behaviour in urban public transport (PT). Its relevance has been presumably shaped by the covid-19 pandemic impacts, which have to be yet fully understood. Real-time crowding information (RTCI) is therefore potentially instrumental in the post-covid recovery of PT ridership. This study investigates the willingness to wait (WTW) to reduce overcrowding in urban PT, analysing pre- vs. post-covid travel behaviour attitudes. Ex-post stated-preference data and (subsequently estimated) choice models indicate, compared to pre-covid findings, a higher propensity to skip overcrowded services with RTCI on seats available in later departures, and lower utility of RTCI on moderately crowded services. The WTW with RTCI seems to have become less dependent on individual characteristics and more prominent for time-critical (obligatory) trips as well. Implications of these findings are discussed in final study sections.

Keywords: public transport; passenger crowding; discrete choice modelling; real-time crowding information; RTCI; willingness to wait; COVID-19

1. INTRODUCTION

Travel behaviour in public transport (PT) systems is shaped by multiple factors, including passenger overcrowding - a recurrent problem in high-density urban transportation networks. Rising (over)crowding influences the relative (un)attractiveness, comfort and safety perceptions of PT travel options. Moreover, it may lead to system failure in oversaturated PT networks - manifested in form of denied boardings, demand-supply feedback deteriorations etc. (Tirachini et al, 2013; Cats, West, Eliasson 2016). Crowding impacts upon travel behaviour have been widely studied in state-of-the-art literature (e.g. Wardman and Whelan, 2011; Tirachini et al, 2013; Hoercher et al, 2017; Yap et al, 2018) and references cited therein).

Meanwhile, the recent COVID-19 pandemic has profoundly impacted the urban PT systems worldwide, with yet lingering ramifications for passengers' travel behaviour (Tirachini and Cats, 2020; Gkiotsalitis, Cats 2021). Its experience has exacerbated the perceived risks of travelling in higher crowding conditions. The emerging stream of literature underlines that crowding valuations have increased by up to 25% compared to pre-pandemic levels (Cho and Park, 2021; Shelat et al, 2022b). Various user groups (e.g. female and elderly travellers, but also across wider population) have become much more apprehensive of exposure to PT overcrowding, especially if considering the associated infection risks (Shelat et al, 2022a; Aghabayk et al, 2021; Basnak et al, 2022).

As a consequence, PT ridership is often still struggling to recover to pre-pandemic levels. This underpins the need for tools addressing the post-pandemic travel safety concerns in urban PT systems. A prospective ITS-based solution emerges in form of providing **real-time crowding information (RTCI)** on (in-vehicle) passenger loads of urban PT services. The RTCI can help passengers mitigate the PT overcrowding experience, as demonstrated in simulation studies (Nuzolo et al, 2016; Noursalehi et al, 2021; Drabicki et al, 2021, 2022). Moreover, RTCI provision may incite a novel (and not fully understood yet) travel behaviour phenomenon in form of willingness to wait (WTW) to reduce overcrowding. Namely, passengers may opt to skip deliberately an (over)crowded departure and wait for a less-crowded service at the same PT stop. This notion has been hitherto explored in a number of studies, though conducted either before the onset of COVID-19 pandemic (Kim et al, 2009; Kroes et al, 2014; Preston et al, 2017; Drabicki et al, 2023), or afterwards (Shelat et al, 2022a; Singh et al, 2023). To the best of our knowledge, no comparative analysis of pre- vs. post-covid changes in passengers’ WTW with RTCI is yet available in state-of-the-art literature.

The objective of this study is to contribute to the above research gap with a pre- vs. post-covid investigation of passengers’ WTW to reduce overcrowding with RTCI in urban PT journeys. To this end, we design a stated-preference (SP) survey and estimate discrete choice models, using a mixed logit specification. A comparison of our investigation results conducted in two stages – pre-covid (2019) and post-covid (2022) – highlights the shifts arising in passengers’ WTW preferences. Findings and conclusions from our research underline how the RTCI may play an even more instrumental and effective role in post-pandemic urban PT networks.

2. METHODOLOGY

Research investigation has been conducted in two data collection stages:

- ‘pre-covid’ investigation – in March 2019 (sample size: $n = 377$ respondents),
- ‘post-covid’ investigation – in May 2022 (sample size: $n = 424$ respondents).

Both stages of our research investigation follow the methodology elaborated in (Drabicki et al, 2023), summarized below. Moreover, timing of both survey stages ensured that the case-study urban PT system was free of any disruptions or social distancing restrictions, which could have impaired the reliability and plausibility of collected responses.

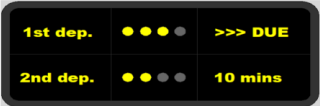
Which of these departures would you choose?		
Trip context:	in-vehicle time:	20 [mins]
	departures every:	10 [mins]
	need to arrive on-time:	YES
	trip purpose:	home → work
	using this route:	2 - 4 [days/week]
no seats available, but can stand comfortably		
seats available		
Answer:	<input type="checkbox"/> departure 1 - NOW	<input type="checkbox"/> departure 2 - WAIT

Figure 1: Example of the SP choice experiment question.

The SP survey has been designed as a field survey, conducted among passengers at urban PT stops within the city of Krakow (Poland), with completion time no greater than 3 – 5 minutes. This allowed a vast majority of interviewees to answer it successfully. A randomized sampling strategy aimed to reflect the typical demand pattern of urban PT users in Krakow.

The SP survey started with questions on the respondents' current trip context – trip motivation, propensity to arrive on-time (at destination), elapsed journey time, service frequency. This followed then with core part of SP survey – i.e., a panel series of stated choice experiments. Respondents were presented with a hypothetical RTCI on the 2 nearest bus/tram departures from their current stop, and were asked to choose the preferred travel option: first departure – departing now, but with higher (over)crowding on-board, vs. second departure – less crowded but requiring a 5- or a 10-minute wait. All the remaining trip characteristics remained equal for both travel options, as specified by respondents themselves. Socio-demographic information (age, gender, PT usage frequency) was also collected for statistical purposes.

In total, each respondent was presented with 6 stated-choice scenarios. Each scenario was set up as a combination of 2 possible waiting time values (5 or 10 minutes) and 3 possible RTCI values of the nearest 2 departures:

- 1st dep. – moderately crowded (RTCI level 3), 2nd dep. – seats available (RTCI level 2),
- 1st dep. – highly overcrowded (RTCI lvl. 4), 2nd dep. – moderately crowded (RTCI lvl. 3),
- 1st dep. – highly overcrowded (RTCI lvl. 4), 2nd dep. – seats available (RTCI lvl. 2),

The SP answers serve next as basis for estimation of discrete choice models of the WTW with RTCI. Our WTW experimental setup essentially reflects a binary choice context, formulated in accordance with the random utility maximization (RUM) theory (Ben-Akiva, Lerman 1985). Choice probability is evaluated between the utility U_1 of boarding now the first departure vs. utility U_2 associated with waiting and boarding (later) the second departure. Once we assume a fixed reference utility rate of $U_1 = 0$, the utility $U_2 = U_{WTW}$ expresses then the relative (dis)utility associated with deliberately waiting for a second, less-crowded PT departure:

$$P(U_2) = \frac{\exp(U_2)}{\exp(U_1) + \exp(U_2)} = \frac{\exp(U_{WTW})}{1 + \exp(U_{WTW})} \quad (1)$$

This U_{WTW} utility is composed of systematic utility V_{WTW} plus a random error term ε_{WTW} (normally distributed with mean equal to zero). The systematic WTW utility is, in turn, a function of a vector of taste (preference) co-efficients β_k and corresponding attribute values X_k :

$$V_{WTW} = \sum_{k=1}^K \beta_k * X_k \quad (2)$$

The attribute set K contains various trip- and population-related characteristics, valid for a given choice situation. We hereby test various model specifications, utilizing the mixed logit (MXL) approach. The MXL allows to capture unobserved heterogeneity effects in our panel survey data. The default MXL model specification consists of RTCI utility $\beta_{RTCI}^s * \delta_{RTCI}^s$ (represented by case-specific dummy variables, denoted by RTCI levels of both PT departures in the choice scenario s) and waiting time (dis)utility $\beta_{wt} * t_{wt}$. The MXL mixing distribution is applied to the waiting time co-efficient, assumed to be a normally distributed variable $\beta_{wt}(\mu, \sigma)$:

$$V_{WTW} = \beta_{RTCI}^{3-2} * \delta_{RTCI}^{3-2} + \beta_{RTCI}^{4-3} * \delta_{RTCI}^{4-3} + \beta_{RTCI}^{4-2} * \delta_{RTCI}^{4-2} + \beta_{wt}(\mu, \sigma) * t_{wt} \quad (3)$$

3. RESULTS AND DISCUSSION

Starting from descriptive statistics, a comparative analysis of both survey stages (2019 vs. 2022) reveals substantial differences in reported WTW with RTCI (Fig. 2). The 2019 pre-covid survey indicates a substantial propensity to avoid high overcrowding (RTCI level 4) in the first vehicle, regardless of crowding level inside the second departure. Ca. 75% of respondents would choose the less-crowded option arriving in 5 [mins], and for a 10-minute wait this rate oscillates around 45%. Meanwhile, the post-covid (2022) findings show that WTW decisions are more dependent on crowding level of the second departure as well. If this involves moderate standing crowding (RTCI level 3), ca. 57% of respondents would wait for 5 [mins], and just above 20% for 10 [mins]. However, if the second arrival has seats available (RTCI level 2), these rates surge to over 90% and 55%, respectively. In the third (alternative) scenario, passengers' willingness to avoid a moderately crowded vehicle (RTCI level 3) in exchange for seat availability (RTCI level 2) remains analogous across both survey samples. Approx. 30% of respondents would accept a 5-minute wait, and ca. 10% would wait for 10 [mins].

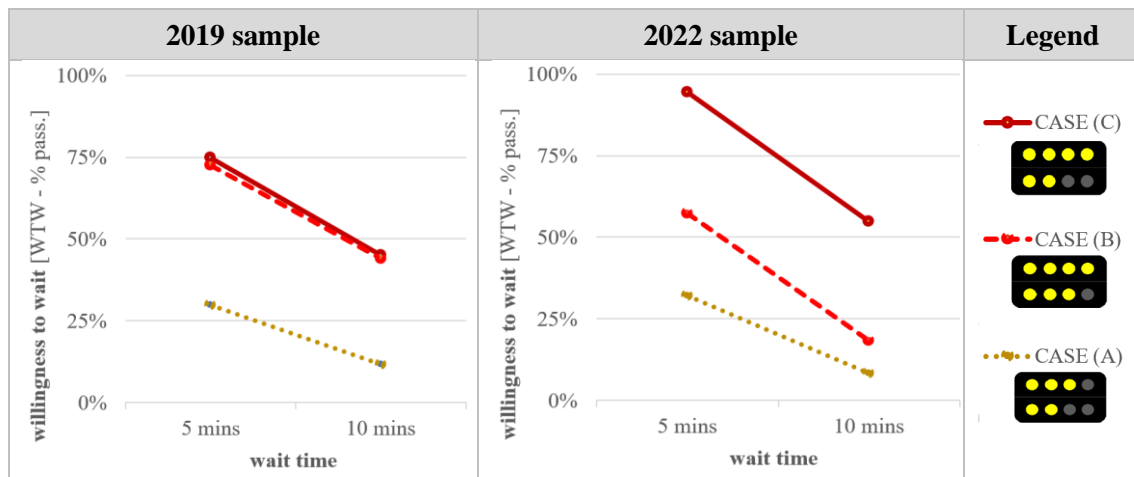


Figure 2: Survey results – overall WTW with RTCI in the pre- (left) vs. post-covid (right) sample.

The post-covid evaluation also indicates a variable and generally lower influence of trip- and demographic-related factors upon WTW with RTCI. For example, the pre-covid survey pointed towards a more substantial role of trip time-criticality, i.e. propensity to arrive on-time at the destination (Drabicki et al, 2023). In the post-covid sample, passengers' preferences are more uniform across the whole sample. This suggests relatively higher WTW probability for time-critical trips, especially with abrupt difference in crowding conditions between consecutive departures (i.e. the RTCI level 4 vs. 2 'scenario'). Otherwise, respondent's age remains a relevant choice factor, as the WTW increases for those aged 50 - 65 years, and even further for the 65+ year-olds.

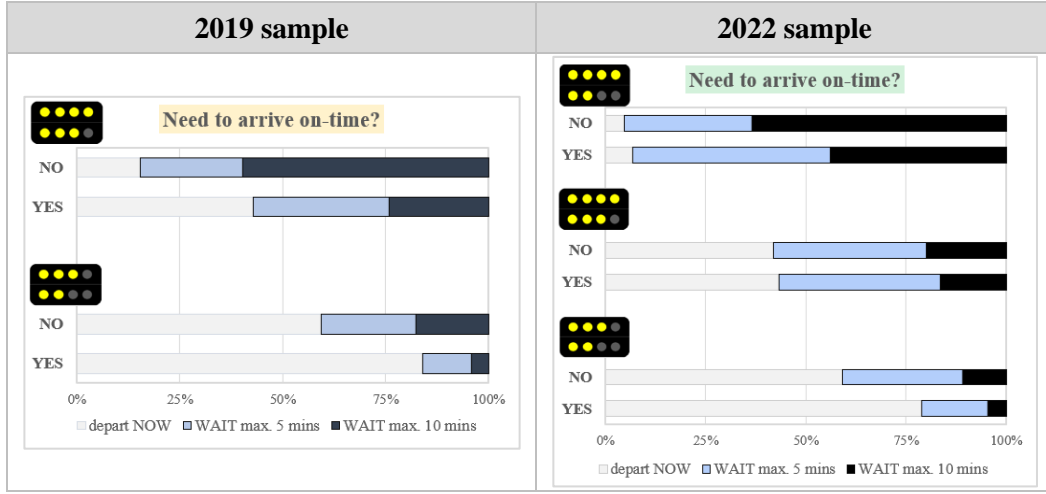


Figure 3: Survey results – reported WTW with RTCI, distinguished by trip time-criticality.

Survey outputs serve then for MXL model estimation purposes (Tab. 1). Model co-efficients essentially represent the β^{s}_{RTCI} RTCI utility versus the β_{wt} waiting time (dis)utility rate. In other words, the former denotes the expected utility gain from reducing the on-board overcrowding if choosing the later departure, whilst the latter reflects the perceived unit utility loss per minute of waiting time (hence the negative symbol). All the co-efficients are statistically significant at $p < 0.05$, and panel effects are included in mixing distribution applied to β_{wt} . In general, post-covid data shows a relative increase in RTCI utility in case of information on seats available in the second departure. This is especially valid if the first departure implies high overcrowding conditions β^{4-2}_{RTCI} . On the other hand, when RTCI indicates only the possibility of decreasing the standing crowding (β^{4-3}_{RTCI}), its utility seems lower compared to pre-covid estimates.




Table 1: Mixed logit estimation results of the WTW with RTCI.

Coefficients mean, (<i>t-stat.</i>)		2019 sample	2022 sample
β^{3-2}_{RTCI}		1.828 (8.09)	2.144 (10.11)
β^{4-3}_{RTCI}		5.294 (15.90)	3.540 (15.14)
β^{4-2}_{RTCI}		5.510 (11.46)	6.598 (18.92)
β_{wt}	μ	- 0.705 (11.82)	- 0.628 (16.62)
	σ	0.286 (6.33)	0.244 (13.10)
initial log-likelihood:		- 1380.9	- 1734.3
final log-likelihood:		- 816.5	- 1141.6
LL ratio test:		1128.4	1243.4
adjusted rho-square:		0.396	0.349
sample size:		377	424

Discrete choice modelling results can be further used to compute the ratio of marginal utilities of RTCI and waiting time. This yields the average acceptable waiting times for a second, less-crowded PT departure (Tab. 2). Seemingly, acceptable WTW thresholds have increased in the




post-covid period with RTCI indicating the possibility to mitigate standing crowding in the second departure. On average, waiting times have risen from (roughly) 3 to 4 [mins] for the β^{3-2}_{RTCI} case, and even more from 9 to 12 [mins] for the β^{3-2}_{RTCI} case. In contrast, mean acceptable waiting time has dropped from ca. 9 to 7 [mins] for the β^{4-3}_{RTCI} case.

Table 2: Acceptable waiting times in [mins] with the RTCI, acc. to MXL modelling results.

RTCI case	2019 sample (mean, (st. dev.))	2022 sample (mean, (st. dev.))
	3.2 (3.5)	4.2 (3.7)
	8.9 (5.8)	6.8 (4.7)
	9.2 (5.9)	12.1 (6.4)

Based on above findings, we compute the value-of-time crowding multipliers for a sample 15-minute journey in urban PT network (Tab. 3). These are calculated according to the methodology in Preston et al (2017) and Drabicki et al (2023)). The post-covid crowding multipliers are, likewise, higher for the β^{4-2}_{RTCI} and β^{3-2}_{RTCI} cases and lower for the β^{4-3}_{RTCI} cases. Relative changes versus the pre-covid rates amount to ca. 5 – 10%. This compares similar (albeit somewhat lower) to findings in the recent literature (Cho and Park, 2021).

Table 3: Crowding multipliers for a 15-minute PT journey, acc. to MXL modelling results.

RTCI case	2019 sample (mean)	2022 sample (mean)
	1.21	1.28
	1.59	1.45
	1.62	1.81

4. CONCLUSIONS

This study contributes with a pre- vs. post-covid analysis of passengers’ willingness to wait (WTW) to reduce overcrowding in urban PT networks. Based on survey data from 2019 and 2022 conducted in Krakow (Poland), we observe how the prospective utility of real-time crowding information (RTCI) has changed in the aftermath of COVID-19 pandemic. While pre-covid estimates showed that the WTW was primarily driven by sole possibility of avoiding overcrowding in the first departure, the expected crowding reduction in the second departure wields greater influence upon post-covid passengers’ preferences. Compared to pre-covid data, fewer passengers are willing to skip a highly overcrowded vehicle and wait for a moderately crowded one. However, the WTW probability has substantially increased with seats available in the later departure. While seat availability itself may not be a crucial decision factor in short-range, urban PT trips, these findings suggest that passengers nowadays attach relatively greater weight to the RTCI content and displayed difference(s) between crowding levels of PT vehicles.

Our findings also underpin the prospective application of RTCI systems in future urban PT networks. The WTW incited by RTCI provision can lead to more balanced distribution of passenger loads between PT vehicles. This will improve operational efficiency and decrease exposure to overcrowding. Hence, timely and accurate RTCI can reassure the crowding-aware passengers about current travel conditions. Moreover, it can serve as an effective travel demand management tool, playing thus an instrumental role in post-covid recovery of PT ridership.

REFERENCES

- Aghabayk, K., Esmailpour, J. and Shiwakoti, N., 2021. Effects of COVID-19 on rail passengers' crowding perceptions. *Transportation Research Part A: Policy and Practice*, 154, pp.186-202.
- Basnak, P., Giesen, R. and Muñoz, J.C., 2022. Estimation of crowding factors for public transport during the COVID-19 pandemic in Santiago, Chile. *Transp. Res. Part A: Policy and Practice*, 159, pp.140-156.
- Ben-Akiva, M.E., Lerman, S.R. and Lerman, S.R., 1985. *Discrete choice analysis: theory and application to travel demand* (Vol. 9). MIT press.
- Cats, O., West, J. and Eliasson, J., 2016. A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. *Transportation Research Part B: Methodological*, 89, pp.43-57.
- Cho, S.H. and Park, H.C., 2021. Exploring the behaviour change of crowding impedance on public transit due to COVID-19 pandemic: before and after comparison. *Transportation Letters*, 13(5-6), pp.367-374.
- Drabicki, A., Kucharski, R., Cats, O. and Szarata, A., 2021. Modelling the effects of real-time crowding information in urban public transport systems. *Transportmetrica A: Transport Science*, 17(4), pp.675-713.
- Drabicki, A., Kucharski, R. and Cats, O., 2022. Mitigating bus bunching with real-time crowding information. *Transportation*, pp.1-28.
- Drabicki, A., Cats, O., Kucharski, R., Fonzone, A. and Szarata, A., 2023. Should I stay or should I board? Willingness to wait with real-time crowding information in urban public transport. *Research in Transportation Business and Management* (in press).
- Gkiotsalitis, K. and Cats, O., 2021. Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. *Transport Reviews*, 41(3), pp.374-392.
- Hörcher, D., Graham, D.J. and Anderson, R.J., 2017. Crowding cost estimation with large scale smart card and vehicle location data. *Transportation Research Part B: Methodological*, 95, pp.105-125.
- Kim, J.K., Lee, B. and Oh, S., 2009. Passenger choice models for analysis of impacts of real-time bus information on crowdedness. *Transportation research record*, 2112(1), pp.119-126.
- Kroes, E., Kouwenhoven, M., Debrincat, L. and Pauget, N., 2014. Value of crowding on public transport in île-de-France, France. *Transportation Research Record*, 2417(1), pp.37-45.
- Noursalehi, P., Koutsopoulos, H.N. and Zhao, J., 2021. Predictive decision support platform and its application in crowding prediction and passenger information generation. *Transportation Research Part C: Emerging Technologies*, 129, p.103139.
- Nuzzolo, A., Crisalli, U., Comi, A. and Rosati, L., 2016. A mesoscopic transit assignment model including real-time predictive information on crowding. *Journal of ITS*, 20(4), pp.316-333.
- Preston, J., Pritchard, J. and Waterson, B., 2017. Train overcrowding: investigation of the provision of better information to mitigate the issues. *Transportation research record*, 2649(1), pp.1-8.
- Shelat, S., Cats, O. and van Cranenburgh, S., 2022. Traveller behaviour in public transport in the early stages of the COVID-19 pandemic in the Netherlands. *Transportation Research Part A: Policy and Practice*, 159, pp.357-371.
- Shelat, S., Van De Wiel, T., Molin, E., van Lint, J.W.C. and Cats, O., 2022. Analysing the impact of COVID-19 risk perceptions on route choice behaviour in train networks. *Plos one*, 17(3), p.e0264805.
- Singh, J., de Almeida Correia, G.H., van Wee, B. and Barbour, N., 2023. Change in departure time for a train trip to avoid crowding during the COVID-19 pandemic: a latent class study in the Netherlands. *Transportation Research Part A: Policy and Practice*, p.103628.

Tirachini, A., Hensher, D.A. and Rose, J.M., 2013. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice*, 53, pp.36-52.

Tirachini, A. and Cats, O., 2020. COVID-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, 22(1), pp.1-21.

Wardman, M. and Whelan, G., 2011. Twenty years of rail crowding valuation studies: evidence and lessons from British experience. *Transport reviews*, 31(3), pp.379-398.

Yap, M., Cats, O. and van Arem, B., 2020. Crowding valuation in urban tram and bus transportation based on smart card data. *Transportmetrica A: Transport Science*, 16(1), pp.23-42.

Quantum choice models leap out of the laboratory: capturing real-world behavioural change.

Thomas O. Hancock, Charisma F. Choudhury¹, Joan Walker², and Stephane Hess¹

¹Choice Modelling Centre and Institute for Transport Studies, University of Leeds, UK

²Civil and Environmental Engineering, University of California, Berkeley, USA

SHORT SUMMARY

Quantum choice models have been recently introduced to travel behaviour modelling, showing significant promise in explaining preferential change as a result of a change in choice context. However, thus far, quantum choice models have only been applied to stated preference (SP) data. This paper focusses on the application of these models to revealed preference (RP) data and the methodological adaptations required to deal with the increased complexities that come with RP data. Using 2-week travel diaries from 273 individuals in/near Leeds, UK, we demonstrate that quantum choice models can effectively capture the impact of behavioural nudges used to shift travellers towards greener travel modes. The results demonstrate that the provision of feedback on behaviour relative to those of a similar demographic reinforces current behaviour: travellers who make more green choices become greener, whilst the converse is true for travellers who use make less green choices than average.

Keywords: Behavioural change; Quantum choice model; Quantum rotations; Revealed preference data.

1 INTRODUCTION

Quantum probability, first developed in theoretical physics, has recently made the transition into cognitive psychology, where it has been used to explain the impact of question order, fallacies in decision-making, and other effects that were previously difficult to explain using classical models (Pothos & Busemeyer, 2022). In our previous work, we demonstrated that quantum probability theory could also make the transition into choice modelling through the development of quantum choice models, and be used to explain route choice behaviour (Hancock et al., 2020b) as well as moral choice behaviour (Hancock et al., 2020a). In particular, the results revealed that the quantum choice models can efficiently capture the effect of a change in choice context through ‘quantum rotations’ in stated preference choice settings.

However, the applicability of the methods in real-world settings, where the context effects are more ‘fuzzy’ and difficult to capture in the data, remains unclear. Furthermore, the model has not previously been applied to datasets where there are different numbers of alternatives in different choice contexts, nor has it been applied to choice contexts where there are different numbers of attributes for different alternatives.

In this paper, we aim to address these research gaps by testing quantum choice models on revealed preference (RP) data from a natural experiment involving behavioural interventions/nudges potentially leading to shifts in the choice context. We aim to test whether quantum choice models better capture behavioural changes arising from interventions and whether they lead to similar or different behavioural insights in comparison to typical modelling approaches.

The remainder of this paper is arranged as follows. First, the current and standard implementations of quantum choice models are described, before new theories and extensions for the model are detailed. Next, we discuss a real-world case study of mode choice with nudge interventions in the UK, before presenting conclusions and possibilities for future research.

2 METHODOLOGY

First we present an overview of quantum concepts. Next, we outline the ‘quantum amplitude model’, specifying different levels of model complexity. Finally, we demonstrate how the model

captures a change in choice context.

Overview of the quantum concepts

As is the case for any choice model, a quantum choice model captures the quality of each alternative and produces a set of probabilities for the likelihood of choosing each alternative. The representation of the alternatives, however, is very different to that of standard econometric choice models. Under quantum choice models, a series of choices (i.e. where, how and when to travel) is represented by a ‘Hilbert space’, H . This space is similar to Euclidean space, except that it uses complex numbers and is of dimension J , where J represents the number of alternatives. Each individual choice (where, how and when) is itself represented by a subspace, L , within the Hilbert space. Each subspace, L , is created by a basis of vectors (e.g. $|x_1\rangle, |x_2\rangle, \dots$) where each vector represents an alternative and has real and imaginary parts. In line with previous notation used in quantum cognition, we use ‘bra-ket’ notation (Trueblood & Busemeyer, 2011), under which a column vector in a Hilbert space is represented by a ‘ket’ vector, $|\cdot\rangle$, with the corresponding row vector (with each element being complex conjugated) a ‘bra’ vector, $\langle\cdot|$. A representation of a subspace is given in Figure 1.

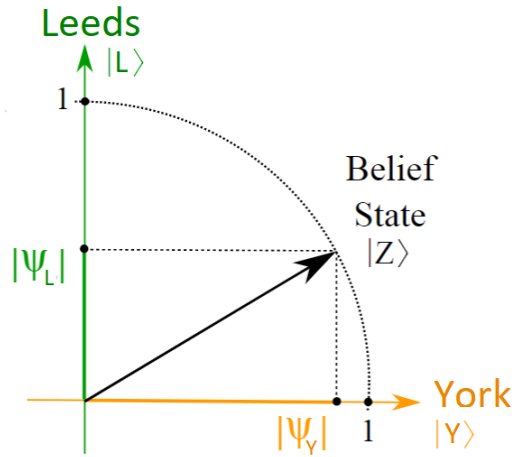


Figure 1: An example of a subspace representing a destination choice, where the individual can choose to travel to Leeds or York (adapted from Hancock et al. 2020b). Leeds is represented by the vector $|L\rangle$ and York is represented by $|Y\rangle$, with the decision-maker’s current preference represented by $|Z\rangle$.

The preferences of the individual (e.g. to choose to travel to York or Leeds in Figure 1) are contained within a ‘state’ vector, $|Z\rangle$, which is a superposition of the alternative vectors and represents the decision-maker’s propensity to choose different alternatives. By ‘choosing’ an alternative, the decision-maker’s ‘state’ vector aligns with the vector representing the chosen alternative through a projection from the state vector. This results in the possibility of complex interactions between pairs of choices (e.g. where and how) as the choice of an alternative for where can impact the probability of choosing different alternatives for how (e.g. a shopper who wishes to travel by car may choose to travel to a place with better parking facilities). A visualisation of this is provided in Figure 2.

The shift in probability is a direct result of the decision being made from a different state. In the example provided in Figure 2, the projection length for Leeds, $|\psi_L|$, is increased if the decision-maker chooses their destination from $|T\rangle$ instead of $|Z\rangle$.

The quantum amplitude model (QAM)

The quantum amplitude model (QAM, Hancock et al. 2020b) defines probabilities for alternatives based on the use of projection lengths for each alternative, $|\psi_i|$. As the state vector representing the decision-maker’s preferences is normalised to be of unit length and the vectors for the different alternatives are orthonormal, the probability for choosing the alternative can be set to its squared projection length. In Euclidean space, this is a direct result of the use of Pythagoras’ Theorem, see

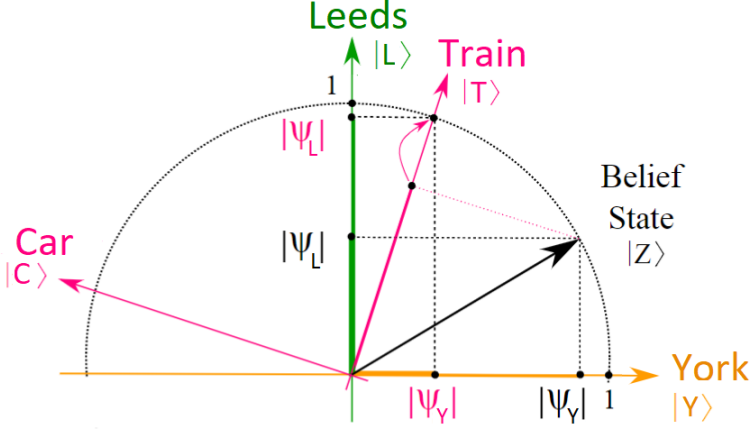


Figure 2: An example of a pair of choices. The choice of either travelling by car or train impacts the probability of choosing to travel to York or Leeds.

Figure 1. In Hilbert space, the use of complex conjugates (Hancock et al., 2020b) ensures that this property holds. Thus, we have:

$$P(j) = |\psi_j|^2 = \psi_j \cdot \text{Conj}(\psi_j), \quad (1)$$

where $|\psi_j|$ is the norm of the ‘amplitude’ and we have:

$$\sum_j^J |\psi_j|^2 = 1. \quad (2)$$

Consequently, we can build a quantum choice model simply by defining a normalised state vector, which, like a standard econometric choice model, can be based on the attributes of the alternatives. The state vector components (with respect to each alternative vector) can be based on utility function differences (e.g. regret functions (Chorus, 2010) work particularly well, Hancock et al. 2020b), as this results in larger projection lengths for alternatives with favourable attributes, which increases the probability of these alternatives being chosen. Thus, under our most basic specification for a quantum choice model, for alternative i , for individual n in choice task t , we define:

$$|\psi_{int}| = \delta_i + \sum_{k=1}^K \sum_{j \neq i}^J (A_{i,j} \cdot \ln(1 + e^{\beta_k(x_{intk} - x_{jntk})})) \quad (3)$$

where δ_i is equivalent to alternative specific constants in econometric choice models, and captures the underlying bias towards an alternative. β_k captures the relative importance of attribute k , the elements x represent the attribute levels and $A_{i,j}$ is an indicator variable taking the value of 1 if alternatives i and j are both available in the current choice context.

In the work using quantum choice models thus far, the number of possible alternatives has always been constant across the dataset. This is not necessarily the case in RP data, where, for example, car availability will impact the number of possible alternatives. As regret functions are used, a correction may be required (Van Cranenburgh et al., 2015). Thus, the second QAM model is defined using average differences through the use of J_{nt} , the number of available alternatives for individual n in choice task t :

$$|\psi_{int}| = \delta_i + \sum_{k=1}^K \sum_{j \neq i}^J (A_{i,j} \cdot \ln(1 + e^{\beta_k(x_{intk} - x_{jntk})})) / (J_{nt} - 1). \quad (4)$$

To properly exploit the fact that the quantum choice model can operate in complex space (with real and imaginary parts), our third version of the model incorporates ‘complex phases’, $e^{i\theta_k}$, to introduce imaginary parts to the amplitudes. This aims to capture the fact that different attributes may not be considered equivalently during the choice deliberation process (Hancock et al., 2020a). Thus we have:

$$|\psi_{int}\rangle = \delta_i + \sum_{k=1}^K \sum_{j \neq i}^J (e^{i\theta_k} \cdot A_{i,j} \cdot \ln(1 + e^{\beta_k(x_{intk} - x_{jntk})))) / (J_{nt} - 1). \quad (5)$$

Finally, further flexibility can be introduced by introducing complex phase multipliers ($e^{i\theta_i}$) to the alternative specific constants:

$$|\psi_{int}\rangle = \delta_i \cdot e^{i\theta_i} + \sum_{k=1}^K \sum_{j \neq i}^J (e^{i\theta_k} \cdot A_{i,j} \cdot \ln(1 + e^{\beta_k(x_{intk} - x_{jntk})))) / (J_{nt} - 1), \quad (6)$$

where the phase for one alternative is fixed to ensure identification.

Capturing a change of choice context

Under quantum choice models, the impact of a nudge can be captured by a ‘quantum rotation,’ which shifts the state vector. Preferences for two separate tasks/actions are represented by the same state vector (i.e. we would have the same state vector representing preferences/the propensity to choose which mode to travel, and where an individual chooses to travel, see Figure 1). Vectors representing alternatives in the first task are not the same as vectors representing alternatives in the second, but are not necessarily orthonormal across tasks, meaning that making one decision (and thus ‘projecting’ onto the vector represented by the first chosen alternative), will shift the decision-maker’s state vector, but ultimately mean that they can still choose any of the alternatives in the second task. Consequently, the probabilities of choosing the different alternatives in the second task may change (i.e. trying to be more environmentally friendly will reduce the likelihood of choosing to drive). This concept of ‘entanglement’ is the key factor in driving the improvement found by adopting quantum models within case studies on ordering effects in cognitive psychology. In the current context, behavioural nudges are represented mathematically through the use of a quantum rotation to the state vector prior to the decision-maker choosing an alternative. For example, an individual may implicitly answer the question ‘Am I an environmentally friendly person?’ if they are nudged towards making greener choices. The more environmentally conscious the individual, the larger the rotation and consequential shift towards greener alternatives. Mathematically, a rotation maps amplitudes onto new ones through multiplication:

$$\psi'_j = \sum_{i=1}^n U_{j,i}^* \cdot \psi_j, \quad (7)$$

where the unitary matrix U must have certain properties to ensure that a new belief state vector maintains unit length. In the case in hand, the use of *Hamiltonians*, (H), control the change of the belief state vector according the dynamics of the Schrödinger equation:

$$U(\tau) = e^{-iH\tau}, \quad (8)$$

where H itself must be Hermitian to ensure that the time evolution (over τ) will conserve the normalisation of the belief state. The size of H will depend on the number of available alternatives. For two alternatives, it is defined:

$$H = \begin{bmatrix} 1 & \delta_{H_{12}} \\ \text{Conj}(\delta_{H_{12}}) & -1 \end{bmatrix}, \quad (9)$$

where $\delta_{H_{ij}}$ is a parameter to be estimated that governs the degree to which an individual’s preference shifts between alternatives i and j .

3 EMPIRICAL APPLICATION

Data

The dataset tested is from the ‘Decisions’ survey (Calastri et al., 2020). In this survey, 273 participants completed a 2-week travel diary with the use of Rmove. An average of 2.5 trips per day across respondents resulted in a total of 9,254 trips.

At the end of the first week, some participants were given feedback as to whether they use more or less CO2 and burn more or fewer calories than other participants with similar demographics.

In total, 128 participants received no feedback, 79 received feedback regarding their own travel, and 66 received feedback showing their own travel relative to those of a similar demographic. An example of the feedback received is shown in Figure 3.

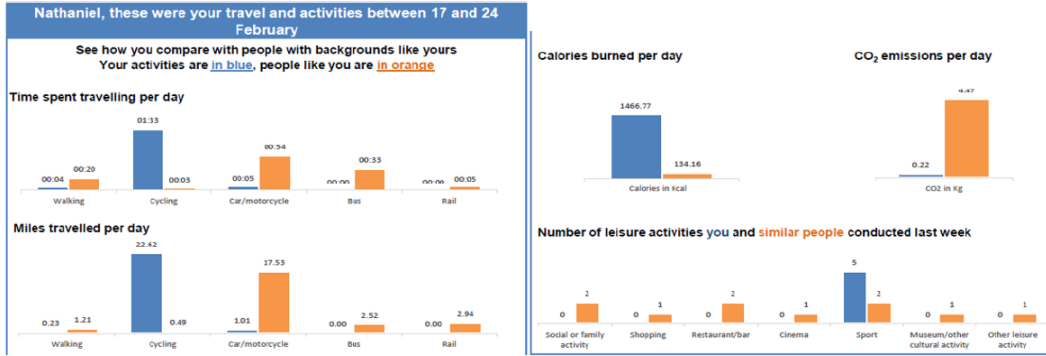


Figure 3: An example of the feedback given after one week to a participant who sees their relative performance in comparison to individuals of a similar demographic.

Corrections for correlation and endogeneity issues suggested that there was no significant impact of feedback for individuals aggregating over all their trips (Palma et al., 2019). However, analysis at the trip-level was not possible at this point as level-of-service data was only obtained for this dataset following later work (Tsoleridis et al., 2022). In the current context, quantum choice models are applied across the choice tasks with ‘quantum rotations’ aiming to capture the ‘nudge’ towards making greener choices if the individual has received feedback.

Model specification

There are up to six possible mode alternatives, listed below in Table 1. The multinomial logit model that is used for comparison is linear sum of alternative specific constants (δ), mode-specific in-vehicle travel times for non-active modes ($\beta_{mode-ivt}$), mode-specific out-of-vehicle times for active modes and public transport ($\beta_{mode-ovt}$), and costs for non-active modes (β_{cost}).

Table 1: The full set of parameters used in the base models.

	Car	Bus	Rail	Taxi	Cycle	Walk
Alternative specific constant	δ_{car}	δ_{bus}	δ_{rail}	δ_{taxi}	δ_{cycle}	δ_{walk}
In-vehicle travel time	$\beta_{car-ivt}$	$\beta_{bus-ivt}$	$\beta_{rail-ivt}$	$\beta_{taxi-ivt}$		
Out-of-vehicle travel time		$\beta_{bus-ovt}$	$\beta_{rail-ovt}$		$\beta_{cycle-ovt}$	$\beta_{walk-ovt}$
Fare/Cost	β_{cost}	β_{cost}	β_{cost}	β_{cost}		

To capture behavioural change in the MNL models, we use shifted alternative specific constants, thus we have:

$$\delta'_{i,n} = \delta_i + \delta_{co2,i} \cdot \zeta_{co2,n} + \delta_{cal,i} \cdot \zeta_{cal,n}, \quad (10)$$

where $\zeta_{co2,n} = 1$ and $\zeta_{cal,n} = 1$ if individual n has received feedback stating that they have used more CO2, and burned more calories, respectively, than individuals of a similar demographic.

For the quantum choice models, the behavioural change is incorporated through the use of Hamiltonians (see Equations 7-9). We focus on the shift towards or away from the use of car, thus have:

$$H_{car,i} = \delta_{co2,i} \cdot \zeta_{co2,n} + \delta_{cal,i} \cdot \zeta_{cal,n}, \quad (11)$$

with the other non-diagonal elements of the Hamiltonian set to a value of zero.

Results for base models

The full model results for the base MNL model and the base quantum amplitude model (QAM, based on Equation 3) are given in Table 2.

For the MNL model, we see reasonable willingness-to-pay outputs, with out-of-vehicle times generally worse than in-vehicle times, and the coefficient for walking time most negative, as would

Table 2: Model outputs and parameter estimates for the base MNL and quantum models.

Model	Base MNL			QAM (Eq. 3)		
free pars.	14			15		
LL(0)	-11,176			-11,176		
LL	-4,057			-4,160		
adj. ρ^2	0.6358			0.6265		
BIC	8,369			8,593		
par.	est	rob. t-rat	VTT (£/hr)	est	rob. t-rat	RI (£/hr)
δ_{car}	0.000	NA		36.383	3.59	
δ_{bus}	-2.908	-11.62		2.512	2.67	
δ_{rail}	-2.850	-7.30		-3.536	-2.69	
δ_{taxi}	-3.780	-10.21		3.522	3.16	
δ_{cycle}	-4.239	-9.68		-4.600	-4.23	
δ_{walk}	0.365	1.47		12.122	3.41	
$\beta_{car-ivt}$	-0.144	-9.27	41.42	-0.395	-3.31	26.93
$\beta_{bus-ivt}$	-0.049	-7.21	14.05	-0.214	-3.49	14.58
$\beta_{bus-ovt}$	-0.097	-3.28	27.81	-0.584	-4.07	39.86
$\beta_{rail-ivt}$	-0.048	-3.29	13.67	-0.157	-2.26	10.71
$\beta_{rail-ovt}$	-0.079	-5.84	22.62	-0.321	-3.53	21.88
$\beta_{taxi-ivt}$	-0.100	-3.32	28.79	-0.916	-3.55	62.50
$\beta_{cycle-ovt}$	-0.080	-5.53	22.95	-0.528	-2.96	36.03
$\beta_{walk-ovt}$	-0.154	-11.94	44.10	-0.347	-4.13	23.65
β_{cost}	-0.209	-6.71		-0.879	-3.53	

be expected. Though the quantum choice model does not produce WTP outputs, the relative importance of attributes can be estimated. These values, though within the same range as MNL WTPs, show some large discrepancies. For example, the relative importance for taxi travel time is approximately doubled, whilst it is halved for walking time.

Next, we compare model fits across the QAM models with updated specifications based on Equations 4-6. These model results are given in Table 3.

Table 3: Model fits from the different versions of quantum amplitude models

Model	MNL	QAM (Eq.3)	QAM (Eq.4)	QAM (Eq.5)	QAM (Eq.6)
Free pars.	14	15	15	17	22
LL(0)	-11,176	-11,176	-11,176	-11,176	-11,176
LL	-4,057	-4,160	-4,158	-4,059	-4,032
adj. ρ^2	0.6358	0.6265	0.6266	0.6353	0.6373
BIC	8,369	8,593	8,589	8,428	8,465

The inclusion of complex phases (Equation 5) results in a substantial improvement in the performance of the QAM model, bringing it in line with the result from the MNL model, though at the cost of 3 additional parameters. The addition of a further 5 complex phases results in the QAM model outperforming MNL by 25 log-likelihood units with a better adjusted ρ^2 , but worse BIC.

The impact of the inclusion of the complex phases is visualised in Figure 4. If two phases overlap, this suggests ‘no interference’ between the two factors, meaning that the factors are fully compensatory and can be traded off against each other. Thus, in this case, as there is no interference between time and bus, this implies that there is no correlation between travel times and unobserved utility contributions towards bus. As a contrast, walk and taxi are most different.

Key differences between models

Notably, the performance of the models looks very different if they are applied to subsets of the dataset where the dataset is split based on the number of available alternatives. The model fit for choice tasks with a different number of available alternatives is given in Table 4, in which the performance of the base MNL and 4th QAM model are compared to their counterparts fitted to the subsets of data.

Most strikingly, it appears that QAM performs better than MNL for 2 or 3 alternatives, and worse for 4 or more alternatives. The QAM model also appears insufficiently adapted for datasets with differing numbers of available alternatives, with the performance of the QAM model improving more significantly than that of the MNL model when separate models are applied to subsets of data.¹

¹The application of correction factors (Van Cranenburgh et al., 2015), where an estimated coefficient

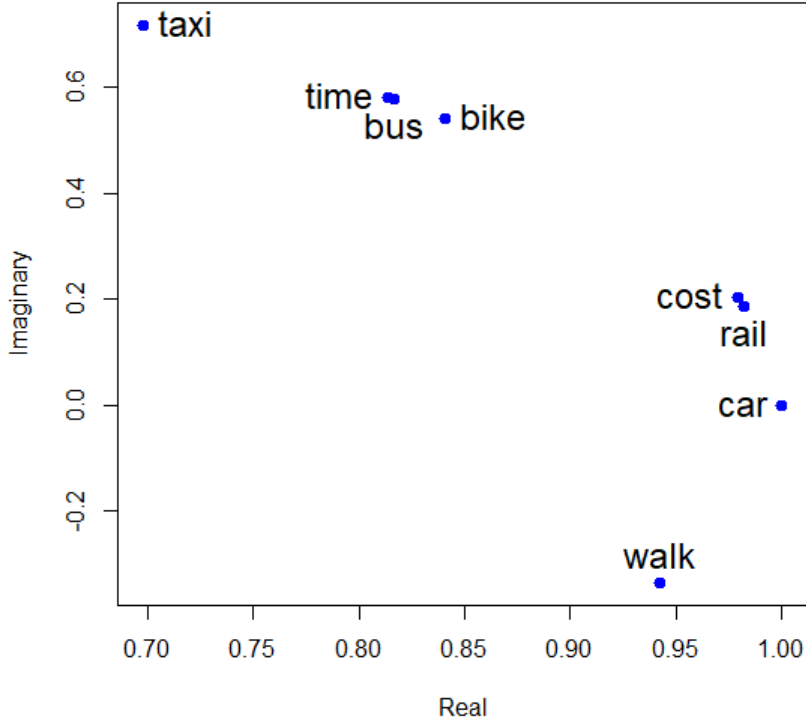


Figure 4: Inference patterns for the 4th QAM model, with the complex phase outputs for each attribute/alternative visualised in real and imaginary space.

Table 4: Contrasting performance of models applied to subsets of data based on the number of available alternatives.

		Same model			Separate models		
Available alts	Obs.	MNL	Quantum	Diff	MNL	Quantum	Diff
2 alts	1,541	-319	-317	2	-301	-281	21
3 alts	2,940	-1,473	-1,400	73	-1,434	-1,337	97
4 alts	3,630	-1,730	-1,756	-25	-1,712	-1,731	-19
5/6 alts	1,143	-534	-559	-25	-501	-511	-11
Sum	9,254	-4,057	-4,032	25	-3,948	-3,859	88

Results for models capturing behavioural change

Next, we add parameters to the MNL and QAM models to capture the shift in preference following the provision of feedback. Using Equations 10 and 11, respectively, the gain in model performance is shown in Table 5.

Positive shifts are observed for $\delta_{cal,i}$, meaning that a participant who is told they burn more calories than individuals of a similar demographic shift away from the use of a car towards the use of mode i (note that the effects for cycling and bus were insignificant and thus dropped). Negative shifts are observed for individuals who are told they burn more CO₂. This implies that current behaviour is reinforced. Though results between the models appear similar, there are small differences, with the QAM in comparison to MNL suggesting a stronger relative shift for walking in comparison to rail (see Figure 5). The converse is true for the shift from taxi.

replaces $J_{nt} - 1$ in Equation 4, results in some improvement (-4,021) though not yet getting close to the gain obtained through the use of separate models (-3,859).

Table 5: Model results after the inclusion of parameters to capture behavioural change.

Model	MNL		QAM	
Base model LL	-4,056.53		-4,031.54	
LL	-4,029.68		-4,009.35	
Difference	26.85		22.20	
Extra pars.	6		6	
Likelihood Ratio Test	8.48E-10		6.214E-08	
par.	est	rob. t-rat	est	rob. t-rat
$\delta_{co2,rail}$	-1.838	-2.01	-0.207	-2.96
$\delta_{co2,taxi}$	-1.669	-2.97	-0.136	-4.49
$\delta_{co2,walk}$	-2.514	-3.58	-0.339	-7.07
$\delta_{cal,rail}$	1.522	2.21	0.221	3.16
$\delta_{cal,taxi}$	1.379	3.39	0.124	4.27
$\delta_{cal,walk}$	1.897	3.19	0.243	6.70

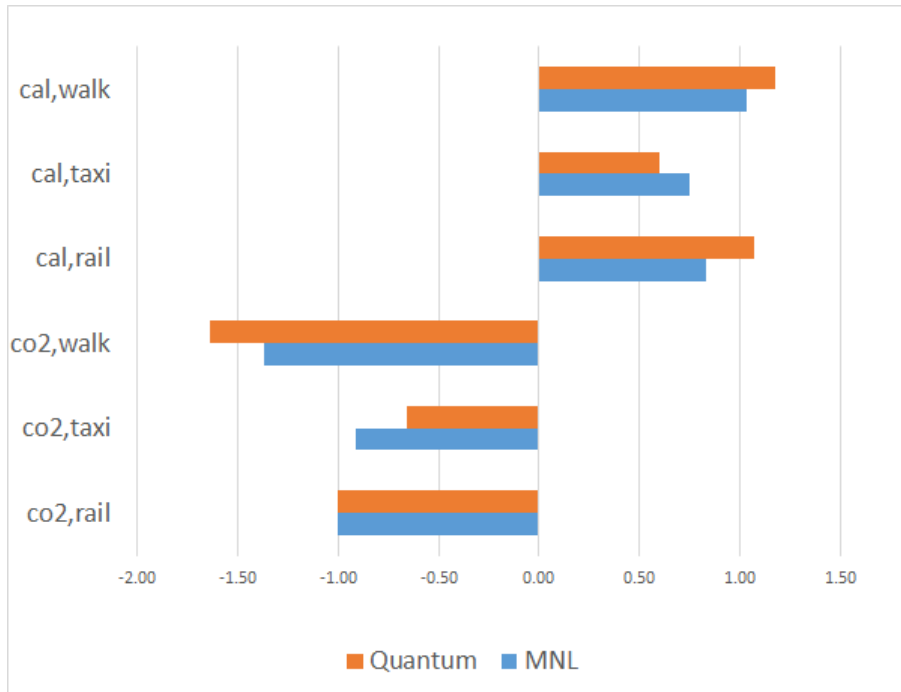


Figure 5: Relative shifts in preferences through fixing the shift from rail to car for those who used more CO2 to a value of -1 in both models.

4 CONCLUSIONS

In this work, we applied quantum choice models to revealed preference data for the first time, utilising ‘quantum rotations’ to capture the impact of behavioural nudges. Contrary to most examples in the literature, we find that, regardless of the modelling paradigm used, feedback reinforces behaviour, as opposed to changing behaviour. In particular, there is a shift towards choosing car if an individual is told they use more CO2 than others, whereas there is a shift away from choosing car if an individual is told that they burn more calories than others. However, these results may be subject to endogeneity biases (Palma et al., 2019), and future work should consider whether these concerns are properly accounted for in the choice models applied here.

Contrary to results from SP data (Hancock et al., 2020b,a), the base QAM model performed worse than the counterpart MNL. However, after utilisation of the full flexibility of QAM models, performance was significantly better than that of MNL, though a full interpretation of what complex phases actually capture remains a direction for future research.

Though it was not evident that QAM models are better suited than MNL models at capturing behavioural change (log-likelihood improvement was similar, see Table 5), it was clear that QAM models appear better when there are fewer available alternatives. Further work is required to establish whether these findings can be generalised.

Finally, it should be noted that extensions are possible for both models in terms of capturing heterogeneity in responses to the provision of feedback: some individuals may be more impressionable

or may simply have more opportunities to change their behaviour. Further analysis is required to better understand these possible individual differences.

ACKNOWLEDGEMENTS

Thomas Hancock and Charisma Choudhury's time were partially supported by UKRI Future Leader Fellowship [MRT020423/1]. Stephane Hess acknowledges the financial support by the European Research Council through the consolidator Grant 615596-DECISIONS.

REFERENCES

- Calastri, C., Crastes dit Sourd, R., & Hess, S. (2020). We want it all: experiences from a survey seeking to capture social network structures, lifetime events and short-term travel and activity planning. *Transportation*, *47*, 175–201.
- Chorus, C. G. (2010). A new model of random regret minimization. *European Journal of Transport and Infrastructure Research*, *10*(2).
- Hancock, T. O., Broekaert, J., Hess, S., & Choudhury, C. F. (2020a). Quantum choice models: a flexible new approach for understanding moral decision-making. *Journal of choice modelling*, *37*, 100235.
- Hancock, T. O., Broekaert, J., Hess, S., & Choudhury, C. F. (2020b). Quantum probability: A new method for modelling travel behaviour. *Transportation Research Part B: Methodological*, *139*, 165–198.
- Palma, D., Crastes dit Sourd, R., Calastri, C., Hess, S., & O'Neill, V. (2019). Can information really change travel behaviour? controlling for endogeneity in modelling the effect of feedback information. *Institute for Transport Studies, University of Leeds. Working paper*.
- Pothos, E. M., & Busemeyer, J. R. (2022). Quantum cognition. *Annual review of psychology*, *73*, 749–778.
- Trueblood, J. S., & Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. *Cognitive science*, *35*(8), 1518–1552.
- Tsoleridis, P., Choudhury, C. F., & Hess, S. (2022). Deriving transport appraisal values from emerging revealed preference data. *Transportation Research Part A: Policy and Practice*, *165*, 225–245.
- Van Cranenburgh, S., Prato, C. G., & Chorus, C. (2015). *Accounting for variation in choice set size in random regret minimization models*. working paper.

Identifying instant utility (latent emotion) triggers using psychophysiological indicators with an Experience-Based Choice Model in a travel experiment

Bastián Henríquez-Jara^{*1}, C. Angelo Guevara², and Angel Jiménez³

¹PhD (c), Department of Industrial Engineering, Universidad de Chile, Chile

²Prof., Department of Civil Engineering, Universidad de Chile, Chile

³Prof., Department of Industrial Engineering, Universidad de Chile, Chile

SHORT SUMMARY

We propose the Experience-Based Choice Model (EBCM), a novel approach capable of: (1) revealing the triggers of instant utility (emotions) in a transportation context, (2) measuring instant utility using psychophysiological indicators, and (3) estimating choices based on experiences. This framework combines the canonical discrete choice modelling, with the cyclical idea of decisions influenced by hedonic measures of experiences. In this article, we apply the components (1) and (2) of EBCM with data from a real-life travel experiment, in which skin temperature (SKT), heart rate (HR), heart rate variation (HRV), and electrodermal activity (EDA) were measured with a specially designed wristband. Using a latent variable approach, the main results show that instant utilities are sensible, to the travel mode; speed; crowding; brightness; and noise. In addition, it is shown that the participants kept a biased memory of the emotions and that EDA and SKT are meaningful indicators of instant utilities.

Keywords: EBCM; instant utility; psychophysiological indicators; travel experiment; emotions.

1 INTRODUCTION

The traditional models used to predict travel behaviour and discrete choice, in general, are based on the classic assumption that people are able to perfectly predict the utility they will perceive once they choose any alternative (Becker, 1996). This rationality argument, which is questionable in various fields including transportation, may be relaxed by introducing the ideas of experienced and instant utility proposed by Kahneman et al. (1997).

Kahneman proposed that when an individual chooses a specific alternative from a set, that alternative is afterwards experienced, causing a set of outcomes in every instant of that experience. The outcomes trigger hedonic feelings at each time point, which are called *instant utilities*. At the end of the experience, the individual has experimented a total utility (or *experienced utility*) which is the area beneath the curve of instant utilities. However, the individual may associate with that experience a biased level of utility due to memory limitations or any other bias source. This is called the *remembered utility* and is a function of the experienced instant utilities.

In the practical implementation of his framework Kahneman et al. (1997) measured instant utility directly from questionnaires of stated emotions, while fantasising with future methods that may instead incorporate *physiological indices of stress and of hedonic states*. Such is the goal of the present paper. After 25 years, we are capable of measuring instant utilities using PPIs in a real-life experiment. For this we adapted Kahneman et al. (1997)'s framework and propose the Experience-Based Choice Model (EBCM), which is capable of (1) revealing the triggers of instant utility, (2) measuring instant utility using psychophysiological indicators (PPIs) and (3) estimating choices based on past experiences. **In this framework, we understand *instant utility* as a latent variable that gives account of the psychophysiological states and emotions triggered by an experience in a specific time point.**

For this purpose, we use a transportation context, plagued with situations where this framework occurs naturally, but the methodology may be straightforwardly applied in several other fields.

Castro et al. (2020) and Hancock & Choudhury (2023) proposed the use of PPIs to correct endogeneity of hybrid models (Walker & Ben-Akiva, 2002) and to measure latent variables. But, no previous studies have used PPIs as indicators of latent variables. However, PPIs have been used in travel context in correlational studies, to predict choices based on stress levels in driving simulators (see Hancock & Choudhury (2023) for a review) and to model stated emotions with logistic

regressions (Barria et al., 2022). To the best of our knowledge, none of them bring to discussion framework of Kahneman et al. (1997).

In this article, we focus on estimating the components (1) and (2) of EBCM. The results show the triggers of instant utility; how PPIs variate with instant utility and the benefit from incorporating PPIs as indicators of instant utility. We also explored how individuals keep biased memories of their experiences.

The data used in this study is retrieved from the experiment reported in Barria et al. (2022). They collected physiological indicators, environmental and travel variables from 44 participants travelling in different modes. The field experiment design is based on the methodological framework proposed and validated by Castro et al. (2020).

The remainder of this article is divided into 4 sections. Section 2 overviews the EBCM. Section 3 shows the methods, Section 4 the main results and Section 5 exposes the main conclusions of this work.

2 THE EXPERIENCE-BASED CHOICE MODEL

We propose an extended Experience-Based Choice Model (EBCM) that integrates Kahneman’s theory with the canonical discrete choice framework. Figure 1 shows a schematic view of the model. We argue, that decision-making is a cyclical process, where the individual learns from his/her experience and also considers exogenous information for making a choice. In this framework, an experience is composed by a mapping of outcomes to instant utilities, which are latent hedonic measures of the outcomes of the decision and can also be understood as latent emotions. Instant utilities are aggregated as a memory into a *remembered utility*. We call this process the *Memory Aggregation Process* (MAP), which could adopt different functional forms (e.g. the mean or the logsum of instant utilities). Instant utility is a latent variable that can be measured with statements of emotions and PPIs, while the remembered utility can be measured with post-experience stated emotions. Remembered utility is a biased measure of the experience, in contrast to experience utility which is the area beneath the curve of instant utilities. When facing the same decision in the future, the decision utility will be a function of exogenous information and the remembered utility from previous experiences with each alternative. In the absence of previous experience, it could be assumed that decision depends just on exogenous information.

Both frameworks complement each other, since Kahneman’s framework did not consider the exogenous information and did not consider instant and remembered utilities as latent variables that can be measured. On the other hand, the canonical decision making framework neglects the weigh of experience on decisions. EBCM is capable of (1) revealing the triggers of instant utility, (2) measuring instant utility using psychophysiological indicators (PPIs) and (3) estimating choices based on past experiences.

When an experience is remembered worse than what it actually was, i.e., negative memory bias exists, the subject is less likely to choose it in the future, despite the actual level of experienced utility (the opposite follows directly). This is the main reason why observed choices should not be used to infer satisfaction with the alternatives but only a biased notion of satisfaction.

In this article, we focus on modelling the relations outside the canonical framework in Figure 1 (relations 4, 5, 6, and 8) in order to: (1) prove the feasibility of measuring instant utility with PPI plus real-time and PPI plus post-trip stated emotions; (2) identify the triggers of emotions; and (4) identify which PPIs are best suited to measure latent emotions. Future efforts should be made to model the complete EBCM.

3 METHODS

In this study we used the data collected in the experiment reported by Barria et al. (2022) to partially model the EBCM. To measure the PPIs, a wristlab called Biomonitor 2.0 was used, which was developed by WeSST Lab at Universidad de Chile (Jimenez-Molina et al., 2018). It can measure electrodermal activity (EDA), heart rate (HR), heart rate variation (HRV), skin temperature (SKT), and acceleration. In addition, an observer collected environmental data using device of sensors called ContextINO, and recorded punctual events that occurred during each trip using a mobile application called PsychoTrans. Specifically, the relations outside the canonical framework were modelled in Figure 1. We estimated the following models:

1. MV: instant utility measured by valence of emotions and PPIs (Figure 2)

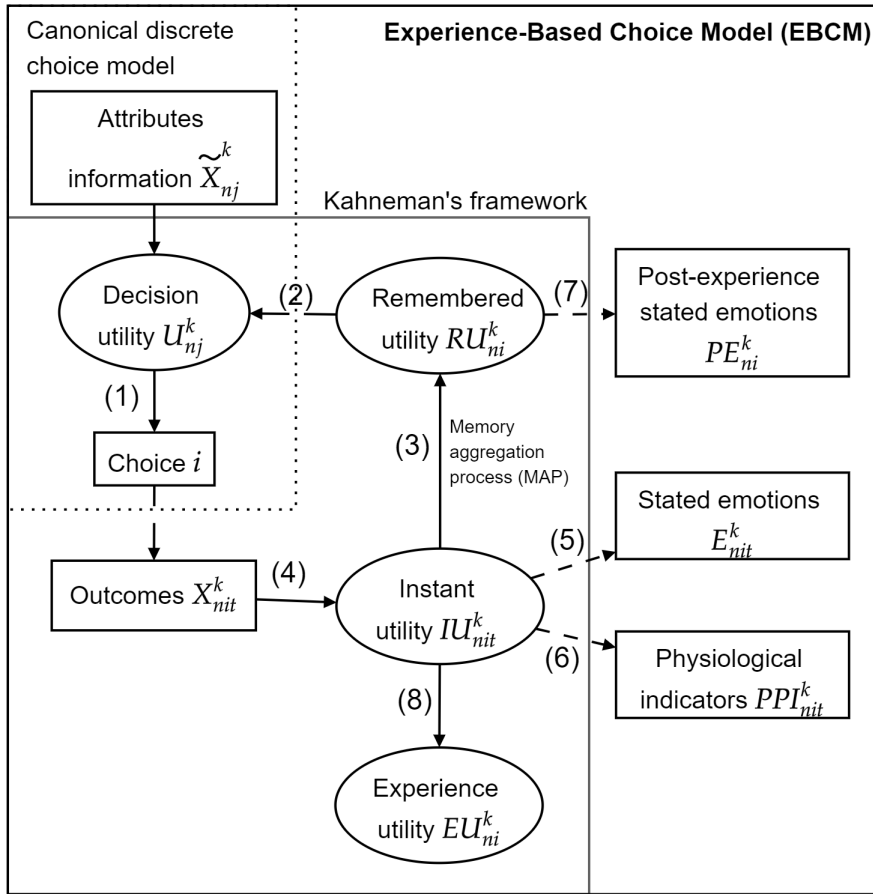


Figure 1: Experience-Based Choice Model (EBCM)

2. MA: instant utility measured by arousal of emotions and PPIs (Figure 2)
3. MV-NoP: MV without PPIs
4. MA-NoP: MA without PPIs
5. MV-NoT: MV without travel data
6. MA-NoT: MA without travel data

The estimation was carried out using *Apollo*, a freeware package for R (Hess & Palma, 2019). After model estimation, we estimated the posterior model parameter distributions using the function *apollo_conditionals* of the *Apollo* package. In addition, onboard stated emotions were compared with ex-post stated emotions.

Then, spatial and temporal profiles of instant utility are generated using the estimated parameters. These profiles allow visualising how did participant's instant utility vary in the different sections of the trip.

4 RESULTS

The results shown are obtained using 5000 Halton draws ($N(0,1)$) for the likelihood simulation. It was verified that estimates are stable between 4000 and 5000 draws. We display in Figure 3 the impact on the valence and on the arousal of the different triggers studied. Since only the ratio of the coefficients can be identified in discrete choice models, the analysis is presented relative to (β_{tt}) , which showed the largest nominal value. Just the statistically significant, and those that were slightly above the significance limit in at least one of the models, are shown.

It was observed that:

- Travelling on BHS (when velocity is above the average) causes more happiness than any of the other variables.

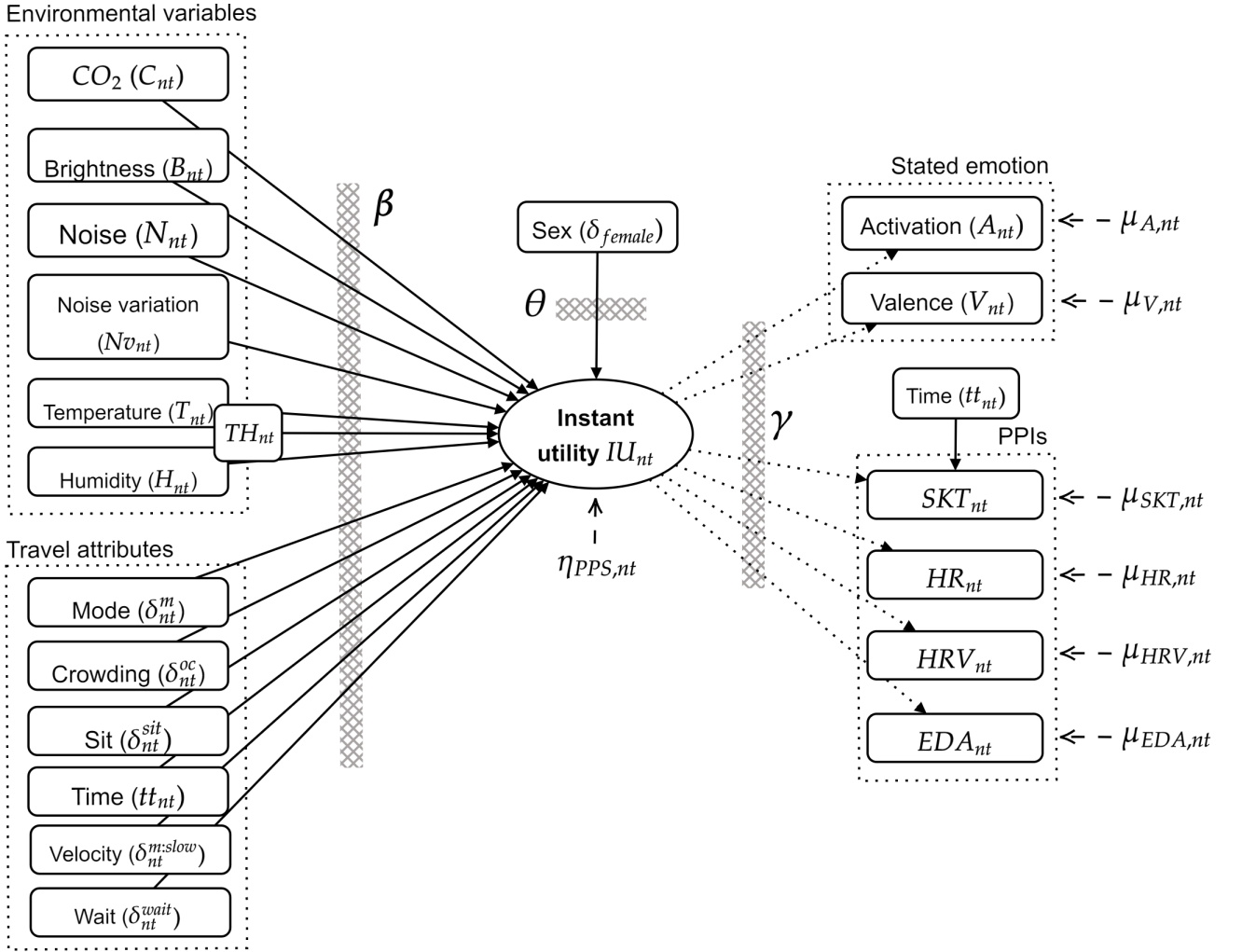


Figure 2: Model diagram of MV and MA

- Travelling on BLS (when velocity is below average) triggers more sadness than any other considered variables.
- When velocity is above the average, the impact of BLS on emotions is similar to walking. This may appear contradictory, but we suggest this is due to an unsafe feeling when riding BLS and higher displeasure caused by vibrations and route accelerations.
- Crowings levels 2 and 1 reduce the arousal.
- Better-illuminated places positively affect emotions.
- Higher humidity and higher noise variation increase the feeling of sadness.

We conducted a similar latent variable model but neglecting PPIs (MV-NoP and MA-NoP). It was found no environmental variables turned statistically significant. Some travel variable are significant.

From the estimation of the parameters of the measurement equations, some relations are derived: At higher valence:

- Higher EDA
- Lower SKT, HR

At higher arousal:

- Higher EDA, HR
- Lower HRV

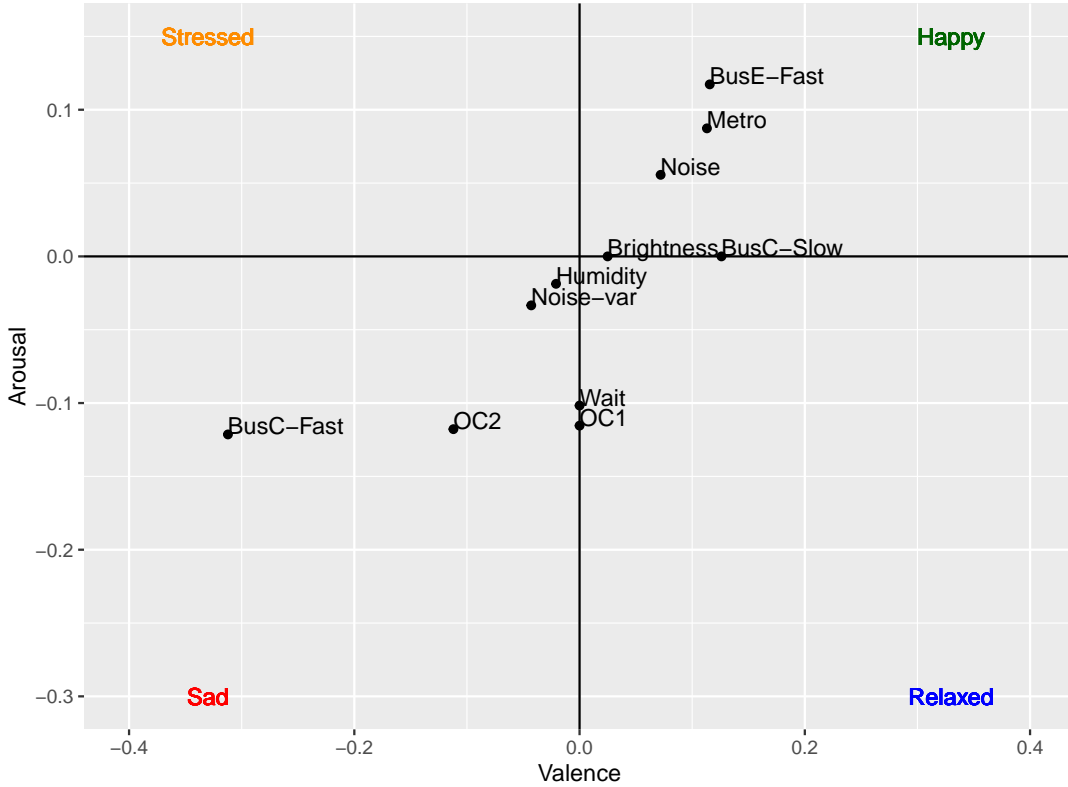


Figure 3: Relative marginal contribution to Valence and Activation

Posterior parameter distribution

This section seeks to answer whether the relations found between instantaneous utility and PPI can be extrapolated to the individual level. Figure 4 and Figure 5 show, for MV and MA models, respectively, the density function of both the population distribution (no filled area) and the density of the expected value of each parameter across individuals (filled area). The mean population distribution of the parameter is shown with a vertical dashed line, while the mean expected value across individuals is shown with a vertical solid line. Under each density plot, a box-plot shows the distribution of the expected parameters for each individual.

Tables 4 and 4 show the analysis of the mean expected value of the valence-PPI and arousal-PPI effects. It can be observed that, on average, higher valence is expected to increase EDA, while it is expected to decrease SKT (-0.75 , $p < 0.001$). Higher arousal significantly increases the EDA (0.56 , $p = 0.01$) and HR (0.37 , $p = 0.02$) while it decreases HRV (-0.17 , $p = 0.01$).

Table 1: Individuals' expected valence-PPI relation: mean, standard deviation, t-test against zero and p-value

	Mean	sd	t-test	p-value
EDA	0.65	0.21	3.12	<0.001
HR	0.19	0.14	1.33	0.19
SKT	-0.75	0.17	-4.55	<0.001
HRV	0.06	0.05	1.41	0.17

Instant utility profile

Figures 6 and 7 show the spatial instant utility profiles. In both figures, higher values of instant utility indicate a higher probability of experiencing a positive valence or arousal, respectively. Also, it was estimated the mean instant utilities and the mean valence and arousal direct from the stated emotions. On the contrary, the comparison between the remembered mean arousal and the mean onboard stated arousal shows a negative bias (-4.447% , $p = 0.0923$). The same way, the mean remembered valence was 5.848% lower than the mean onboard stated valence ($p = 0.0267$). This means that participants, after the experiment, associate to the travel sadder feelings than what they actually experienced.

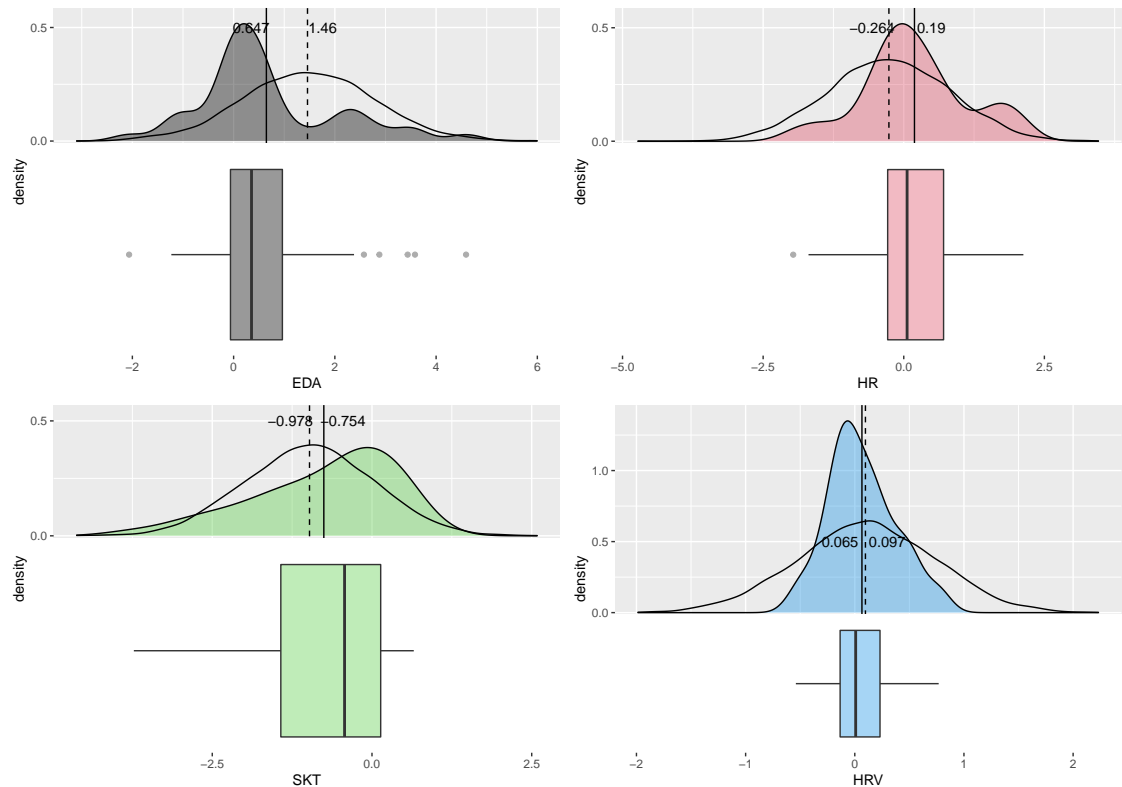


Figure 4: Distribution of the relation valence-PPI for the population (no filled density plot) and distribution of the expected parameters for each individual (filled density and boxplot)

Table 2: Individuals' expected Arousal-PPI relation: mean, standard deviation, t-test against zero and p-value

	Mean	sd	t-test	p-value
EDA	0.56	0.21	2.63	0.01
HR	0.37	0.15	2.49	0.02
SKT	-0.19	0.23	-0.81	0.42
HRV	-0.17	0.06	-2.88	0.01

5 CONCLUSIONS

A novel methodology was implemented to estimate the latent psychophysiological state of 44 participants in an on-road travel experiment. A novel approach to discrete choice is proposed: the Experience-Based Choice Model (EBCM). This model is partially estimated incorporating environmental, travel, and physiological variables.

It was shown that the participants' emotions were sensible environmental and travel variables. The low-standard bus has, better effects on emotions when velocities are slow, which can be associated with insecurity or discomfort feelings.

Also, it was shown statically significant mean effects of instant utility physiological indicators, despite that for different individuals this effect may be in opposite directions. It was also shown that the estimation of the instant utilities without the physiological indicators is futile due to the loss of statistical efficiency.

In addition, it was shown that the participants hold biased memories of what they actually experienced on the public transport trip. In fact, the bias on valence was negative and significant, which we suggest may be due to an over-weighting of past experiences or to a social tendency to negatively evaluate the use of public transport.

Policy-makers should consider the potential benefit of psychological well-being associated with less stressful travel conditions. Future research should consider including more heterogeneity in the sample and assess more interactions between variables. Also, the instant utility discussion opens

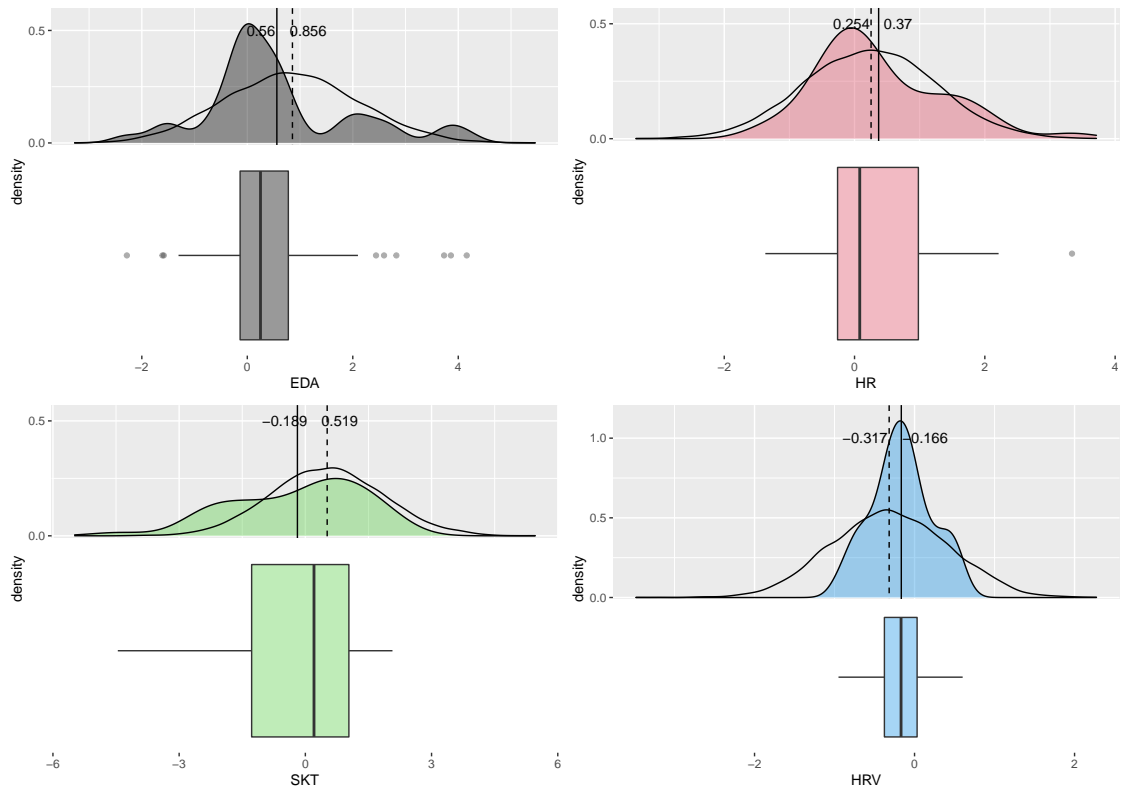


Figure 5: Distribution of the relation arousal-PPI for the population (no filled density plot) and distribution of the expected parameters for each individual (filled density and boxplot)



Figure 6: Spatial latent emotion profile estimated from model MV results

at least two research lines, namely: to explore whether memory bias influences modal choice and

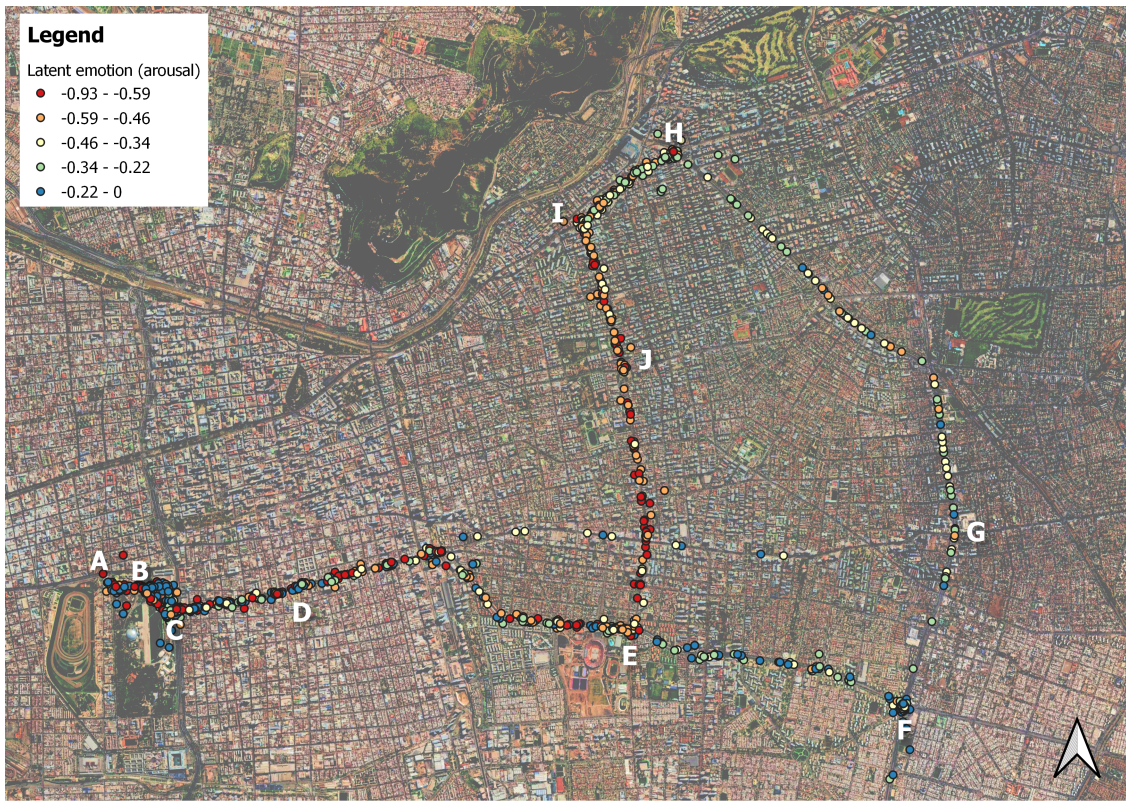


Figure 7: Spatial latent emotion profile estimated from model MA results

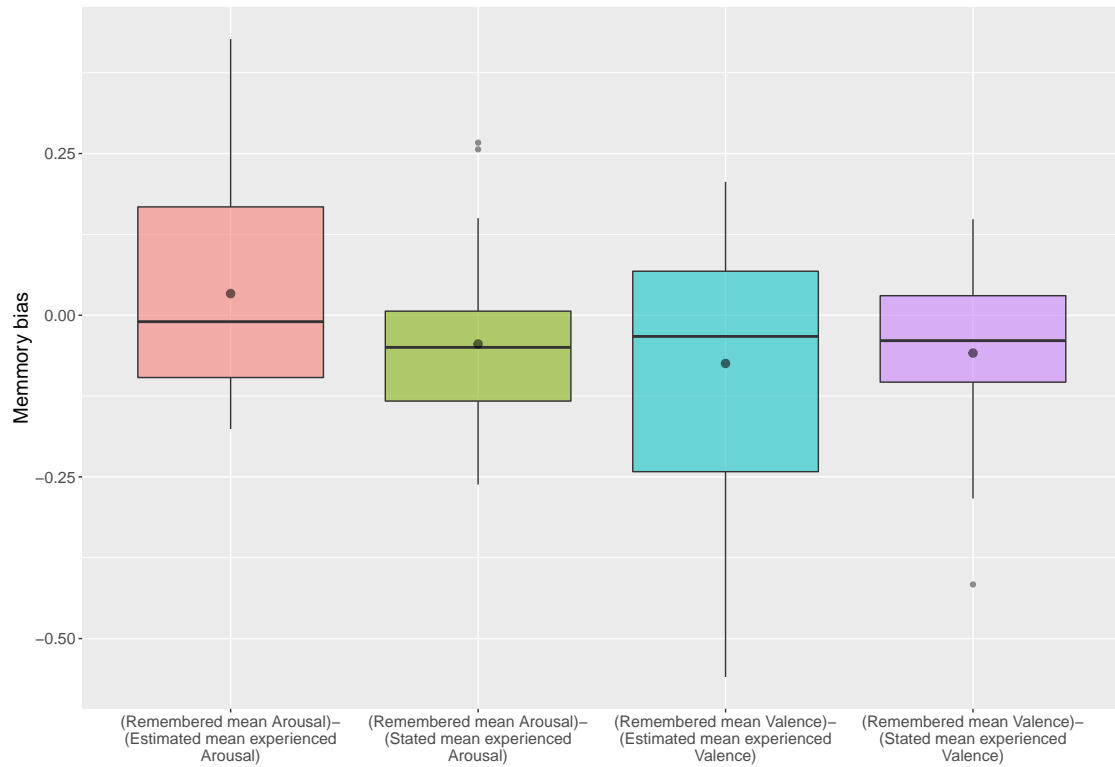


Figure 8: Estimated memory bias as the difference of the mean arousal and valence stated at the end of the trip (remembered valence and arousal) and the estimated or the stated valence during the trip

to explore whether this bias holds when travelling by other modes and if it is necessary to design public policies that tackle a possible loss of (biased) public transport users.

ACKNOWLEDGEMENTS

This work was partially funded by ANID through projects FONDECYT 1191104, 1231584, PIA/PUENTE AFB220003 and IDEA IT21I0059.

REFERENCES

- Barria, C., Guevara, A., Jimenez, A., & Seriani, S. (2022). *Relating Emotions, Psychophysiological Indicators and Context in Public Transport Trips: Case Study and a Joint Framework for Data Collection and Analysis*. [Unpublished manuscript]. doi: 10.2139/ssrn.4133094
- Becker, G. S. (1996). *Accounting for Tastes*. Cambridge, MA: Harvard University Press.
- Castro, M., Guevara, C. A., & Jimenez-Molina, A. (2020). A methodological framework to incorporate psychophysiological indicators into transportation modeling. *Transportation Research Part C: Emerging Technologies*, 118, 102712. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0968090X20306276> doi: <https://doi.org/10.1016/j.trc.2020.102712>
- Hancock, T. O., & Choudhury, C. F. (2023). Utilising physiological data for augmenting travel choice models: methodological frameworks and directions of future research. *Transport Reviews*. doi: 10.1080/01441647.2023.2175274
- Hess, S., & Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, 32(June), 100170. Retrieved from <https://doi.org/10.1016/j.jocm.2019.100170> doi: 10.1016/j.jocm.2019.100170
- Jimenez-Molina, A., Retamal, C., & Lira, H. (2018). Using psychophysiological sensors to assess mental workload during web browsing. *Sensors (Switzerland)*, 18(2), 1–26. doi: 10.3390/s18020458
- Kahneman, D., Wakker, P., & Sarin, R. (1997). Back to Bentham ? Explorations of Experienced Utility. *The Quarterly Journal of Economics*, 112(2), 375–405. Retrieved from <https://www.jstor.org/stable/2951240>
- Walker, J., & Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical social sciences*, 43(3), 303–343. doi: 10.1016/S0165-4896(02)00023-9

A model for Robust Rolling Stock Scheduling

Joris Wagenaar¹, Evelien van der Hurk², Marie Schmidt³, and Richard Lusby²

¹TiSEM: Department of Econometrics and Operations Research, Tilburg University, The Netherlands

²Management Science, DTU Management Engineering, Technical University of Denmark

³Faculty of Mathematics and Computer Science, Institute of Computer Science, University of Würzburg

Short summary

Major disruptions render the schedules of public transport operators infeasible. The majority of the recently developed algorithms for updating these schedules assume the duration of the disruption is known when they occur. However, in practice, this is generally untrue. This paper compares three different models for Rolling Stock Rescheduling under uncertainty: an optimistic approach, a strict-robust model inspired by the definition of disruption of Ben Tal et al, and a Light-Robustness approach that aims to provide a middle way between the two.

The models are evaluated on a realistic case study of the Netherlands Railways. Initial results indicate that building robustness against different disruption durations is worthwhile when alternative scenarios are associated with a sizable probability mass. The best approach depends on the probability distribution over the different scenarios.

Keywords: public transport, passenger train rolling stock rescheduling, robust optimization, mathematical combinatorial optimization

1 Introduction

Railway operators all over the world transport millions of passengers on a daily basis. During actual operations, railway operators may face major *disruptions* as a result of, e.g., a system malfunction, an accident, or the complete blockage of a track segment by a fallen tree. In such cases, the current operational plan becomes infeasible, and needs to be rescheduled. Research on effective disruption management has led to many algorithmic tools for rescheduling the timetable, rolling stock and crew schedule (see Cacchiani et al. (2014)). These types of advances support operators in increasing the reliability of their operations' reliability, hopefully positively impacting their passenger volumes.

Our work focuses on robust rolling stock rescheduling assuming the duration of the disruption is uncertain. Previous research on this topic, such as Nielsen (2011), Løve et al. (2002), and Wagenaar (2016), have always assumed the duration of the disruption is known. Although such algorithms could be used in a rolling horizon setting, this means that in practice they will often have either over or underestimated the disruption duration.

Considering only a single disruption duration when rescheduling the rolling stock could result in myopic, irreversible decisions that may negatively impact passenger comfort. For example, it might be impossible to provide capacity for all passengers on a trip if the disruption lasts longer or shorter than expected, trips may have unnecessarily been cancelled, or additional trips need to be cancelled due to unforeseen shortages of rolling stock. If disruption duration variability is considered when rescheduling the first time, then the quality of the service that the operator is able to provide can be improved. Moreover, a minimization in the number of required updates has a practical advantage. At many operators, the updating of the schedule requires manual updates and communications, and the risk of errors is minimized when the frequency of these changes is minimized.

The main contribution of this paper is a model that is able to reschedule the rolling stock schedule in a robust way. This means that the rolling stock schedule requires no or small additional changes in case the disruption duration turns out to be longer than originally expected. Depending on the level of robustness required, our model is able to give a full robust or a semi-robust solution. In case of a full robust solution all important trips and all composition changes are robust against different

disruption durations and in case of a semi-robust solution at least a given percentage of the trips must be robust against different disruption durations. We also demonstrate that robustness comes at a price and show what this price is for different practical settings.

2 Methodology

Problem definition

We consider the variant of the rolling stock scheduling problem as defined in Fioole et al. (2006), with the additional complication that, in contrast to (Fioole et al., 2006), the duration of the disruption is uncertain. Rolling stock scheduling thus entails finding a minimum cost assignment of train *compositions* (a specific ordering of specific rolling stock type units) to a set of timetabled *trips*. A trip refers to the movement of a train between two successive stops. It is assumed that both the departure time and the arrival time of any trip are known. Associated with each trip is a known, forecast demand that indicates the number of passengers who wish to make the trip. A *service* refers to the movement of a train between two terminal stations and comprises a sequence of trips. For any trip, its predecessor trip and its successor trip are specified in the timetable. Figure 1 illustrates an example timetable, adjusted after a major disruption. Stations are depicted as vertical layers, and time runs from left to right. A total of 44 trips is depicted, examples of which include t_1 , t_2 , and t_3 . The sequence $t_1 - t_2 - t_3$ provides an example of a service between Amsterdam (Asd) and Arnhem (Ah).

The general objective of the rolling stock rescheduling problem is to minimize a weighted combination of the number of cancelled trips, the number of additional composition changes and shunting movements in comparison to the planned rolling stock schedule, carriage kilometers, seat-shortages, and end-of-day imbalances in rolling stock depot inventories.

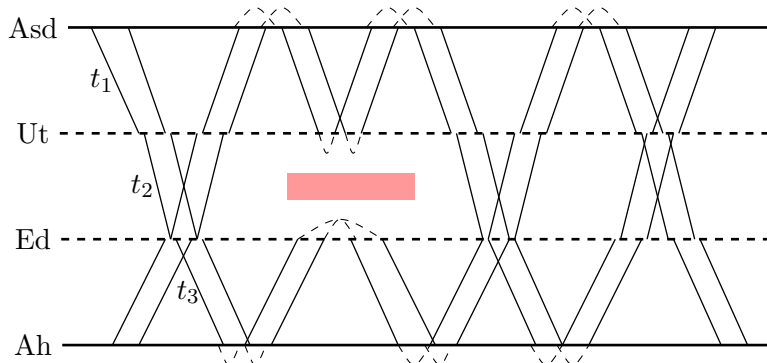


Figure 1: Timetable with disrupted area and short-turning.

Three approaches to robust rolling stock rescheduling

We propose three approaches for rolling stock rescheduling under uncertainty: a hopeful approach (HOPE) that always assumes the shortest-possible disruption duration and updates the rolling stock schedule step-by-step whenever this assumption is incorrect; a strict composition robust approach (STR), inspired by robustness as defined in Ben-Tal & Nemirovski (2002); and a light trip robust approach (LTR), inspired by light robustness as defined in Fischetti & Monaci (2009). All models are mixed-integer-programming models customized for rolling stock rescheduling under uncertainty. Unfortunately, the length of the abstract does not allow us to present them in detail here.

3 Results and discussion

Case study

2 depicts our Netherlands Railways(NS) based case study, spanning with 3 lines a significant and busy part of the network consisting of 1094 trips. We consider 16 different compositions, and thus

223 different composition changes, and assume that composition changes may occur after every trip.

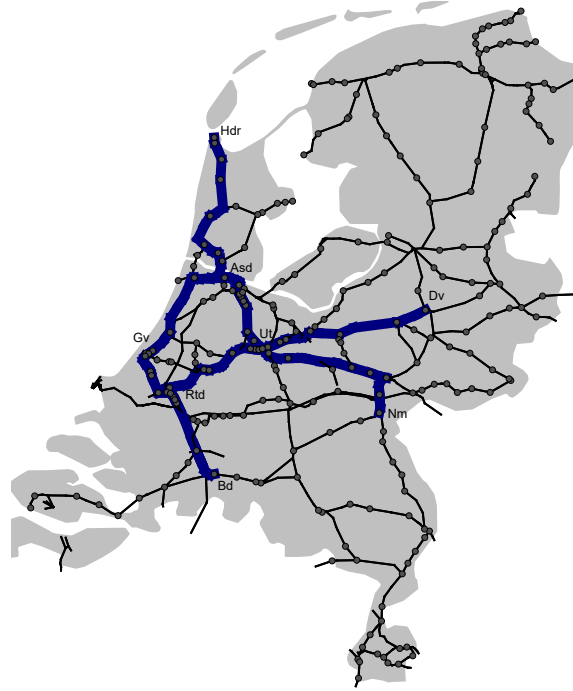


Figure 2: Train lines in the Netherlands. The lines in blue are the ones considered in the case study.

Evaluation function

The quality of a schedule is defined, for each possible disruption scenario, by the weighted sum of schedule changes. Table 1 gives an overview of the objective coefficients that we use in our experiments to evaluate the rescheduling approaches, set in discussion with NS. We make a distinction between rescheduling the first time (update λ_0) and rescheduling thereafter again (update λ_i for $i \geq 1$).

Costs	λ_0	λ_i for $i \geq 1$
Seat shortages: SS	0.1	0.1
Carriage costs p km: Carr.	0.01	0.01
Unplanned Shunting	10	100
Cancelled Shunting	5	20
Difference in end of day balance: EOD	10	10
Change in composition type: Composition	5	10
Cancellation of trip: Cancel	100000	100000

Table 1: Overview of costs

Detailed discussion of a single instance

There is a disruption between Utrecht (Ut) and Amsterdam (Asd) starting at 7:00 in the morning. We have the scenario set $\mathcal{S} = \{[(7 : 00, 9 : 00)], [(7 : 00, 9 : 00), (9 : 00, 11 : 00)]\}$. We show and compare results of four of our rescheduling approaches; the HOPE approach, the LTR approach with $\alpha = 0.4$ and $\alpha = 0.8$, and the STR approach.

The results of our comparison regarding the different cost components and the overall objective are given in Table 2. Here, the columns denote the total penalty paid for each different components of the evaluation. Only 'Shunting' denotes a combination of the penalties for both unplanned and cancelled shunting movements. The first three rows of the table show the evaluation value

and its components when rescheduling the original rolling stock schedule at time $\lambda_0 = 7 : 00$ for disruption end time $d_0 = 9 : 00$, that is, for update $u_0 = (7 : 00, 9 : 00)$. In the second set of rows we see the rescheduling costs that we incur if we need to reschedule for a second time, that is, for update $u_1 = (9 : 00, 11 : 00)$. The third set of rows depict the total costs for scenario $S_1 = [(7 : 00, 9 : 00), (9 : 00, 11 : 00)]$. The final set of rows shows the expected evaluation value and its cost components for the scenario set $\mathcal{S} = \{S_0, S_1\}$ with $S_0 = [u_0]$ and $\pi_0 = \pi_1 = 0.5$.

Evaluation	Approach (A)	Evaluation value	SS	Carr	Shunting	EOD	Cancel	Comps
'disr. ends at 9:00'	<i>HOPE</i>	3066	1285	1181	65	100	0	435
	<i>LTR</i> ^{0.4}	3199	1280	1184	140	100	0	495
	<i>LTR</i> ^{0.8}	3470	1300	1180	185	100	0	705
	<i>STR</i>	3658	1231	1183	150	200	0	894
'rescheduling costs'	<i>HOPE</i>	4468	1385	1175	600	100	0	1210
	<i>LTR</i> ^{0.4}	3908	1310	1178	260	100	0	1060
	<i>LTR</i> ^{0.8}	3054	1150	1174	140	100	0	490
	<i>STR</i>	2664	1281	1183	0	200	0	0
$GC(A, R_0, S_1)$ 'disr. ends at 11:00'	<i>HOPE</i>	4968	1385	1175	665	100	0	1645
	<i>LTR</i> ^{0.4}	4543	1310	1178	400	100	0	1555
	<i>LTR</i> ^{0.8}	3944	1150	1174	325	100	0	1195
	<i>STR</i>	3708	1281	1183	150	200	0	894
$GC(A, R_0, \mathcal{S})$ 'expected value'	<i>HOPE</i>	4017	1335	1177	365	100	0	1040
	<i>LTR</i> ^{0.4}	3875	1300	1180	270	100	0	1025
	<i>LTR</i> ^{0.8}	3707	1225	1177	255	100	0	950
	<i>STR</i>	3683	1256	1183	150	200	0	894

Table 2: Evaluation values and their components for the single instance considered in Section 3.

For short disruption lengths, the robust approach provides a cost that can be avoided using the HOPE approach. The main difference in schedules stems from the penalty for having more changed shunting operations and the number of different compositions appointed. However, when the disruption last long (second scenario) the rescheduling for the HOPE approach is much more expensive, in particular with respect to using different shunting operations and different compositions than in the initial schedule. Whether building in this robustness is worthwhile, depends on the probability on each of the two scenarios occurring.

The last set of columns denotes the expected evaluation value assuming a probability of 50% for each scenario, indicating the STR approach as optimal in this example. Figure 3 shows the expected evaluation values depending on the probabilities that we assign to the two scenarios 'disruption is over at 9:00' (π_0) and 'disruption is over at 11:00' ($\pi_0 = 1 - \pi_1$). As can be seen, in this example, each of the strategies is optimal for a certain range of probabilities: for a low probability of a later disruption end time, i.e., $\pi_1 \leq 0.22$, the HOPE approach is best, for $0.22 \leq \pi_1 \leq 0.3$ *LTR*^{0.4} is best, for $0.3 \leq \pi_1 \leq 0.45$ *LTR*^{0.8} is best, for $\pi_1 \geq 0.45$ the STR approach should be used.

In general, we observe that the penalty for the additional canceling of trips is chosen so high that the model avoids this measure completely. Composition changes, changes in the shunting movements, and in the end of day balance, which are penalized moderately in the model, have most influence on the evaluation. Seat shortages and costs for operating the carriages play a minor role for this parameter setting.

Experiments concerning multiple scenarios, and different ways of dealing with uncertain disruption length, are underway. Preliminary results indicate that whether the build-in of initial robustness against a longer disruption length is worthwhile when the probability of longer disruption lengths represents a significant probability mass.

4 Conclusions

We proposed three approaches to rolling stock rescheduling under uncertainty: HOPE, *LTR* and *STR*. The first represent current approach, the second a Light-robustness approach, and the latter a

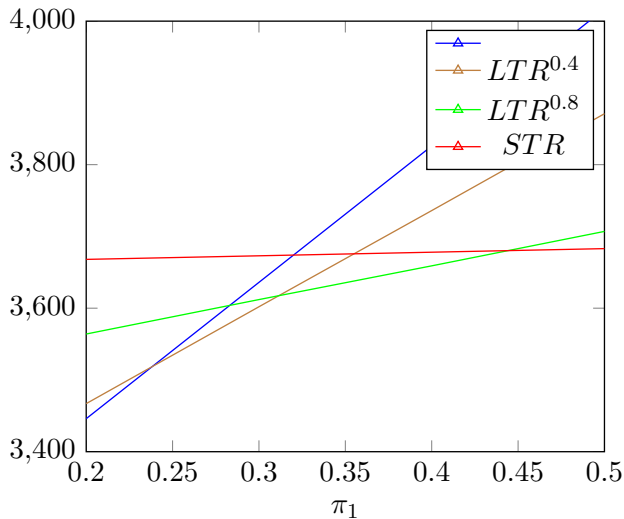


Figure 3: Expected evaluation values. The figure zooms in on the interesting part between $0.2 \leq \pi_1 \leq 0.5$

strict robustness version in the spirit of Ben-Tal & Nemirovski (2002). Mixed-Integer programming formulations have been developed for all three, that within a reasonable time can be solved with a commercial solver like CPLEX, Gurobi, etc. These approaches are evaluated on a realistic case study for Netherlands Railways. The results indicate that it is worthwhile to build in robustness for multiple possible durations of the disruption when there exists a reasonable probability for an alternative scenario. Which approach in expectation is best depends on the probabilities associated with each disruption scenario.

Current work is evaluating the model for multiple cases and multiple scenarios. Moreover, a sensitivity analysis for the selection of objective-parameters is on its way.

References

- Ben-Tal, A., & Nemirovski, A. (2002). Robust optimization—methodology and applications. *Mathematical Programming*, 92(3), 453–480.
- Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., & Wagenaar, J. (2014). An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63, 15–37.
- Fioole, P.-J., Kroon, L., Maróti, G., & Schrijver, A. (2006). A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research*, 174(2), 1281–1297.
- Fischetti, M., & Monaci, M. (2009). Light robustness. In *Robust and online large-scale optimization* (pp. 61–84). Springer.
- Løve, M., Sørensen, K. R., Larsen, J., & Clausen, J. (2002). Disruption management for an airline—rescheduling of aircraft. In *Workshops on applications of evolutionary computation* (pp. 315–324).
- Nielsen, L. K. (2011). *Rolling stock rescheduling in passenger railways: Applications in short-term planning and in disruption management* (No. EPS-2011-224-LIS).
- Wagenaar, J. (2016). *Practice oriented algorithmic disruption management in passenger railways* (No. EPS-2016-390-LIS).

Can Bayesian Optimization be the Last Puzzle for Automatic Estimation of Neural Network Discrete Choice Models? An experiment

Rui Yao¹ and Renming Liu¹

¹Technical University of Denmark, Denmark

SHORT SUMMARY

This study investigates the performances of Bayesian optimization (BO) and random grid search methods for tuning neural network hyper-parameters in the context of discrete choice modeling. Specifically, the fully-connected feed-forward (FNN) and alternative-specific-utility neural networks (ASU) are tuned. Results show that BO outperforms random grid search for both FNN and ASU models in terms of out-of-sample log-likelihood. Furthermore, it is illustrated that BO has higher sample efficiency and is relatively more robust to different random initialization. Our experiments show that the Bayesian hyper-parameter tuning framework could accommodate and complement existing neural network models that are cast for automatic utility function specifications, and create a fully automatic estimation workflow.

Keywords: Bayesian optimization, Discrete choice modeling, Hyper-parameter tuning, Neural network models

1 INTRODUCTION

Discrete choice modeling typically requires prior knowledge of the utility functions (Han et al., 2022), which are often, however, specified using trial-and-error based on researcher’s interpretations and experiences. Recent studies have adapted neural network (NN) methods for automatic utility function specifications and shown greater predictive power of NN models (e.g., Lee et al., 2018; Wang, Wang, & Zhao, 2020).

However, general-purpose neural networks tend to over-fitting. Consequently, their out-of-sample performances (Han et al., 2022), as well as their interpretability (Wang, Mo, & Zhao, 2020), could be poor. To tackle these issues, several studies have incorporated domain-knowledge-based regularization methods with NN models, by designing specific sparse NN architectures. For example, Sifringer et al. (2020) and Han et al. (2022) assign single-layer sparse (e.g., linear-in-parameter) neural network for the *interpretable* component of the systematic utility, and report improved predictive power and retained interpretability. Wang, Mo, & Zhao (2020) also show that their proposed alternative-specific-utility (ASU) NN modeling framework, in which only attributes associated with the same alternative are connected, generally improves model performances compared to fully-connected feed-forward NN. Yao & Bekhor (2022) applies the variational autoencoder neural network to estimate the implicit availability perception of alternatives and embed it in the utility computation.

Although these domain-specific NN for choice modeling have demonstrated the automatic feature-learning power of NN, their performances still heavily depend on the hyper-parameters. Wang, Mo, & Zhao (2020) shows that poorly tuned hyper-parameters can hinder the performances of NN models due to their large estimation error.

The challenges in selecting the NN hyper-parameters are two-fold: 1) Large number of hyper-parameters; 2) Bi-level problem structure. The lower-level model estimation depends on the upper-level hyper-parameter selection; whereas the upper-level hyper-parameter tuning typically requires evaluating the lower-level objective (i.e., out-of-sample log-likelihood in the context of choice modeling). In the case of large datasets, which is typical for applying data-driven methods, the lower-level model estimation could be time-consuming for complex NN models.

One classic approach for hyper-parameter tuning is to perform a random grid search on the hyper-parameter space and select the set of hyper-parameters with high out-of-sample log-likelihood (Bergstra & Bengio, 2012). Recently, Snoek et al. (2012) shows that, compared to random search, the Bayesian optimization (BO) method is more efficient and improves state-of-the-art performances of many machine-learning tasks.

The primary objective of BO is to identify the global optimum of an unknown function with only a small number of evaluations. This is achieved by modeling the unknown function as a Gaussian process (GP) and selecting the next sample point that maximizes an acquisition function derived from the GP. The acquisition function is designed to balance the trade-off between exploitation, where the mean is high, and exploration, where the uncertainty is high. BO is therefore employed for solving problems that are expensive to evaluate, have an unknown structure, such as concavity or linearity (i.e., the function is a black-box), and possess a continuous objective function (Frazier, 2018).

This study aims to evaluate and compare the performances of the Bayesian optimization method and random grid search for improving the model out-of-sample log-likelihood, which could provide a promising direction for a fully automatic neural network discrete choice model estimation pipeline.

2 METHODOLOGY

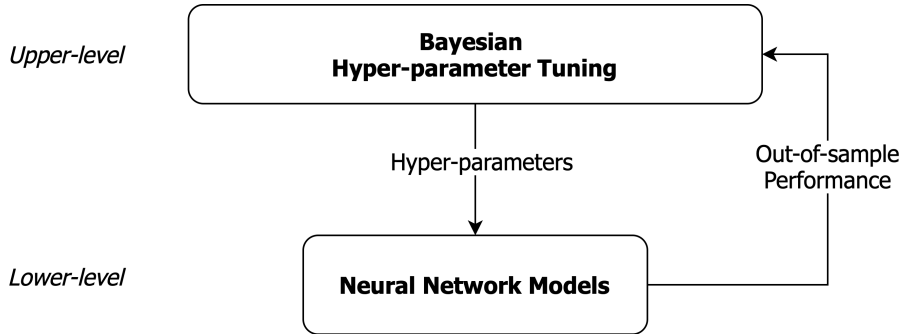


Figure 1. Bayesian hyper-parameter tuning framework

We show in Figure 1 the overall Bayesian hyper-parameter tuning framework. In the following subsections, we first detail the lower-level neural network models used in this study for discrete choice modeling. Next, the Bayesian hyper-parameter tuning method is introduced. Lastly, we briefly introduce the benchmark random grid search method.

Neural networks for choice modeling

Two neural network models, fully-connected feed-forward neural network (FNN) and alternative-specific utility neural network (ASU, Wang, Mo, & Zhao (2020)), are selected in this study for comparing the performances of Bayesian hyper-parameter tuning and random grid search in the context of choice modeling.

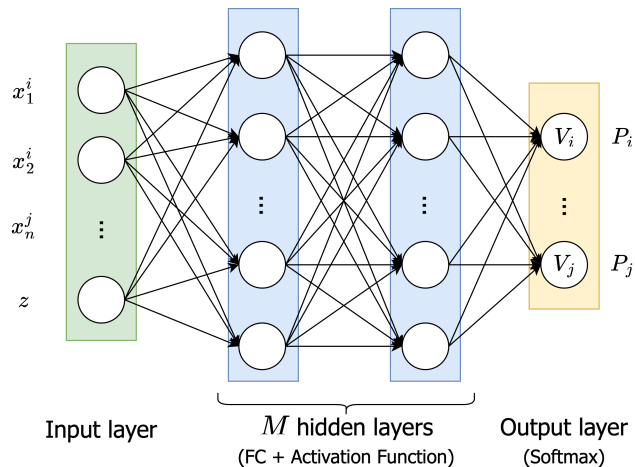


Figure 2. FNN architecture for discrete choice modeling

We show in Figure 2 the FNN architecture, in which there are two types of inputs, namely, alternative-specific attributes x^i for alternative i , and individual-specific attributes z . The FNN

model connects all input attributes to the hidden layer neurons through fully-connected layers (FC) and activation functions. Correspondingly, the systematic utility of an alternative i , V_i , is function of attributes of *all* alternatives $\mathbf{x} = \{\mathbf{x}^i, \forall i\}$ and the individual-specific attributes \mathbf{z} . Mathematically, systematic utility V_i of FNN can be defined recursively as:

$$V_i = V(\mathbf{z}, \mathbf{x}) = w_i^\top (g_M \circ \dots \circ g_1)(\mathbf{z}, \mathbf{x}) \quad (1)$$

where, w_i denote the weights on the output layer (i.e., readout) for alternative i before applying softmax (Logit function), $g_m(y) = \Phi(W_m^\top y)$ denote the hidden layer $m \in [1, M]$, $\Phi(\cdot)$ denote the activation function, W_m is the weight on the FC layer m , and \circ denote function composition. We refer to Wang, Mo, & Zhao (2020) for detailed proof on the connection between FNN and utility function specification.

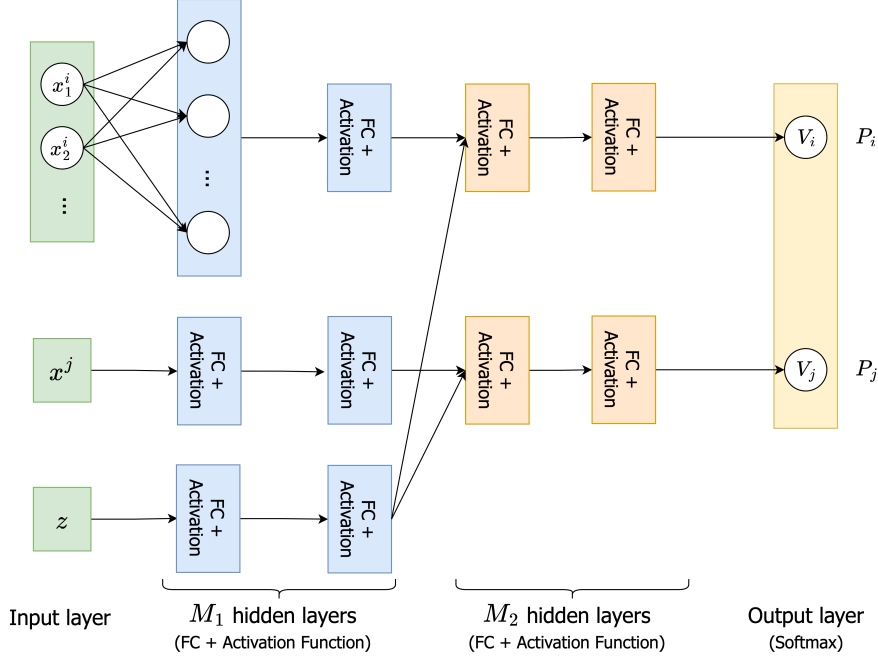


Figure 3. ASU architecture for discrete choice modeling (adapted from Wang, Mo, & Zhao (2020))

The ASU architecture is shown in Figure 3. Different from the FNN architecture, the ASU model first transforms independently each set of alternative-specific attributes \mathbf{x}^i and the individual-specific attributes \mathbf{z} with separate NNs. Only after M_1 layers of transformation, \mathbf{z} enter the systematic utility V_i . As a result, V_i is only function of alternative-specific attributes \mathbf{x}^i and the individual-specific attributes \mathbf{z} , which can be formally defined as:

$$V_i = V(\mathbf{z}, \mathbf{x}^i) = w_i^\top (g_{M_2}^i \circ \dots \circ g_1^i) ((g_{M_1}^z \circ \dots \circ g_1^z)(\mathbf{z}), (g_{M_1}^i \circ \dots \circ g_1^i) \mathbf{x}^i) \quad (2)$$

where, $g_{m_1}^i$ and $g_{m_1}^z$ denote the hidden layers for separate transformation of alternative-specific attributes \mathbf{x}^i and sociodemographic attributes \mathbf{z} , and $g_{m_2}^i$ denote the additional transformation after \mathbf{z} enters the computation of V_i .

We summarize a list of selected FNN and ASU hyper-parameters for comparing the performances of Bayesian optimization and random grid search in Table 1:

Table 1. Hyper-parameters for FNN and ASU

FNN	
M	Number of FNN hidden layers
n	Width (number of neurons) of each FNN hidden layer
ASU	
M_1	Number of ASU hidden layers for separate transformations
M_2	Number of ASU hidden layers after z enters
n_1	Width of ASU M_1 hidden layer
n_2	Width of ASU M_2 hidden layer
Generic	
l_1	L1 regularization parameter to control model sparsity
l_2	L2 regularization parameter to control coefficient magnitudes
Dropout rate	Probability of dropping some coefficients for model sparsity
Learning rate	Step size in stochastic gradient descent (SGD)
Batch size	Number of observations per batch in SGD
Batch normalization	Normalizing each batch of observations in SGD

Note that, we set the number of iterations as 20,000 and employ an early-stopping strategy for training both FNN and ASU, for which the algorithm stops if the log-likelihood of the validation set does not improve in 50 consecutive iterations. Other hyper-parameters are set as recommended values in the literature.

Bayesian hyper-parameter tuning

A BO framework comprises two primary steps (Frazier, 2018). The first step involves updating a Bayesian statistical model which approximates the complex mapping from the hyper-parameters (denoted by θ), to the objective values (i.e., out-of-sample log-likelihood l). The second step is to select a candidate hyper-parameter vector that optimizes the acquisition function and evaluates its performance.

The Gaussian process is often chosen as the prior for the statistical model, because of its tractability in computing posterior and predictive distributions. The GP is characterized by its mean function $\mu_0(\theta)$, and its covariance kernel function, denoted by $k(\theta, \theta')$. Given a set \mathcal{D}_m containing hyper-parameters and their corresponding objective values, that is, $\mathcal{D}_m = \{\theta_{1:m}, l_{1:m}\}$, where subscript $1 : m$ represents $\|m\|$ variables.

The joint distribution of $l_{1:m}$ is Gaussian:

$$l_{1:m} \sim \mathcal{N}(\mu_0(\theta_{1:m}), K(\theta_{1:m}, \theta'_{1:m})), \quad (3)$$

where μ_0 is the prior mean function, which is usually set as a constant value (0 in this study), and $K(\theta_{1:m}, \theta'_{1:m})_{i,j} = k(\theta_i, \theta'_j)$, for $i, j \in \{1, 2, \dots, m\}$, is the covariance matrix. We also assume k as the commonly used Matern kernel.

The posterior distribution of l_{m+1} can be computed using Bayes' theorem:

$$l_{m+1}|l_{1:m} \sim \mathcal{N}\left(\mu(\theta_{m+1}), \sigma^2(\theta_{m+1})\right), \quad (4)$$

where $\mu(\theta_{m+1}) = \mathbf{k}^T \mathbf{K}^{-1} l_{1:m}$ and $\sigma^2(\theta_{m+1}) = k(\theta_{m+1}, \theta_{m+1}) - \mathbf{k}^T \mathbf{K}^{-1} \mathbf{k}$.

The fitted GP, which serves as a surrogate model of the lower-level problem objective (i.e., out-of-sample log-likelihood in our paper), is capable of predicting the value of the objective function at un-evaluated hyper-parameter locations based on previously collected data points $\mathcal{D}_m = \{\theta_{1:m}, l_{1:m}\}$. The goal is to select the next vector of hyper-parameters with the highest gained value of information. Such value of information is measured by an acquisition function, which is a proxy function derived from the mean and variance of the objective function values and directs the next sample point. This study uses a popular acquisition function, Expected Improvement (EI), which computes the expected improvement with respect to the current maximum $l^* = \max l_{1:m}$, i.e.,

$$EI(\theta) = \mathbb{E}([l_{m+1} - l^*]^+ | \theta_{1:m}, l_{1:m}). \quad (5)$$

The next point to be evaluated is determined by:

$$\theta_{m+1} = \arg \max_{\theta} EI(\theta) \quad (6)$$

This optimization problem can be efficiently solved by local solvers like L-BFGS-B (Liu & Nocedal, 1989) with multiple restarts.

Random grid search

For a given grid of hyper-parameter values, a random grid search selects random combinations of these values to train the models. The set of hyper-parameters with the best out-of-sample performance is chosen in a post-hoc manner. Although the independent sampling procedure suggests that random grid search can be performed in parallel, its efficiency (in terms of computational costs) and effectiveness (in terms of *best* performance) could be low (Snoek et al., 2012).

3 RESULTS AND DISCUSSION

Dataset and experiment setup

Our experiment is based on the swissmetro dataset (Bierlaire et al., 2001), for which 6,768 observations with trip purpose of *commute* and *business* are selected for model estimation. In the dataset, respondents choose among 3 alternative modes: train, swissmetro, and car. Data statistics of the attributes used for estimation are summarized in Table 2:

Table 2. Data statistics for the selected swissmetro dataset

Attribute	Mean	Std.
Train time [min]	166.63	77.35
Train cost [CHF]	514.34	1088.93
Swissmetro time [min]	87.47	53.55
Swissmetro cost [CHF]	670.34	1441.59
Car time [min]	123.80	88.71
Car cost [CHF]	78.74	55.26
Number of observations: 6,768		
Number of choices: Train (13.26%), Swissmetro (57.94%), Car (28.71%)		

The dataset is divided into training, validation, and out-of-sample sets in the ratio 4: 1: 1. A five-fold cross-validation is used for model selection. That is, each model (i.e. one set of hyper-parameters) is trained 5 times with different data folds, and evaluated with the out-of-sample set. The model performance is taken as the average of 5 out-of-sample log-likelihoods.

We consider the following hyper-parameter space for FNN and ASU (Table 3):

Table 3. Hyper-parameter space for FNN and ASU

Hyper-parameter	Values
FNN	
M	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}
n	{60, 120, 240, 360, 480, 600}
ASU	
M_1	{1, 2, 3, 4, 5, 6}
M_2	{0, 1, 2, 3, 4, 5, 6}
n_1	{10, 20, 40, 60, 80}
n_2	{10, 20, 40, 60, 80, 100}
Generic	
l_1	$[1^{-20}, 1.0]$
l_2	$[1^{-20}, 1.0]$
Dropout rate	$[1^{-20}, 0.1]$
Learning rate	$[1^{-5}, 0.5]$
Batch size	{50, 100, 200, 500, 1000}
Batch normalization	{True, False}
Invariant hyper-parameters	
Activation function	ReLU (Rectified linear unit) and Softmax (Logit)
Loss function	Log-likelihood
Weight initialization	He initialization (He et al., 2015)

Note that, we consider discrete architecture parameters for FNN and ASU, while selected parameters for the training algorithm, namely, l_1, l_2 , dropout rate, and learning rate, are considered as continuous values. For the random grid search method, all discrete values are uniformly sampled, whereas log-uniform sampling is applied for the training parameters in order to draw values at different magnitudes.

Both FNN and ASU are implemented in PyTorch 1.12 with Adam optimizer (Kingma & Ba, 2014). The Bayesian optimization is implemented with the BoTorch package (Balandat et al., 2020). We set the number of hyper-parameter tuning iterations as 100 for ASU, and 80 for FNN (due to fewer hyper-parameters to be tuned). In addition, 2 replications with different random seeds are performed for both the BO and random grid search. That is, for each hyper-parameter tuning method, we estimate 200 ASU models and 160 FNN models.

Comparison of out-of-sample performances

We report out-of-sample performance and architecture parameters of the *best* models obtained by BO and random grid search methods in Table 4.

Table 4. Comparison of the *best* model of Bayesian optimization and random grid search

	Bayesian Optimization	Random grid search
FNN		
M	5	11
n	240	60
Out-of-sample performance	-741.17	-768.60
ASU		
M_1	1	3
M_2	1	1
n_1	40	10
n_2	80	20
Out-of-sample performance	-762.03	-787.94

As shown in Table 4, the BO outperforms random grid search in terms of out-of-sample log-likelihood for both the FNN and ASU models. Moreover, compared to random grid search, the BO method is able to find network architecture with fewer layers yet stronger predictive power on the out-of-sample dataset. This is consistent with the literature (e.g., Hillel, 2019; Han et al., 2022), for which shallower but wider NN empirically performs better in choice modeling tasks.

We further examine the hyper-parameter optimization iterations of BO and random grid search methods in Figure 4, where *iteration numbers* for the random grid search are sorted models.

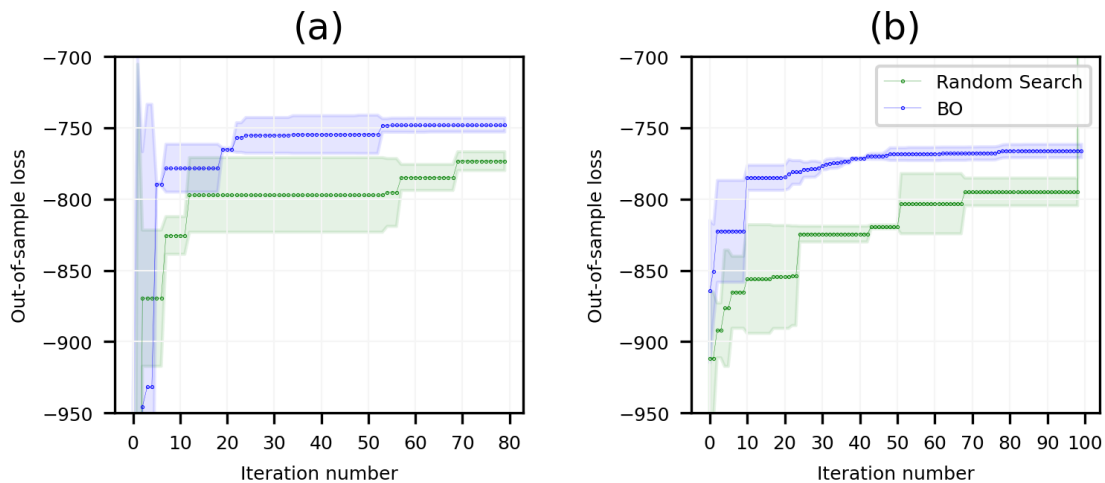


Figure 4. The optimization for hyper-parameters of (a) FNN and (b) ASU architectures (shaded area represents 95% confidence interval)

As shown in Figure 4, for both the FNN and ASU models, the BO method outperforms the random grid search after the initial 10 iterations. This suggests that, compared to random grid search, the BO method has higher sample efficiency, i.e., the BO candidate hyper-parameters have a higher potential to improve the out-of-sample performance. Furthermore, the variance of out-of-sample performance of BO iterations is smaller than random grid search (as indicated by the smaller shaded area of BO), which suggests the BO method could be more robust to (initial) randomization. In the following subsection, we further illustrate the sample efficiency of BO and random grid search.

Illustration of sample efficiency between Bayesian and random search methods

We next investigate the sample efficiency of random search and BO methods by illustrating the distributions of sample points for the ASU models, as shown in Figure 5 and 6. The diagonal plots show histograms for each hyper-parameter, and the lower triangle shows two-dimensional scatter plots of all sample points for each pair of the hyper-parameters. The red points represent the set

of hyper-parameters with the best out-of-sample performance (termed *optimal* hereinafter) found by each method.

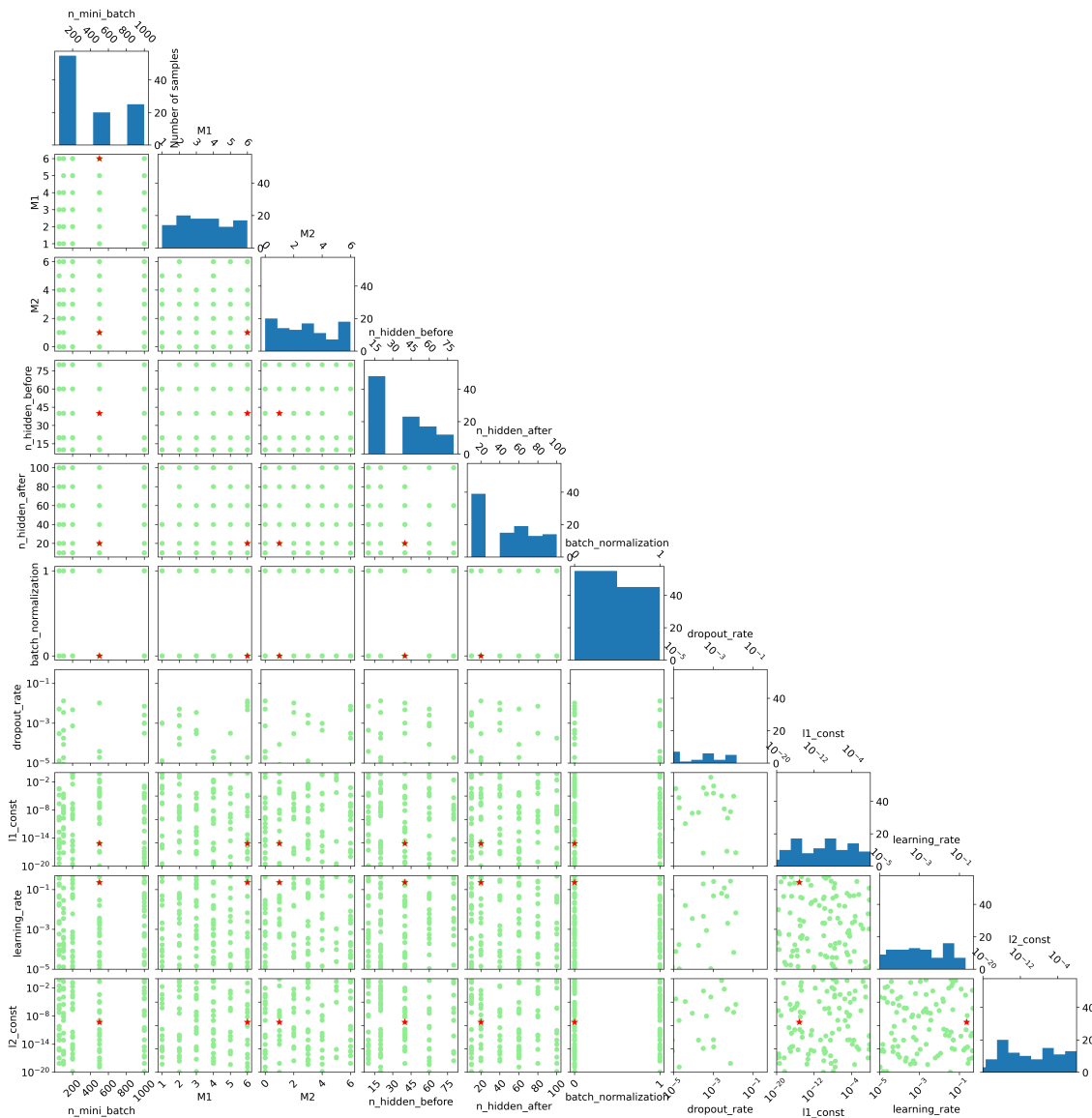


Figure 5. Sample points of the random search for ASU architecture. (Red points represent the *optimal* hyper-parameters)

As shown in Figure 5 for the random grid search method, all sample points are distributed almost uniformly across the search space. Although random grid search has been proven to be more efficient than brutal-force full grid search (Bergstra & Bengio, 2012), this uniform distribution of sample points still suggests the out-of-sample performance might not be improved in consecutive iterations, resulting in relatively poorer sample efficiency. This can also be verified by the longer and flatter platoons of random grid search in Figure 4 for out-of-sample performance versus iterations.

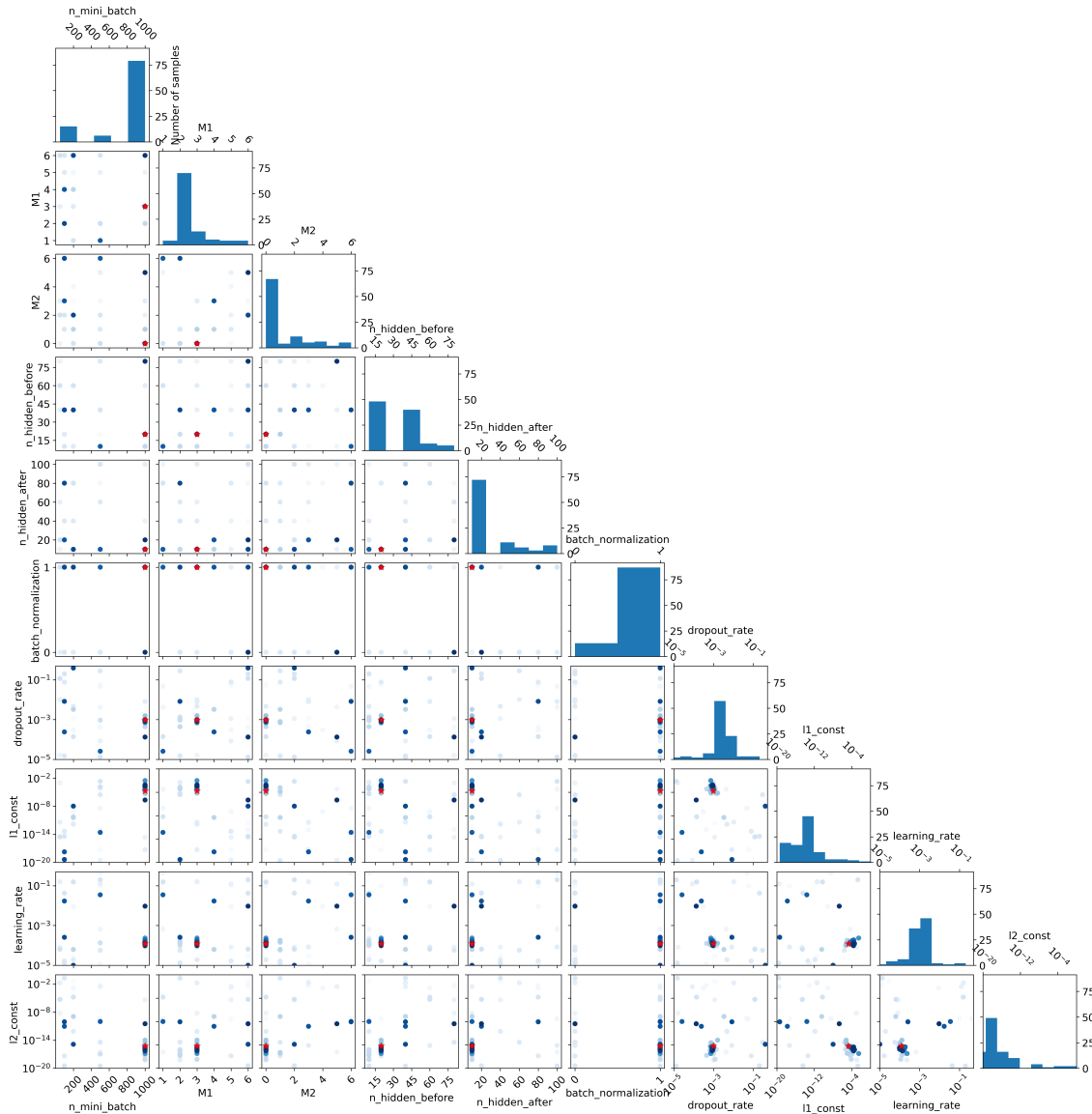


Figure 6. Acquisition points of the Bayesian optimization for ASU architecture.
(Red points represent the *optimal* hyper-parameters,
and darker colors correspond to later samples)

On the other hand, BO follows a different pattern. As shown in Figure 6, acquisition points obtained by BO gradually cluster around the *optimal* hyper-parameters, as indicated by the darker points centered around the red points. This clustering effect is also evident in the histograms, in which the highest frequencies are observed around the *optimal* hyper-parameters. These results indicate that the BO method has higher sample efficiency, compared to random grid search.

The behavior of the BO method can be explained by its optimization procedures. As more sample points are collected/evaluated, the surrogate model has higher confidence about its estimation of out-of-sample performance at unexplored hyper-parameter points. As a result, the associated acquisition function can direct more efficiently the search towards the *optimal* hyper-parameters with the improved surrogate model estimations.

Note that, similar patterns of sample points for random grid search and BO are also observed of FNN models, which suggests the BO method is expected to have relatively higher sample efficiency.

4 CONCLUSIONS

This study investigates the performances of Bayesian optimization (BO) and random grid search methods for tuning neural network hyper-parameters in the context of discrete choice modeling. Specifically, the fully-connected feed-forward (FNN) and alternative-specific-utility neural networks

(ASU) are tuned, and the out-of-sample performances as well as the sample efficiencies of the BO and random grid search methods are compared.

Results show that BO outperforms random grid search for both FNN and ASU models in terms of out-of-sample log-likelihood. Furthermore, it is illustrated that BO has higher sample efficiency and is relatively more robust to different random initialization.

Our experiments show that BO provides a promising direction for a fully automatic estimation workflow of neural network discrete choice models. Future research will extend the experiment to larger datasets, transferability of BO among datasets, as well as other data-driven machine learning models.

REFERENCES

- Balandat, M., Karrer, B., Jiang, D. R., Daulton, S., Letham, B., Wilson, A. G., & Bakshy, E. (2020). BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Advances in neural information processing systems 33*. Retrieved from <http://arxiv.org/abs/1910.06403>
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).
- Bierlaire, M., Axhausen, K., & Abay, G. (2001). The acceptance of modal innovation: The case of swissmetro. In *Swiss transport research conference*.
- Frazier, P. I. (2018). A tutorial on bayesian optimization. *arXiv preprint*.
- Han, Y., Pereira, F. C., Ben-Akiva, M., & Zegras, C. (2022). A neural-embedded discrete choice model: Learning taste representation with strengthened interpretability. *Transportation Research Part B: Methodological*, 163, 166–186.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the ieee international conference on computer vision* (pp. 1026–1034).
- Hillel, T. (2019). *Understanding travel mode choice: A new approach for city scale simulation* (Unpublished doctoral dissertation). University of Cambridge.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Lee, D., Derrible, S., & Pereira, F. C. (2018). Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling. *Transportation Research Record*, 2672(49), 101–112.
- Liu, D. C., & Nocedal, J. (1989). On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1-3), 503–528.
- Sifringer, B., Lurkin, V., & Alahi, A. (2020). Enhancing discrete choice models with representation learning. *Transportation Research Part B: Methodological*, 140, 236–261.
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25.
- Wang, S., Mo, B., & Zhao, J. (2020). Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions. *Transportation Research Part C: Emerging Technologies*, 112, 234–251.
- Wang, S., Wang, Q., & Zhao, J. (2020). Multitask learning deep neural networks to combine revealed and stated preference data. *Journal of choice modelling*, 37, 100236.
- Yao, R., & Bekhor, S. (2022). A variational autoencoder approach for choice set generation and implicit perception of alternatives in choice modeling. *Transportation Research Part B: Methodological*, 158, 273–294.

In-depth, Breath-first or Both? Toward the Development of a RUM-DFT Discrete Choice Model

Nova G.¹, Guevara C.A.*^{1,2}

¹ Department of Civil Engineering, University of Chile, Santiago, Chile

² Institute for Complex Engineering Systems (ISCI), Santiago, Chile

SHORT SUMMARY

Development in discrete choice modelling has been dominated by Random Utility Maximization approaches due to their ease of application and high economic interpretability. However, this model assumes that decision-makers perform an in-depth information search process (ISP) implicitly and instantaneously. It has not been investigated in detail whether the ISP of transport users is in depth or breadth-first in a public transport choice context, a gap that this research aims to fill. To this end, the ISP of public transport users has been characterized in three SP surveys with click-tracking, which were pivoted concerning commute and varied in the number of dimensions. The results allow us to conclude that the ISP is part of a heuristic, heterogeneous, complex, and mixed deliberation process, which depends on the dimensions of the choice tasks. However, breadth-first searches predominate, i.e., the evaluation of information is done by comparing alternatives under one attribute in each search.

Keywords: Information Search Process, Breadth-first, Public Transport SP, RUM -DFT

1. INTRODUCTION

Making a decision implies a deliberation process in which individuals, based on their preferences and valuations, choose an alternative among a given set (Engel et al., 1968). Several areas of knowledge have a great interest in comprehending the decision process, ranging from transportation and behavioural economy to psychology. To explain and predict how individuals make such decisions, choice models have been developed and widely used. Busemeyer and Townsend (1993) propose a classification of choice models into static and dynamic.

Static models define the choice probabilities in a way that is independent of the cognitive process and that does not vary in the choice process. Among the advantages, their simplicity of implementation and great explanatory power stand out, with the Random Utility Maximization (RUM) approach being the most widely used member of this group (McFadden, 1976). Static choice models consider that individuals are rational agents, possess complete information, and choose the alternative with the highest utility within the choice set by performing an in-depth information search of the attributes. In other words, RUM models assume an in-depth search first in which individuals consider all the attributes of an alternative to construct the utility of that alternative, which is then used to make a comparison and the choice.

RUM models are versatile, practical and provide transparent statistical a microeconomic framework for the analysis of discrete choices. However, they fail to consider the true cognitive process that individuals go through when deciding. RUM models assume that decision-makers somehow

instantaneously evaluate all the attributes of all the alternatives involved in the choice situation, but evidence suggest that the process is dynamic. Individuals focus on some attributes of some alternatives in different stages, acquiring information sequentially to make a final decision. Among others, Noguchi, and Stewart (2014); Stewart et al., (2016a); Stewart et al., (2016b); Sui et al. (2020) have evidenced, through eye-tracking those preferences vary across attribute evaluation and that this behaviour occurs in simple, risky, strategic and with multi-attributes choices.

On the other hand, dynamic models consider that preferences change over the time of choice due to the cognitive process. This approach has been represented, among others, by Decision Field Theory (DFT) (Busemeyer and Townsend, 1993). The DFT considers that the cognitive process is iterative and that the sequence of information search is breadth-first. This means that individuals begin with some initially preconceived preferences toward the available alternatives. Then, they focus on and evaluate one specific attribute at a time among all alternatives (breadth-first), and then iteratively update preferences by looking at other attributes until they finally choose an alternative when they reach their internal (preference) or external (time) limit. Recent studies show that the DFT model fits the data often better than conventional static models [Qin et al. (2013); Hancock et al. (2018)]. However, the DFT model presents important limitations, such as that it relies on ad-hoc matrix implementation, lacks a robust statistical theoretical framework that allow transparent identification, and lacks compatibility with microeconomic theory.

The trade of between RUM and DFT motivates the development of a RUM based model that could account for the dynamics of the decision-making process that is captured by the DFT model. As a first stage toward this overall research goal, this article is devoted to the collection an analysis of data on the sequential process of attributes evaluation within a public transport stated preference experiment. The experiment has the purpose of confirming or rejecting the breath-first behavioural hypothesis that is behind the DFT model and to study the impact of various contextual settings that may influence such behaviour. Besides, the data collected and studied in this research will be used to generate a database on which, later, different practical RUM-DFT models can be assessed.

2. LITERATURE REVIEW

The Information search process is defined as the stages where an individual performs cognitive tasks, such as searching their memory, acquiring new information and processing the data to carry out their choice (Payne, 1992; Riedl et al., 2008). Figure 1, Xie et al., (2019) presents this process that could be incorporated into discrete choice models.

On the one hand, there is the internal search related to retrieving information stored by individuals. Quite a few studies have been conducted to try to understand this process, but it is still unclear how to apply the findings of people's memory in choice models. And on the other hand, there is the external search, which corresponds to the stage of acquisition and processing of new information that individuals obtain from external sources (Holland et al.,1994). According to Schulte-Mecklenbeck et al.,(2017) this sub-process is defined by the fixations on the attributes during the deliberation time that the decision-maker performs before making the choice. The transitions between these attentions allow the construction of the information search patterns that individuals perform in prior to the choice.

Therefore, efforts will be made to understand, analyse, and characterise the external information search process performed by public transport users. This is crucial for modelling discrete choices

since ignoring the real dynamics behind this process will result in inconsistent model parameter estimators due to endogeneity (Guevara, 2015).

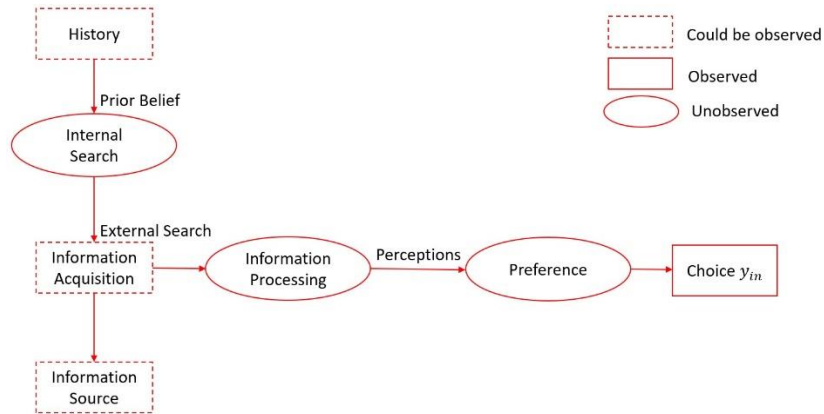


Figure 1: The information search process Xie et al., (2019).

Specifically, four different patterns of information search are considered in the analysis that depend on the transitions of the attributes addressed by the decision-makers in a sequential manner. Definitions like Bettman (1976), Payne (1976), Johnson et al., (2008), Noguchi et al., (2014), Jiang et al., (2016).

- Depth search occurs when the individual conceptualises all the attributes of an alternative before making comparisons with the rest of the options. Thus, attention is expected to fluctuate under different attributes but within the same alternative. The RUM models implicitly include this pattern in the calculation of the choice probabilities since it is assumed that individuals construct the utilities of the alternatives considering the value of all available attributes before choosing.
- Breadth-first search occurs when the individual focuses on a particular attribute and simultaneously updates the value of all available alternatives for comparison. Therefore, it is expected to focus more on one attribute, and transitions occur more frequently between the alternatives. This pattern has been incorporated into different discrete choice models. On the one hand, the RRM model implicitly includes it in the modelling since the probabilities of choice depend on the regret calculated through the bilateral comparisons of attributes (Chorus, 2010). On the other hand, the DFT model explicitly includes this pattern since updating preconceived preferences during the deliberation process is made concerning one attribute at a time (Hancock et al., 2018). It should be noted that this behaviour has been evidenced and supported by the findings of Noguchi et al., (2014), which indicate that comparisons of a pair of alternatives under the same attribute dimension occur more frequently. This is why psychological choice models should be modelled in such a way.
- Unusual searches occur when attention is more erratic than usual. On the one hand, the adjacent diagonal searches capture the transitions that occur towards contiguous attributes and alternatives. In contrast, non-adjacent diagonal transitions occur when attention goes between non-contiguous attributes and alternatives that imply greater cognitive cost.

The figure 2 shows the information search patterns adopted in the different discrete choice models (Inspired from Chorus (2012)). The solid arrows represent the conceptualisations, and the dotted arrows represent the information search process and comparisons of alternatives.

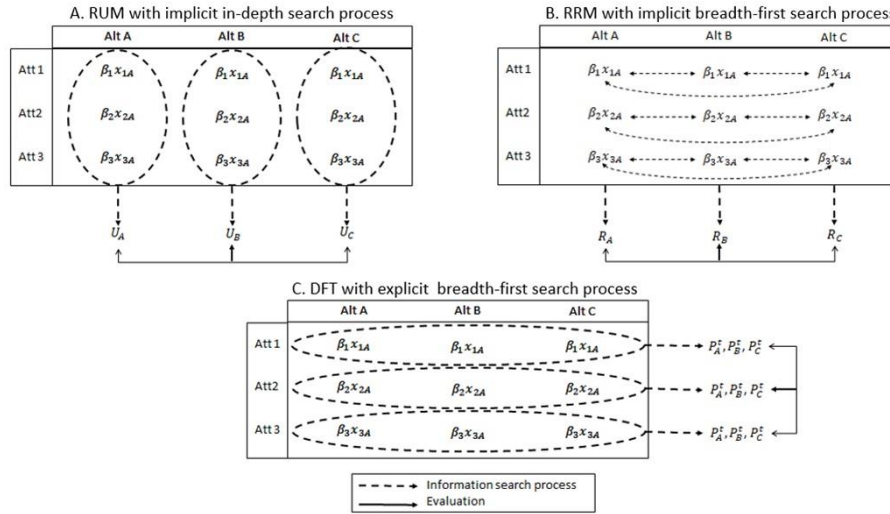


Figure 2: Choice process based on RUM (A), RRM (B) and DFT (C).

3. METHODOLOGY AND PROCEDURES

For studying the information search process, in this research a methodology was developed and applied to collect behavioural evidence of sequential attribute evaluation. Three click-tracking surveys are built that vary in the number of alternatives and attributes displayed as areas of interest (AOI), which are not visible and only one can be displayed at a time, following a Payne (1976)'s information board format.

Click tracking was used to evaluate the information search process as decision makers respond to pivoted Stated Preference (SP) surveys Public Transport trips occurring around the morning peak hour transit route choice study (click tracking), where the areas of interest shown (AOIs) representing the attributes of available alternatives are varied. In general, respondents were asked to report their socioeconomic characteristics, their typical commute, and then to choose one of the hypothetical public transport routes based on their walking time, waiting time, travel time, cost, number of transfers, and seat availability.

After the surveys are designed, they are implemented on a web page and evaluated by applying them in a focus group to verify that the sequence of attributes clicked, the time of each click, and the choices in each task are properly obtained. Likewise, to improve the instrument, comments and feedback are received. Figure 3.A shows the section of socioeconomic questions, figure 3.B shows the mobility questions (revealed preferences), and figure 3.C shows an example of a choice task; containing 3 alternatives and 6 attributes in a click-tracked information board format (declared preferences).

In particular, we measured (1) the amount of information search, which corresponds to the number of fixations on an attribute; (2) the pattern of information search, which reflects the search transitions in the attribute evaluations (in breadth-first search), which can be either by attribute (in-

depth search), by alternative, or by attribute and alternative (diagonal searches); (3) the duration of fixation over the time course of a choice; and (4) the time spent on each choice task.

Figure 3: A) Socio-economic questions. B) Revealed preference questions: Mobility. C) Revealed questions: Information board with Click-Tracking

4. 4.RESULTS AND DISCUSSION

Amount of information search

The information searches carried out by the respondents before the choice, and the mean values, together with their standard deviation, are summarised in the table 1. On average, more information searches were carried out than the total number of areas of interest shown in each instrument, realising that public transport users do not capture the value of the attributes in the first instance and need to reconceptualise these values to include them in their utilities or preferences. Moreover, this construct grows significantly as the number of attributes or alternatives increases. On the other hand, when considering the AIS normalised by the AOI in the different surveys and performing a test of means, it can be noted that a more significant increase in the number of searches is generated when attributes are added ($t=6.36$) compared to when the alternatives are increased ($t=2.71$). Therefore, the amount of information search increases at decreasing rates with the AOIs shown in the surveys, but to a greater extent when attributes are added.

Table 1: Amount of information search in surveys

Survey	J	K	AOI	AIS	AIS/AOI
CT36	3	6	18	19.5 (11.8)	1.1 (0.7)
CT26	2	6	12	14.6 (10.1)	1.2 (0.8)
CT23	2	3	6	9.2 (5.4)	1.6 (0.9)
no differentiated AOI				14.4 (10.2)	1.3 (0.8)

The boxplot with the fixations made by the respondents normalised by the amount of AOI corresponding to each survey reinforces the previous finding since the confidence intervals are different from each other. Furthermore, it is shown that less information is sought by AOI when cognitive load increases. Also, the greatest difference is between the CT23-CT26 surveys, more than the CT26-CT36 surveys, showing that the number of searches intensifies to a greater degree with the attributes increase.

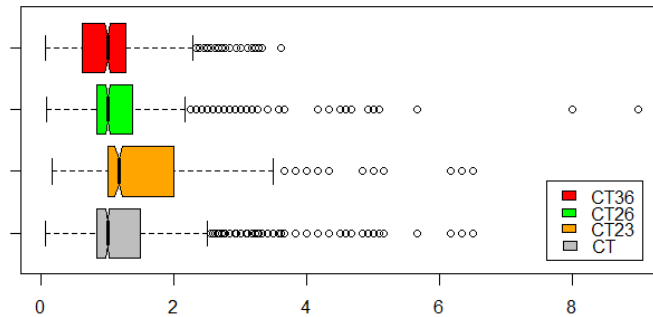


Figure 4: Amount of information search standardised by AOI.

To close this construct analysis, it can be commented that these results are consistent with those found by Meißner et al, (2020), who conclude that the dimensions of the election situation affect the information search process. The findings show that increasing the number of attributes and alternatives leads to an increase in the information search and induces certain filtering of attributes. The novelty of this research lies in the fact that these results are replicated in the public transport route choice SP context, and a greater impact is evidenced in the deliberation process when adding attributes to the number of alternatives.

Filtration

Figure 5 shows the percentage of areas of interest that were not fixed during the deliberation process in each of the surveys presented. In this image, it can be seen, in general, that increasing the number of alternatives or attributes leads to a significant increase in filtered information. This was already evident with the other results for amount information search, which show observations that did not fix on certain attributes or alternatives at any point in time. It can also be seen that the percentage of neglected areas of interest increases as the decision-maker progresses through the choice tasks.

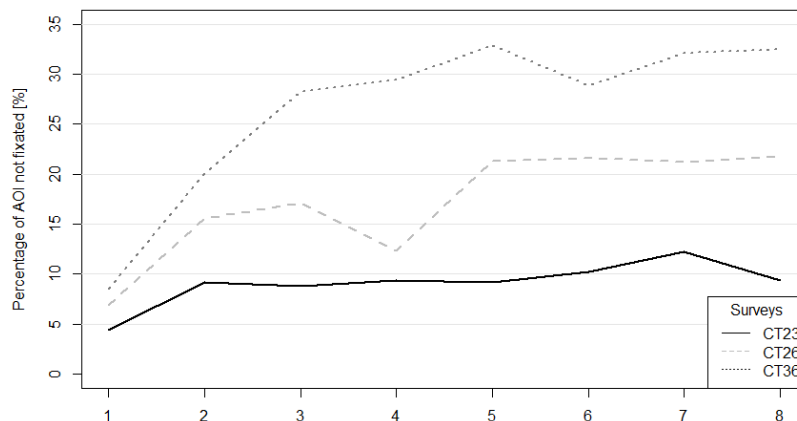


Figure 5: Percentage of areas of interest not fixed in surveys.

Search pattern order

Figures 6, 7, and 8 plot the curves that represent the times in which respondents performed the different information search patterns, at each step of the deliberation process. From this analysis, it is possible to deduce the order in which these types of searches are carried out.

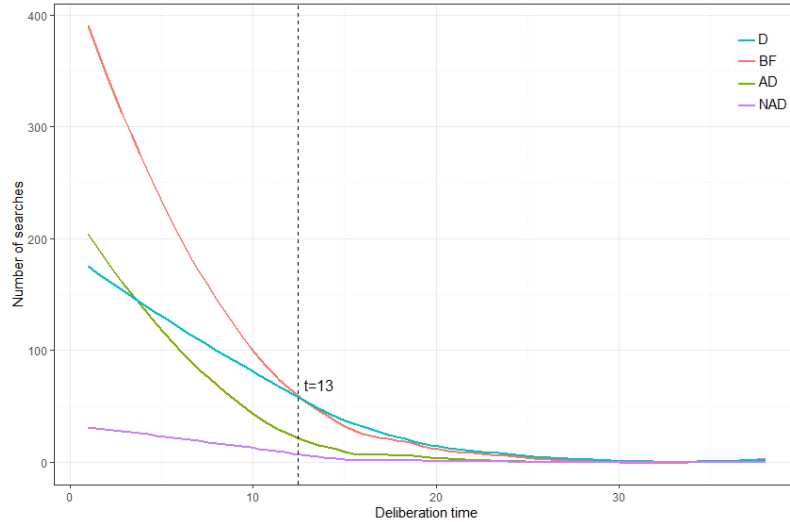


Figure 6: Information search pattern order in CT23.

Firstly, from the CT23 survey (figure 6), there is a greater number of respondents who preferably perform breadth-first searches; in the first instance, and this is maintained during most of the deliberation time (up to $t=13$ there are 83.5% of respondents who have already decided on an alternative). Then until $t=4$, it is followed by adjacent diagonal searches and depth transitions for the rest of the time (until $t=13$). The non-adjacent patterns are smaller and correspond to the minimum value in 92% of the deliberation time concerning the other transitions.

It should be noted that more depth comparisons are made as the deliberation time progresses, which can be evidenced for two reasons. First, the slope of the curve is less than the decay rate of the breadth-first transition. And secondly, the number of depth searches is greater than the rest of the patterns from the first third of the process onwards.

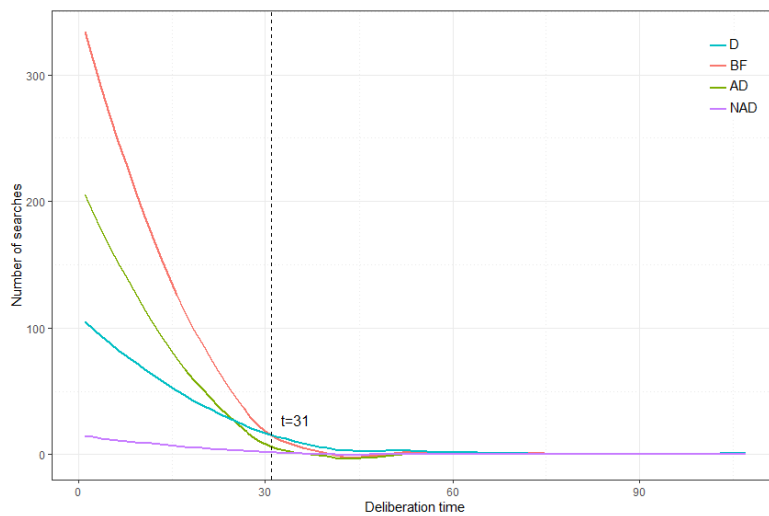


Figure 7: Information search pattern order in CT26.

On the other hand, in the CT26 and CT36 survey, there is a clear predominance of breadth-first search occurrences performed in the first stage of the deliberation time. It stands out, as in the previous case, as time progresses, more depth searches are generated concerning the rest of the patterns, and there is a turning point after the first third of the deliberative process.

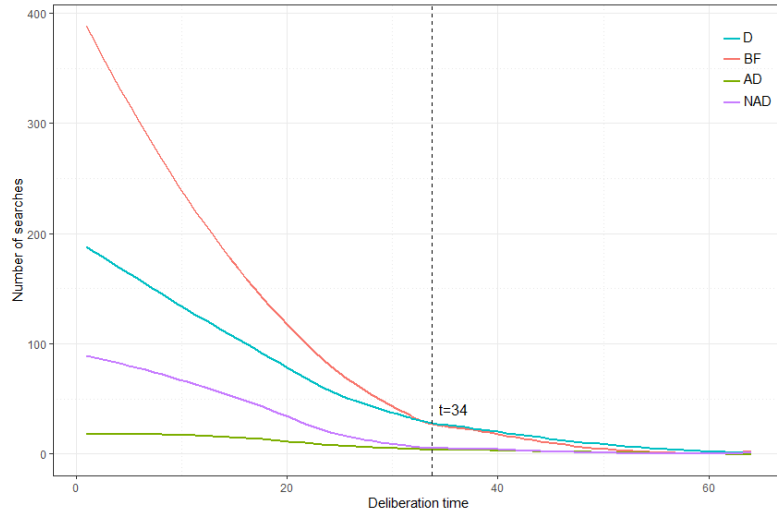


Figure 8: Information search pattern order in CT36.

Transition Matrix

Transition matrices show the relative frequencies between consecutive fixations. These results are summarised in figures 9, 10 and 11, which reveal the probabilities that respondents attend a particular area of interest in timestep t conditional on a previous AOI. From the figures, two important findings can be obtained. First, it is possible to know the aggregate information search patterns that predominate among the participants to reach their final decision. Secondly, it can empirically know the probability of observing an attribute in the next step in the information search process. This allows to include how the evidenced information is acquired and its attention weights in the model to be proposed, which incorporates the sequential evaluation of attributes. Therefore, this modelling allows for the first time to adequately integrate the updating of utilities or preferences for the subsequent comparison of alternatives and choices.

Attribute observed in step t		Attribute observed in step t+1						
		Alt. A			Alt. B			
		WT	TT	C	WT	TT	C	
Alt. A	WT Waiting Time	0	6	1	13	1	0	21%
	TT Travel Time	3	0	4	1	12	1	20%
	C Cost	1	1	0	2	0	8	13%
Alt. B	WT Waiting Time	4	9	0	0	5	0	19%
	TT Travel Time	2	3	6	2	0	4	18%
	C Cost	2	1	3	1	1	0	9%

Figure 9: Transition matrix of CT23

		Attribute observed in step t+1												
		Alternative A						Alternative B						
Attribute observed in step t		WKT	WT	TT	C	T	SA	WKT	WT	TT	C	T	SA	
Alternative A	WKT Walking Time	0	2	0	0	0	0	8	0	0	0	0	0	11%
	WT Waiting Time	1	0	2	0	0	0	0	7	0	0	0	0	11%
	TT Travel Time	0	1	0	1	0	0	0	0	8	0	0	0	11%
	C Cost	0	0	0	0	1	0	0	0	0	6	0	0	8%
	T Transfer	0	0	0	0	0	1	0	0	0	0	5	0	7%
	SA Seat Availability	0	0	0	0	0	0	0	0	0	0	0	4	5%
Alternative B	WKT Walking Time	2	6	0	0	0	0	0	2	0	0	0	0	10%
	WT Waiting Time	1	2	5	0	0	0	1	0	1	0	0	0	10%
	TT Travel Time	0	0	3	4	0	0	0	1	0	1	0	0	10%
	C Cost	0	0	1	3	4	0	0	0	0	0	1	0	8%
	T Transfer	0	0	0	0	1	4	0	0	0	0	0	1	6%
	SA Seat Availability	0	0	0	0	0	1	0	0	0	0	0	0	2%

Figure 10: Transition matrix of CT26.

		Attribute observed in step t+1																		
		Alternative A						Alternative B						Alternative C						
Attribute observed in step t		WKT	WT	TT	C	T	SA	WKT	WT	TT	C	T	SA	WKT	WT	TT	C	T	SA	
Alternative A	WKT Walking Time	0	2	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	7%
	WT Waiting Time	1	0	2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	7%
	TT Travel Time	0	1	0	2	0	0	0	0	4	0	0	0	0	0	1	0	0	0	7%
	C Cost	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	1	0	0	6%
	T Transfer	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	5%
	SA Seat Availability	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	3%
Alternative B	WKT Walking Time	1	0	0	0	0	0	0	2	0	0	0	0	4	0	0	0	0	0	7%
	WT Waiting Time	0	1	0	0	0	0	1	0	2	0	0	0	0	3	0	0	0	0	7%
	TT Travel Time	0	0	1	0	0	0	0	1	0	1	0	0	0	0	3	0	0	0	7%
	C Cost	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	2	0	0	6%
	T Transfer	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	2	0	5%
	SA Seat Availability	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3%
Alternative C	WKT Walking Time	1	2	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	7%
	WT Waiting Time	0	1	2	0	0	0	0	1	0	0	0	0	1	0	2	0	0	0	6%
	TT Travel Time	0	0	1	2	0	0	0	0	1	0	0	0	0	1	0	1	0	0	6%
	C Cost	0	0	0	1	2	0	0	0	0	1	0	0	0	0	0	0	1	0	5%
	T Transfer	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	4%
	SA Seat Availability	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2%

Figure 11: Transition matrix of CT36.

Duration

The boxplots in figure 12 show the average information search time for each of the 8 choice tasks of the CT23, CT26 and CT36 surveys, respectively. Also, the average of these values is highlighted in red. These values are obtained as the amount of time the clicked areas of interest remain visible. From these values, there is a steady decline in average click durations as respondents progress through the survey. This suggests two possible reasons. The respondents acquire knowledge to use the instrument and click faster, or participants begin to memorise the location of the relevant areas of interest to make their choice and perform certain information search heuristics acquired during the previous answered tasks.

Therefore, the duration is not constant during the deliberation process and contradicts the results shown by Stewart et al., (2016). This is because the experiments carried out in this work are more complex. As has been evidenced from the different constructs, different information search patterns are involved, showing instability in information processing. This implies differences in the duration of information acquisition and processing (Rayner et al., 2012). However, short durations and variability throughout the deliberation are more consistent with automatic processes such as accumulating models (Glockner et al., 2011).

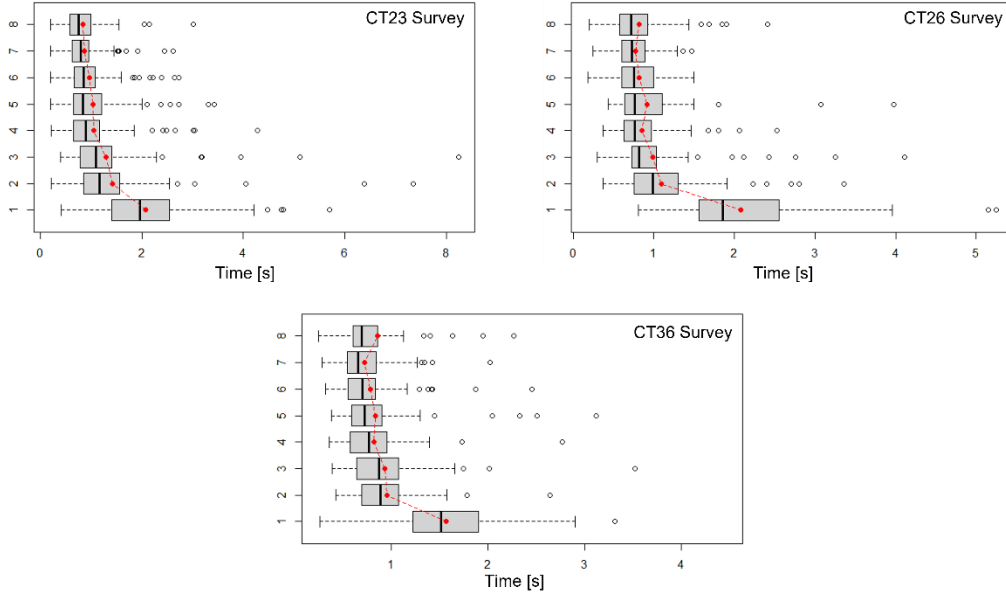


Figure 12: duration per task and survey.

5. CONCLUSIONS AND FUTURE RESEARCH

Based on this analysis, we found 3 main findings. The first, that transit users have a change in the pattern of information search that depends on the number of alternatives and attributes (AOI) shown in the choice situation. On the one hand, as AOIs increase, people search more frequently for information in breadth-first. In addition, we find that users, on average, perform breadth-first and then in-depth search to validate their chosen alternative. A second finding, we found that the sequential evaluation of attributes, number of steps, increases with increasing AOI and a higher number of diagonal searches is observed. Third, doing an analysis of the transition matrices of the AOIs, we find that the most likely transition corresponds to a search for information in breadth-first in all experiments, and that the effect becomes more acute with the number of AOIs.

The evidence found suggests that there is a predominance of information search in breadth-first, so that the RUM model would not be able to adequately describe the choice process and that the DFT approach would be more appropriate, but not necessarily fully comprehensive, for these purposes. Evidence shows that the search for information in breadth-first is not total and this behavior becomes more acute with increasing AOI. Given these results, it seems that a latent class model, incorporating both types of searches (in-depth and breath-first), may be the more suitable to address the problem. The development an assessment of this and other models with the collected data remains as future work.

Acknowledgements

This research was partially funded by the National Fund for Scientific and Technological Development (ANID, FONDECYT) 1191104 and through the grant ANID PIA/BASAL AFB180003.

REFERENCES

- Bettman, J.R., Jacoby, J., 1976. Patterns of processing in consumer information acquisition. *ACR North American Advances*.
- Busemeyer, J., y Townsend, J. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, 100(3), 432.
- Chorus, C.G., 2010. A new model of random regret minimization. *European Journal of Transport and Infrastructure Research* 10.
- Chorus, C., 2012. Random regret minimization: an overview of model properties and empirical evidence. *Transport reviews* 32, 75–92.
- Engel, J., Kollat, D., y Blackwell, R. (1968). *Consumer behavior*. holt. New York: Rinehart and Winston Marketing Series.
- Guevara, C.A., 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice* 82, 240–254.
- Glöckner, A., Herbold, A.K., 2011. An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making* 24, 71–98.
- Hancock, T. O., Hess, S., y Choudhury, C. F. (2018). Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. *Transportation Research Part B: Methodological*, 107, 18–40.
- Hulland, J.S., Kleinmuntz, D.N., 1994. Factors influencing the use of internal summary evaluations versus external information in choice. *Journal of Behavioral Decision Making* 7, 79–102.
- Johnson, E.J., Schulte-Mecklenbeck, M., Willemsen, M.C., 2008. Process models deserve process data: Comment on brandstätter, gigerenzer, and hertwig (2006).
- Jiang, T., Potters, J., Funaki, Y., 2016. Eye-tracking social preferences. *Journal of Behavioral Decision Making* 29, 157–168
- McFadden, D. L. (1976). Quantal choice analysis: A survey. *Annals of Economic and Social Measurement*, Volume 5, number 4, 363-390.
- Meißner, M., Oppewal, H., Huber, J., 2020. Surprising adaptivity to set size changes in multi-attribute repeated choice tasks. *Journal of Business Research* 111, 163–175.
- Noguchi, T., y Stewart, N. (2014). In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition*, 132 (1), 44–56.
- Payne, J.W., 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational behavior and human performance* 16, 366–387.

- Payne, J.W., Bettman, J.R., Johnson, E.J., 1992. Behavioral decision research: A constructive processing perspective. *Annual review of psychology* 43, 87–13.
- Rayner, K., Pollatsek, A., Ashby, J., Clifton Jr, C., 2012. *Psychology of reading*. Psychology Press.
- Riedl, R., Brandstätter, E., Roithmayr, F., 2008. Identifying decision strategies: A process-and outcome-based classification method. *Behavior research methods* 40, 795–807.
- Schulte-Mecklenbeck, M., Johnson, J.G., Böckenholt, U., Goldstein, D.G., Russo, J.E., Sullivan, N.J., Willemsen, M.C., 2017. Process-tracing methods in decision making: On growing up in the 70s. *Current Directions in Psychological Science* 26, 442–450.
- Stewart, N., Gächter, S., Noguchi, T., y Mullett, T. (2016). Eye movements in strategic choice. *Journal of behavioral decision making*, 29 (2-3), 137–156.
- Stewart, N., Hermens, F., y Matthews, W. (2016). Eye movements in risky choice. *Journal of behavioral decision making*, 29 (2-3), 116–136.
- Sui, X.-Y., Liu, H.-Z., y Rao, L.-L. (2020). The timing of gaze-contingent decision prompts influences risky choice. *Cognition*, 195, 104077.
- Xie, Y., et al., 2019. Choice modeling with observed and unobserved information search. Ph.D. thesis. Massachusetts Institute of Technology.
- Qin, H., Guan, H., y Wu, Y.-J. (2013). Analysis of park-and-ride decision behavior based on decision field theory. *Transportation research part F: traffic psychology and behaviour*, 18 , 199–212.

Power sector effects of alternative options for electrifying heavy-duty vehicles

Carlos Gaete-Morales^{*1}, Julius Jöhrens², Florian Heining², and Wolf-Peter Schill¹

¹Energy, Transportation, Environment Department, German Institute for Economic Research (DIW Berlin), Mohrenstraße 58, Berlin, 10117, Germany

²ifeu - Institute for Energy and Environmental Research, Wilckensstraße 3, Heidelberg, 69120, Germany

SHORT SUMMARY

In the passenger car segment, battery-electric vehicles have emerged as the most promising option to decarbonize transportation. For heavy-duty vehicles, the technology space still appears to be more open. Aside from stationary-charged battery-electric trucks, electric road systems (ERS) for dynamic power transfer to electric vehicles are also discussed, as well as trucks that use hydrogen fuel cells or e-fuels. Here we investigate the power sector implications of these different options. We apply an open-source power sector capacity expansion model to future scenarios of Germany with high renewable energy shares, drawing on detailed route-based truck mobility data. Results show that power sector costs are highest in the case of e-fuels, and lowest for battery-electric and ERS trucks. The latter technologies can generally provide more temporal flexibility to the power sector than battery-electric and ERS trucks. Yet, these flexibility benefits do not outweigh their disadvantages in terms of energy efficiency. In equilibrium, the different flexibility characteristics lead to higher capacity expansion and use of solar PV for battery-electric and ERS trucks, and to a higher use of wind power for hydrogen and e-fuel trucks. If battery-electric and ERS trucks are charged in a non-optimized manner, power sector costs increase, but still remain below those of hydrogen and e-fuel trucks.

Keywords: Battery-electric vehicles, Catenary, Electrification and decarbonization of transport, Heavy-duty vehicles, Hydrogen, Power sector modeling.

1 INTRODUCTION

Making energy consumption climate neutral in all end-use sectors is of paramount importance for mitigating climate change de Coninck et al. (2018). A key strategy for achieving this is to substitute fossil fuels by renewable electricity, facilitated by direct or indirect electrification of end uses in mobility, heating, and industrial applications Shukla et al. (2022). In the transportation sector, battery-electric vehicles (BEV) have emerged as the most promising option for the passenger car segment. Already today, BEV can lead to sizeable greenhouse gas emission reductions compared to internal combustion engines Hoekstra (2019), which will further increase when the electricity mix becomes cleaner. In many countries, markets for electric passenger cars have been soaring in the past years, and are expected to continue to grow strongly in the near future IEA (2022). For heavy-duty vehicles (HDV), however, the technology space still appears to be more open. While the feasibility of pure battery-electric HDV has been assessed to be increasing Nykvist & Olsson (2021), they compete with other options. This includes electric road systems (ERS), which allow for dynamic power transfer to electric vehicles on the road (Boltze et al., 2020; Speth & Funke, 2021); trucks with hydrogen fuel cells; or conventional HDV with internal combustion engines that use liquid e-fuels which are produced with renewable electricity Hannula & Reiner (2019); Lajevardi et al. (2022); Plötz (2022); Li et al. (2022).

These options of direct or indirect electrification of HDV have different properties concerning, on the one hand, energy efficiency, and, on the other hand, temporal flexibility of electricity use. For example, direct electrification via BEV is more energy efficient compared to indirect electrification via electrolysis-based hydrogen or e-fuels Ueckerdt et al. (2021); Lajevardi et al. (2022). Yet, the temporal flexibility of BEV may be constrained by charging availability and limited battery capacities, as vehicle batteries are costly and heavy. In contrast, indirect electrification via hydrogen

or e-fuels may entail large-scale and low-cost storage options Taljegard et al. (2017); Stöckl et al. (2021), but the overall energy efficiency of these supply chains is lower compared to BEV. Temporal power sector flexibility becomes increasingly important with growing shares of renewables, as the potential for firm renewable generation such as hydropower, bioenergy, or geothermal power is limited in many countries. In contrast, wind and solar power potentials are often abundant, but they have variable generation profiles that depend on weather conditions and daily and seasonal cycles (López-Prol & Schill, 2021). Integrating growing shares of such variable renewables thus requires an increasing use of flexibility options in the power sector (Kondziella & Bruckner, 2016).

Against this background, we investigate the power sector implications of different options for (in-)directly electrifying HDV, particularly focusing on the trade-off between energy efficiency and temporal flexibility. To do so, we apply an open-source capacity expansion model Zerrahn & Schill (2017); Gaete-Morales, Kittel, et al. (2021) to 2030 scenarios of the Central European power sector with high renewable energy shares. We focus on the domestic traffic of HDV in Germany with a gross vehicle weight above 26 tonnes, drawing on a detailed data set of truck trips on inner-German origin-destination pairs. We include stationary-charged BEV trucks as well as hybrid battery-catenary trucks as a particular example of an electric road system technology (ERS-BEV), fuel-cell hydrogen electric trucks (FCEV), and such with internal diesel combustion engines powered by e-fuels (ICEV PtL). For hydrogen, we further differentiate two domestic supply chains, either decentralized electrolysis at filling stations, which is temporally inflexible, or centralized electrolysis and transport via gaseous hydrogen, which also comes with low-cost storage opportunities and is thus more flexible. We compare the power sector costs of these options, as well as their repercussions on the optimal power plant fleet, under different assumptions on the temporal flexibility of the electric load of electric HDV usage.

While there is a broad literature on the potential power sector impacts of battery-electric passenger cars Richardson (2013); Muratori & Mai (2020); Mangipinto et al. (2022), according research for electric HDV is sparse Schill & Gerbaulet (2015); Gnann et al. (2018); Sadeghian et al. (2022); Pickering et al. (2022).

We contribute to the literature by providing, to the best of our knowledge, the first analysis that co-optimizes the charging and discharging operations (including V2G) of different types of electrified HDV with capacity and dispatch decisions in the power sector. We do so for a wide range of HDV technologies, including dynamic power supply via electric road systems. We use a power sector model that fully captures the hourly variability of load and renewable generation over all hours of a full year, and apply it to a future scenario with high shares of variable renewables. The model code and all input data, including detailed hourly HDV mission profiles for domestic transport in Germany, are provided open source for transparency and reproducibility.

2 METHODOLOGY

The power sector model DIETER

We use the open-source power sector model *Dispatch and Investment Evaluation Tool with Endogenous Renewables* (DIETER). It is a linear program that minimizes power sector costs by optimizing capacity and dispatch decisions for a full year in an hourly resolution Zerrahn & Schill (2017); Gaete-Morales, Kittel, et al. (2021). Its objective function includes fixed and variable costs of all electricity generation and storage technologies, electrolysis and PtL plants, as well as hydrogen or e-fuel transportation. It does not include the costs of charging or catenary infrastructure, hydrogen filling stations, or PtL filling stations. Accordingly, the power sector cost figures provided above do not include the costs of HDV electrification infrastructure. We further do not consider the option of hydrogen imports, as these are likely to be unavailable at scale by 2030. In general, the global scaling up of green hydrogen supply remains uncertain Odenweller et al. (2022).

Endogenous model variables include power sector costs, optimal generation and electricity storage capacities (Germany) and their hourly use (all countries), hourly decisions for HDV charging and discharging, as well as the capacity and operational decisions of electrolysis and PtL generation and storage infrastructure. In addition, we interpret the marginals of the hourly energy balance as wholesale prices (compare (Brown et al., 2018)).

Exogenous model inputs include fixed and variable costs of all electricity generation and storage technologies, efficiency parameters, as well as time-series variable renewable energy availability profiles and electric load. In the case of inflexible HDV charging (BEV Inflex, ERS-BEV Inflex), we assume that the vehicles always start charging as soon as an opportunity arises, and that vehicles batteries are fully charged by the time the next trip starts, if possible. The charging power is set to facilitate exactly this for each charging period, i.e., charging power is generally lower, the longer a vehicle is connected to the grid. This resembles the “balanced” charging profile defined in Gaete-Morales, Kramer, et al. (2021). For Germany, we further assume upper limits for investments in fossil generation capacities as given by the federal Grid Development Plan 2030.

The geographic scope of the model version used here includes Germany and its neighboring countries plus Italy. In order to reduce numerical complexity and improve tractability, we allow for endogenous generation capacity investment only in Germany, and fix the power plant portfolio for the other countries to values derived from ENTSO-E’s Ten Year Network Development Plans ENTSOE (2018). The model is required to satisfy at least 80 percent of the load in Germany with domestic renewable electricity generation. This includes the additional load related to directly or indirectly electrifying HDV. This reflects the current German government’s target for 2030 that has also been set out in the Renewable Energy Sources Act.

Mobility data of heavy-duty vehicles

We generate synthetic truck usage patterns that are intended to approximate the German fleet of HDV larger than 26 tons. The main data source for the usage patterns is the traffic model PTV Validate, from which we extract a database of daily truck trips in domestic German road freight transport.

For these profiles, the time series of charging availability (in the depot, during idle and driver’s resting times) and of the electricity demand of pure BEV-HDV (500 km battery range) and ERS-HDV (150 km battery range) are calculated. The resulting electricity demands and charging availabilities are used as inputs for the DIETER model.

3 RESULTS AND DISCUSSION

Lowest power sector costs and electricity prices for BEV with V2G

Compared to the reference case without electrified HDV, yearly power sector costs increase in all scenarios with electrified HDV (Figure 1, upper panel). That is, the cost of the additional electricity demand induced by HDVs always outweighs their potential flexibility benefits. Cost effects, however, vary strongly between different options. Flexible BEV with V2G incur the lowest additional power sector costs (1.8 bn Euros/year, or around 5,600 Euros/year per vehicle), followed by BEV without V2G (2.3 bn Euros/year, or around 7,200 Euros/year per vehicle). If BEV charging is not optimized, system costs are markedly higher (3.8 bn Euros/year, or 11,900 Euros/year per vehicle). Results are qualitatively similar for ERS-BEV, but on a slightly higher cost level. The differences between the three ERS-BEV cases are much less pronounced than for pure BEV, as their temporal flexibility potential is much smaller. The battery capacity of an ERS-BEV fleet is only around a quarter of that of an alternative pure BEV fleet (655 kWh usable capacity per pure BEV and 181 kWh per ERS-BEV truck). In contrast, power sector cost are substantially higher for FCEV (12.6 or 12.7 bn Euros/year, for decentralized or centralized hydrogen provision, i.e. around 39,700 Euros/year per vehicle) and even more so for PtL (16.8 bn Euros/year, or 52,700 Euros/year per vehicle). This is a direct consequence of high conversion losses of hydrogen and PtL supply chains and vehicles drive trains. Because of these losses, the two hydrogen supply chains increase the electricity demand more than twice as much as the battery-electric options. The electricity demand of PtL-HDV is nearly four times as high as in the case of BEV. Notably, the cost differences between BEV and ERS-BEV are much smaller than the differences between these direct-electric options and indirect electrification via hydrogen or PtL.

Complementary to power sector costs, we also evaluate average yearly wholesale electricity prices

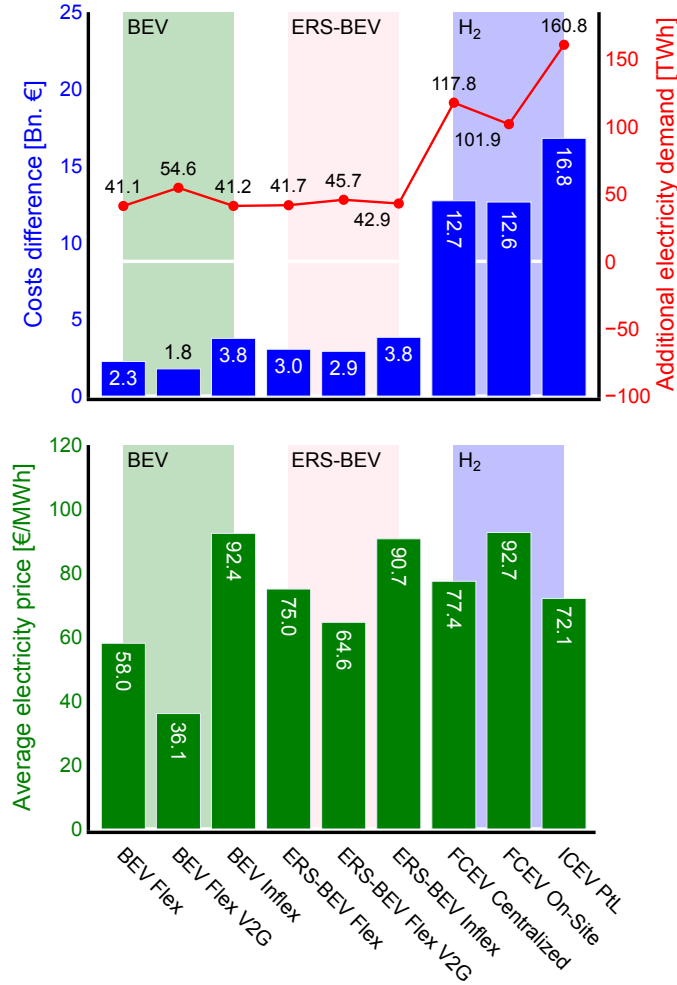


Figure 1: Changes in yearly power sector costs and electricity demand induced by different HDV options (upper panel), and average wholesale market prices of charging electricity (lower panel).

for HDV electricity (Figure 1, lower panel). This allows to largely separate the differences in overall electricity consumption of the various HDV options from their ability to make use of low-cost electricity. Average prices are calculated by multiplying hourly wholesale prices of electricity consumed by the different HDV options or fed back to the grid with respective hourly quantities, summing up over the whole year, and dividing by the overall electricity consumption of the fleet. That is, the numbers also account for revenues of electricity sold via V2G.¹ BEV with V2G face the lowest average electricity prices, as these also benefit from revenues of feeding back to the grid, followed by BEV without V2G. Average electricity prices paid by ERS-BEV are somewhat higher, as their smaller batteries limit the ability for temporally optimizing their charging and V2G decisions. In contrast, pure BEV can leverage their larger battery capacity to make better use of hours with low electricity prices.

In contrast, average electricity prices faced by inflexibly charged BEV are high, and even slightly above those of inflexible ERS-BEV. Note that inflexibly charged BEV generally benefit less from cheap electricity prices around midday related to abundant PV feed-in than inflexible ERS-BEV, while they are driving. With flexible charging, BEV can compensate for this charging availability disadvantage at midday by making better use of low prices in other periods, e.g., in windy nights, leveraging their larger batteries. In the case of inflexible charging, however, this is no longer possible, and the higher day-time grid availability of ERS-BEV gives them a slight competitive edge over BEV in terms of average charging prices. Prices for electricity used in hydrogen and

¹Here we assume that HDV operators receive the respective hourly wholesale price whenever they feed electricity back to the grid.

PtL supply chains are in the same range as those of ERS-BEV options. Centralized hydrogen and PtL supply can make use of lower prices than decentralized supply, because their low-cost storage options allow for higher temporal flexibility. Electricity prices of centralized hydrogen supply chains and PtL are also cheaper than those faced by inflexibly operated BEV or ERS-BEV. In terms of overall costs, these temporal flexibility benefits are, however, by far outweighed by the higher overall energy consumption of the FCEV and PtL options (Figure 1, upper panel).

Capacity and dispatch effects

The upper panel of Figure 2 shows optimal generation capacities in the reference (on the left) and the changes induced by HDV (on the right) in Germany, where minimum renewable energy share of 80% applies. In the reference, variable solar PV and onshore wind power dominate the capacity mix. These are complemented by smaller firm capacities of natural gas and bioenergy. The capacity additions related to the electrification of HDVs are predominantly a mix of solar PV and onshore wind power. Flexible BEV and ERS-BEV lead to the highest PV shares in the capacity additions, especially if combined with V2G. BEV with V2G essentially serve as short-duration grid storage, which favors the expansion of solar PV. HDV options that are temporally less flexible or that, overall, require more electricity favor higher onshore wind power capacities. If, alternatively, more PV was built, this would lead to increasing amounts of unused renewable surplus energy. Offshore wind power is not added here because of relatively unfavourable costs. FCEV and PtL options have the highest capacity needs because of substantial conversion losses. Decentralized electrolysis further requires a substantial addition of long-duration electricity storage capacity (9.4 GW) to compensate for the temporal inflexibility of the additional electricity demand (10.3 TWh). Aside from natural gas and oil, no fossil fuel generation capacities are used, partly due to a CO₂ price of 100 Euros/ton which discourages such investments. However, in the more inflexible scenarios (BEV Inflex, ERS-BEV Inflex and FCEV Distributed), investments into natural gas and oil generation capacity are at the assumed maxima.

The changes in yearly electricity generation are shown in the lower panel of Figure 2. On the left-hand side the overall electricity generation of the reference scenario is displayed. Whereas the right-hand side figure shows electricity additions or reductions according to the HDV options and generating technologies.² Here, the share of wind power in additional electricity generation is higher than in additional capacity because of its higher full-load hours as compared to solar PV.

Flexible BEV and ERS-BEV options show large solar PV generation, the lowest gas and oil generation and an inferior electricity import that led them to reach a higher renewable share (above 81%) than the reference case (80%). Centralized FCEV and PtL increased the gas power output of the most efficient plants, the combined-cycle ones, even without investing in more capacity leading to an increase in the capacity factor and maintaining the same low renewable share as the reference scenario (80%). As the coal power investment was disincentivized due to the high carbon price and having reached the maximum investment in gas and oil power, FCEV distributed option was forced to increase the share of renewables to 82.6% to fulfil the highest energy consumption. It also replaced in part power import by long-term storage, P2G2P, to overcome the inflexibility that distributed hydrogen entailed to the power sector.

Direct carbon emissions

Among all scenarios, CO₂ emissions increase the most if the HDV fleet uses e-fuels or hydrogen (Figure 3). This is a consequence of additional electricity generation from natural gas in these scenarios. Among the two hydrogen cases, on-site electrolysis at filling stations leads to lower emissions impacts compared to centralized electrolysis, as its temporal flexibility limitations require additional long-duration electricity storage, which in turn is charged to a substantial extent by renewable surplus energy. Emission effects are smaller for BEV and ERS-BEV, and even negative for flexibly charged BEV, especially if combined with V2G. The latter is driven by an additional expansion of solar PV facilitated by V2G, as shown above. For neighboring countries, relative emission effects are smaller, as by assumption they have lower renewable energy shares and, in turn, higher emissions, as well as no electrified truck fleets.

²This chart also shows the output power associated with storage options. These figures do not represent generation as long as it corresponds to energy throughput.

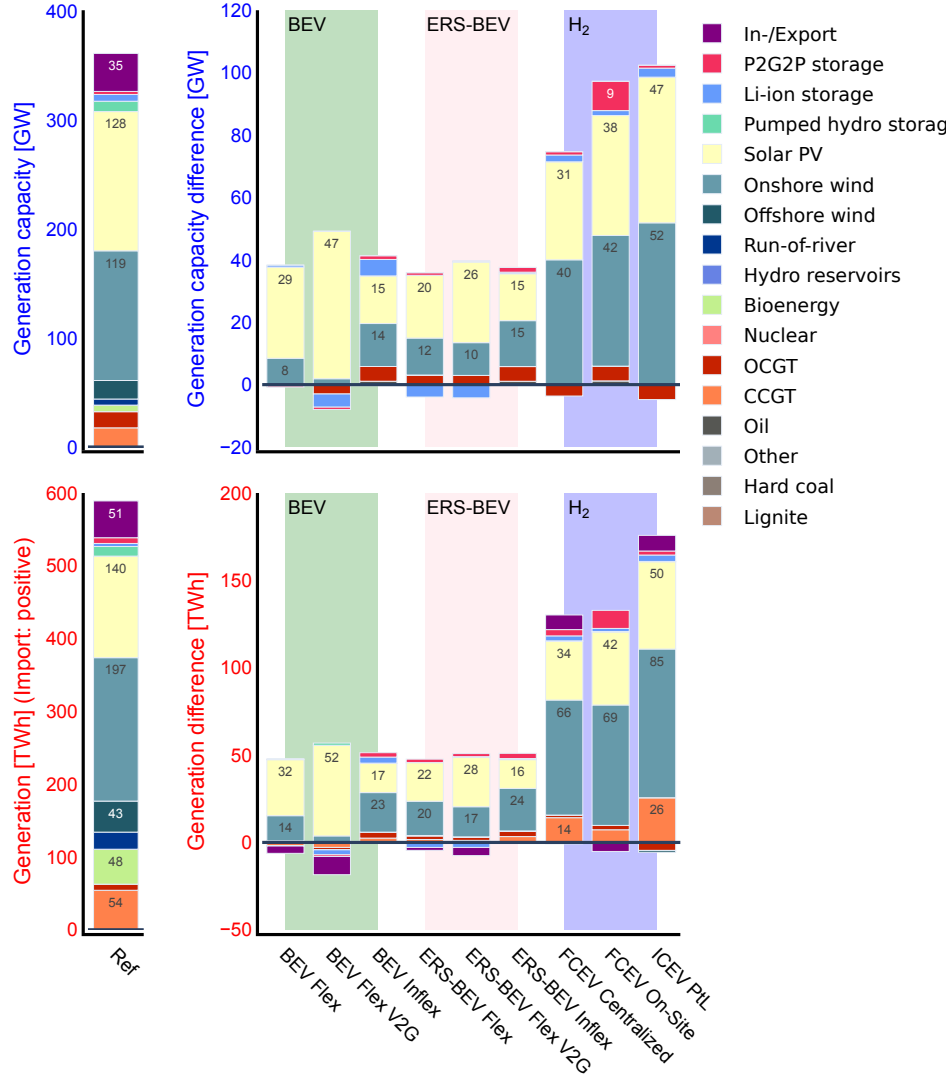


Figure 2: Effects of different HDV scenarios on optimal generation capacity (upper panel) and on yearly generation (lower panel) in Germany

Time series reveal differences in flexibility characteristics

Figure 4 illustrates that flexible BEV HDV are able to charge their batteries in hours of low residual load, and especially to make use of renewable surplus energy to a substantial extent. They also make use of the V2G option to some extent to feed back renewable surplus energy to the grid, whenever the battery capacity is not needed for driving. In the exemplary illustration, this is visible particularly in summer (right panel). Note that the time series shown begins on a Saturday.³ BEV with V2G store a substantial amount of renewable (i.e., solar PV) surplus energy on Saturday afternoon, and feed it back to the grid in the night between Saturday and Sunday (top right time series shown in Figure 4). This is possible because HDV are not used on Sunday, so the battery capacity is idle. Note that this is different in the following days, as HDV are used between Monday and Friday, and much less battery capacity is available for V2G. If BEV charging is not optimized, but follows an inflexible, pre-determined pattern, charging profiles are less peaky and much more balanced (second panels from top). This especially means that BEV are not able to make much use of cheap renewable surplus electricity in this case, but also carry out a lot of charging in hours with positive residual load.

ERS-BEV generally follow similar patterns as non-catenary BEV. Yet, their smaller batteries make ERS-BEV temporally less flexible, so they can make less use of renewable surplus events and also

³For simplification, we assume that both Saturday and Sunday are truck-free.

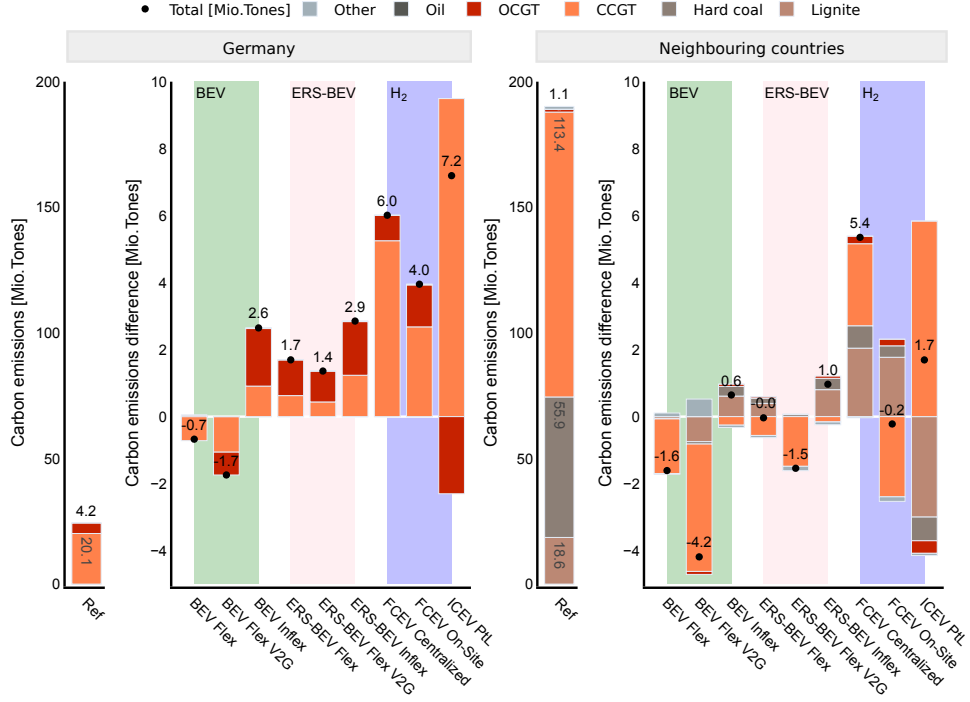


Figure 3: Direct CO₂ emissions from electricity generation. The left panel shows annual emissions from Germany, while in the right panel are the emissions from Germany’s neighbouring countries. Each panel contains, on the left, the overall emissions of a reference scenario with no trucks. On the right is the emissions difference of the scenarios with direct or indirect power demand.

have to draw electricity from the grid during hours with positive residual load to some extent. For the same reason, their potential for feeding electricity back to the grid is also much smaller than in the case of pure BEV. Note that ERS-BEV partly also charge their batteries during driving.

Hydrogen and PtL supply chains show very different patterns of electricity use compared to BEV or ERS-BEV. FCEV with centralized hydrogen supply chains (i.e., centralized electrolyzers with large-scale hydrogen storage capacities) generally have a flat consumption profile in many hours, as their high fixed costs make it optimal to use them with relatively high full load hours. This also limits their ability to make use of renewable surplus energy (lower electricity consumption on first summer day shown in the graph compared to BEV V2G). Yet, they can use the temporal flexibility provided by centralized, large-scale hydrogen storage to reduce electricity consumption in hours of high positive residual load, i.e., high prices. In contrast, on-site electrolysis follows the actual hydrogen demand much more closely. This is because decentralized electrolyzers sited at filling stations by assumption only come with very limited hydrogen storage. Accordingly, they can avoid electricity consumption in hours of high residual load only to a minor extent, and much less than centralized electrolyzers. The PtL supply chain has a relatively similar pattern as the one for centralized hydrogen. The peak load is however higher because of higher overall energy consumption which also goes along with the higher PEM electrolysis capacity (24.4 GW).⁴

4 CONCLUSIONS

We analyze the power sector effects of alternative options for electrifying heavy-duty vehicles in Germany, focussing on power sector costs, investment decisions, dispatch and direct carbon emission in the power sector. Temporal flexibility and energy efficiency are important drivers of results.

⁴PEM electrolyzer capacity. FCEV centralized: 13.5 GW, FCEV distributed: 19.1 GW, ICEV PtL: 24.4 GW.

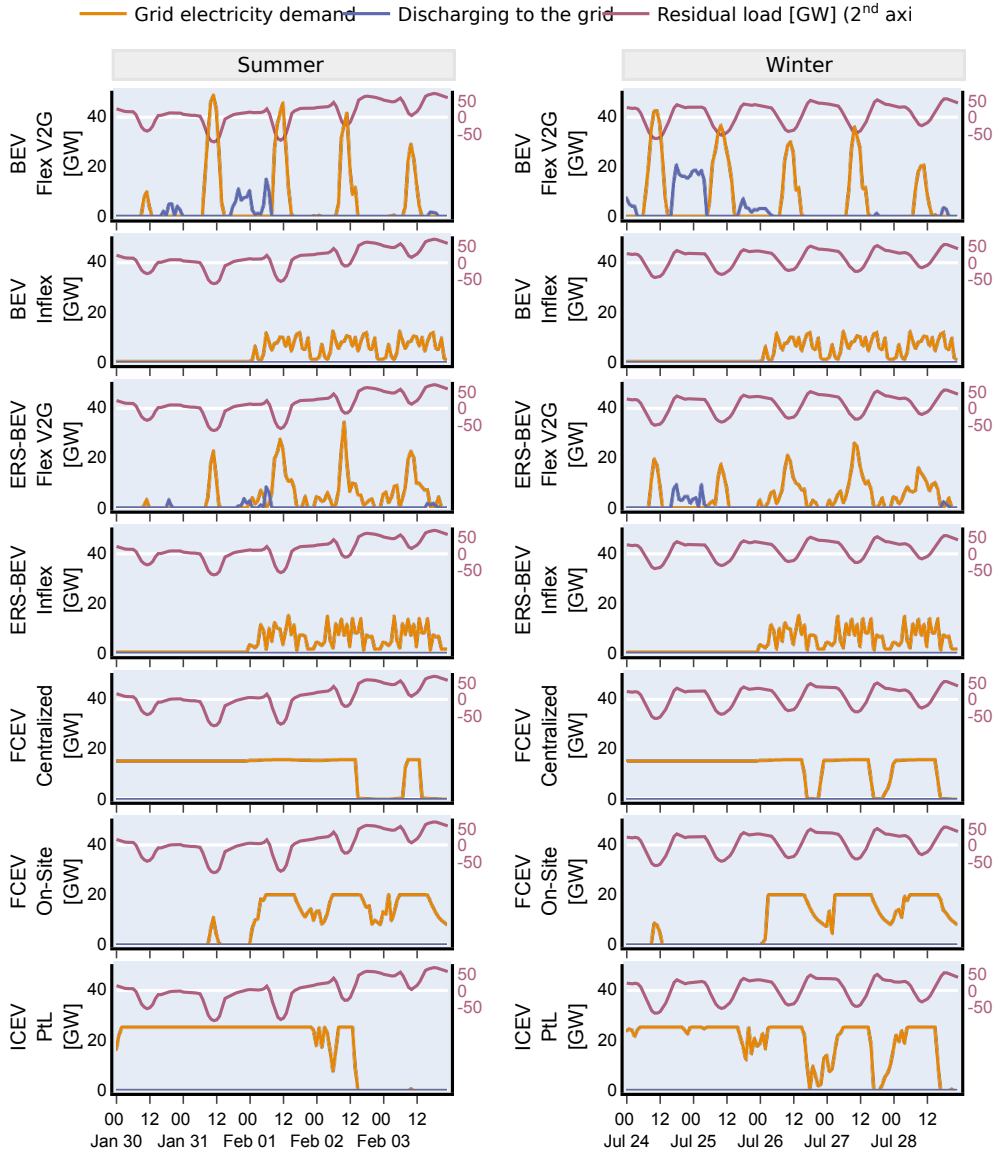


Figure 4: Five days sample for winter and summer of electricity flow time series. It contains residual load, electricity consumed to fulfil the demand for mobility (loading) and electricity returned to the grid in case of Vehicle to Grid (V2G) in GW. The samples start on Saturday.

Flexibly operated BEV and ERS-BEV, especially if combined with vehicle-to-grid, lead to the lowest power sector costs because of their energy efficiency benefits. In contrast, FCEV and PtL are temporally more flexible, but this does not outweigh their energy efficiency drawbacks.

From a pure power sector perspective, direct electrification of the truck fleet would be clearly preferable. Moreover, temporally flexible charging, including V2G, is desirable, as this leads to the lowest electricity sector costs and carbon emissions, and the highest use of renewable electricity. In contrast, inflexible charging should be discouraged as it performs poorly compared to flexible and bi-directional charging.

Future research may investigate overall system cost effects, also considering cost differences of charging and ERS infrastructures, as well as purchase cost differences of the trucks, which was beyond the scope of this study.

ACKNOWLEDGEMENTS

This work has been supported by a research grant of the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) via the projects “My eRoads”.

REFERENCES

- Boltze, M., Linke, R., Schöpp, F., Wilke, J., Öztürk, O., & Wauri, D. (2020, 05). Insights into the operation of overhead line hybrid trucks on the elisa test track. In *Ersc2020 - 4th electric road systems conference* (p. 1-6).
- Brown, T., Schlachtberger, D., Kies, A., Schramm, S., & Greiner, M. (2018). Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system. *Energy*, *160*, 720-739. doi: 10.1016/j.energy.2018.06.222
- de Coninck, H., Revi, A., Babiker, M., Bertoldi, P., Buckeridge, M., Cartwright, A., ... Sugiyama, T. (2018). Strengthening and Implementing the Global Response. In V. Masson-Delmotte et al. (Eds.), *Global Warming of 1.5°C. An IPCC Special Report on the Impacts of Global Warming of 1.5°C Above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (pp. 313-444). Cambridge University Press. doi: 10.1017/9781009157940.006
- ENTSOE. (2018). *TYNDP 2018. Project Sheets*. <https://tyndp.entsoe.eu/tyndp2018/projects/projects>.
- Gaete-Morales, C., Kittel, M., Roth, A., & Schill, W.-P. (2021). DIETERpy: a Python framework for the Dispatch and Investment Evaluation Tool with Endogenous Renewables. *SoftwareX*, *15*, 100784. doi: 10.1016/j.softx.2021.100784
- Gaete-Morales, C., Kramer, H., Schill, W.-P., & Zerrahn, A. (2021). An open tool for creating battery-electric vehicle time series from empirical data, emobpy. *Scientific Data*, *8*(1), 152. doi: 10.1038/s41597-021-00932-9
- Gnann, T., Klingler, A.-L., & Kühnbach, M. (2018). The load shift potential of plug-in electric vehicles with different amounts of charging infrastructure. *Journal of Power Sources*, *390*, 20-29. doi: 10.1016/j.jpowsour.2018.04.029
- Hannula, I., & Reiner, D. M. (2019). Near-term potential of biofuels, electrofuels, and battery electric vehicles in decarbonizing road transport. *Joule*, *3*(10), 2390-2402. doi: 10.1016/j.joule.2019.08.013
- Hoekstra, A. (2019). The underestimated potential of battery electric vehicles to reduce emissions. *Joule*, *3*(6), 1412-1414. doi: 10.1016/j.joule.2019.06.002
- IEA. (2022). *Global EV outlook 2022*. <https://www.iea.org/reports/global-ev-outlook-2022>. (Paris)
- Kondziella, H., & Bruckner, T. (2016). Flexibility requirements of renewable energy based electricity systems - a review of research results and methodologies. *Renewable and Sustainable Energy Reviews*, *53*, 10-22. doi: <https://doi.org/10.1016/j.rser.2015.07.199>
- Lajevardi, S. M., Axsen, J., & Crawford, C. (2022). Simulating competition among heavy-duty zero-emissions vehicles under different infrastructure conditions. *Transportation Research Part D: Transport and Environment*, *106*, 103254. doi: 10.1016/j.trd.2022.103254
- Li, S., Djilali, N., Rosen, M. A., Crawford, C., & Sui, P.-C. (2022). Transition of heavy-duty trucks from diesel to hydrogen fuel cells: Opportunities, challenges, and recommendations. *International Journal of Energy Research*, *46*(9), 11718-11729. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/er.8066> (_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/er.8066>) doi: <https://doi.org/10.1002/er.8066>

- López-Prol, J., & Schill, W.-P. (2021). The economics of variable renewable energy and electricity storage. *Annual Review of Resource Economics*, 13(1), 443–467. doi: 10.1146/annurev-resource-101620-081246
- Mangipinto, A., Lombardi, F., Sanvito, F. D., Pavičević, M., Quoilin, S., & Colombo, E. (2022). Impact of mass-scale deployment of electric vehicles and benefits of smart charging across all european countries. *Applied Energy*, 312, 118676. Retrieved 2022-10-06, from <https://www.sciencedirect.com/science/article/pii/S0306261922001416> doi: 10.1016/j.apenergy.2022.118676
- Muratori, M., & Mai, T. (2020, December). The shape of electrified transportation. *Environmental Research Letters*, 16(1), 011003. Retrieved 2022-10-06, from <https://doi.org/10.1088/1748-9326/abc38-9326/abc38> doi: 10.1088/1748-9326/abc38
- Nykvist, B., & Olsson, O. (2021). The feasibility of heavy battery electric trucks. *Joule*, 5(4), 901–913. doi: 10.1016/j.joule.2021.03.007
- Odenweller, A., Ueckerdt, F., Nemet, G. F., Jensterle, M., & Luderer, G. (2022). Probabilistic feasibility space of scaling up green hydrogen supply. *Nature Energy*, 7(9), 854–865. doi: 10.1038/s41560-022-01097-4
- Pickering, B., Lombardi, F., & Pfenninger, S. (2022). Diversity of options to eliminate fossil fuels and reach carbon neutrality across the entire european energy system. *Joule*, 6(6), 1253–1276. doi: 10.1016/j.joule.2022.05.009
- Plötz, P. (2022). Hydrogen technology is unlikely to play a major role in sustainable road transport. *Nature Electronics*, 5(1), 8–10. doi: 10.1038/s41928-021-00706-6
- Richardson, D. B. (2013). Electric vehicles and the electric grid: A review of modeling approaches, impacts, and renewable energy integration. *Renewable and Sustainable Energy Reviews*, 19, 247–254. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1364032112006557> doi: <https://doi.org/10.1016/j.rser.2012.11.042>
- Sadeghian, O., Oshnoei, A., Mohammadi-ivatloo, B., Vahidinasab, V., & Anvari-Moghaddam, A. (2022). A comprehensive review on electric vehicles smart charging: Solutions, strategies, technologies, and challenges. *Journal of Energy Storage*, 54, 105241. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2352152X22012403> doi: <https://doi.org/10.1016/j.est.2022.105241>
- Schill, W.-P., & Gerbaulet, C. (2015). Power system impacts of electric vehicles in germany: Charging with coal or renewables? *Applied Energy*, 156, 185–196. doi: 10.1016/j.apenergy.2015.07.012
- Shukla, P., et al. (Eds.). (2022). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press. doi: 10.1017/9781009157926
- Speth, D., & Funke, S. A. (2021). Comparing options to electrify heavy-duty vehicles: Findings of german pilot projects. *World Electric Vehicle Journal*, 12(2), 67. doi: 10.3390/wevj12020067
- Stöckl, F., Schill, W.-P., & Zerrahn, A. (2021). Optimal supply chains and power sector benefits of green hydrogen. *Scientific Reports*, 11(1), 14191. doi: 10.1038/s41598-021-92511-6
- Taljegard, M., Göransson, L., Odenberger, M., & Johnsson, F. (2017). Spatial and dynamic energy demand of the E39 highway – Implications on electrification options. *Applied Energy*, 195, 681–692. doi: 10.1016/j.apenergy.2017.02.025
- Ueckerdt, F., Bauer, C., Dirnaichner, A., Everall, J., Sacchi, R., & Luderer, G. (2021). Potential and risks of hydrogen-based e-fuels in climate change mitigation. *Nature Climate Change*, 11(5), 384–393. doi: 10.1038/s41558-021-01032-7
- Zerrahn, A., & Schill, W.-P. (2017). Long-run power storage requirements for high shares of renewables: review and a new model. *Renewable and Sustainable Energy Reviews*, 79, 1518–1534. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1364032116308619> doi: <https://doi.org/10.1016/j.rser.2016.11.098>

Detection of Bus Bunching through the Analysis of Prevalent Public Transport Control Data

Leon Weinsziehr*¹, Antonios Tsakarestos², Frederik Bachmann², Klaus Bogenberger²

*Corresponding author: leon.weinsziehr@q-perior.com

¹ Consultant, Topic Chapter Travel Transport and Logistics, Q_PERIOR AG

² Chair of Traffic Engineering and Control, Technical University of Munich (TUM), Germany

SHORT SUMMARY

Bus bunching describes a phenomenon that is familiar to many public transport users. Two buses, running according to a scheduled frequency, arrive at a stop in immediate succession. In most cases, the leading vehicle is delayed. The delay causes an increasing number of waiting passengers at the stops. Through this higher number of boarding and alighting passengers, the dwell time of the leading bus lengthens and by that also its delay. This problem is made visible using freely available public transport control data of two routes from Sydney, Australia. To validate the bunching events captured from the bus control data, General Transit Feed Specification (GTFS) data is used. The buses' positioning logs are traced to determine the distance between bunched vehicles. Additionally, a direct association between late departures of buses induced by delay propagation from one direction and increased bunching occurrence in the opposite direction is observed.

Keywords: Big Data, Bus Bunching, Public Transport

1. INTRODUCTION

Bus bunching resembles a phenomenon that is frequently observed in urban bus operations. Whereas measuring or detecting its characteristics has not led to a thorough definition of the underlying problem, most literature determines bus bunching by the immediate queuing of two consecutive buses at one particular stop (ILIOPOULOU et al., 2020). Bus bunching resembles a phenomenon that is frequently observed in urban bus operations. Whereas measuring or detecting its characteristics has not led to a thorough definition of the underlying problem, most literature determines bus bunching by the immediate queuing of consecutive buses at one particular stop (MOREIRA-MATIAS et al., 2014; XIN et al., 2021; YU et al., 2016).

This paper has three objectives. First, to develop a methodology for identifying and analysing the spatiotemporal dimensions of bus bunching using publicly available data. Second, to investigate whether strict adherence to the timetable, as well as knock-on delays, also lead to bunching. Third, an overarching objective of this research is to ensure that potential results are obtained as accurately as possible, but with a minimum of effort for the operator.

Datasets revealing the punctuality of buses are already available to many transport companies worldwide. The common procedure is to analyse the actual arrival and departure times at bus stops and to calculate the deviation from the schedule. This paper shows how additional findings can be drawn from such analysis and headway deviation calculations. The methodology process

proves how common practices considering solely schedule adherence are easily extended to capture bus bunching. Service regularity gains importance compared to schedule adherence, especially in dense headways. It is noteworthy that these insights do not require supplementary data sources. Automatic Vehicle Location (AVL) is not necessarily available to all operators, such data is used in this paper as a means of validation of the methodology.

2. METHODOLOGY

The approach used for this project involves a total of four steps (see Table 1). Firstly, with the research questions concerning bus bunching identification and its spatiotemporal analysis are outlined. Secondly, the data acquisition determines the required level of detail to analyse bus bunching. Thirdly, the evaluation framework is described, which is crucial for clarifying which key performance indicators (KPI) will subsequently allow the interpretation of the results of the data analysis. The goal is to create a schedule adherence index. The index does not measure bus bunching directly but bus regularity, which is closely associated with bus bunching. Fourthly, several proposed solution approaches are compared regarding both the determined KPIs and the available data basis. A presented validation method helps to underline the results. In an additional step, the position and time of the bunching events are compared against a second data stream that tracks the actual coordinates of the vehicles.

Table 1: Methodological Steps for Bus Bunching Identification

Step	Task	Question / Decision to be dealt with	Outcome
1	Data Acquisition	How can the available data basis be evaluated in terms of the project's feasibility?	Level of detail of data basis and its feasibility for the intended analysis
2	Evaluation Framework	Which KPIs can be measured with the data basis?	Choice of KPIs
3	Choice of Method	Which solution approach from literature appears to be suitable?	Choice of Solution Approach (Algorithm) to capture chosen KPIs
4	Data Analysis	How can the desired results be obtained from the available data?	Specification of measurement tools and techniques

Data acquisition

The data used for this work is acquired from the Bus Opal Assignment Model (BOAM) hosted by Transport for New South Wales in Australia (TfNSW), which makes a wide variety of public transport (PT) related datasets publicly available. For buses, only actual arrival times at stops are recorded so that the actual departure times remain unknown. Figure 1 shows exemplarily those four columns that are relevant for the bunching identification and its spatiotemporal analysis:

Table 2: Excerpt of BOAM Daily Dataset Relevant for Bunching Identification

Trip ID_Date	Stop	Scheduled Arrival	Actual Arrival
179839616_2020-02-03	1	06:00:00	06:03:34
179839727_2020-02-03	1	06:10:00	06:10:58
179839617_2020-02-03	1	06:20:00	06:22:34
179839618_2020-02-03	1	06:30:00	06:30:18
179839730_2020-02-03	1	06:38:00	06:37:19
179839732_2020-02-03	1	06:45:00	06:45:09

From public transport control data to bus bunching analysis

The real-time extension of the General Transit Feed Specification (GTFS) differentiates trip updates, service alerts, and vehicle positions. The vehicle positions feed is of vital importance for the identification of bus bunching, it depicts the current location and movement parameters of vehicles (BARBEAU., 2018). AVL data is the primary of three forms of PT control data (alongside Automatic Fare Collection and Automated Passenger Counting) and typically involves information in three dimensions (latitude, longitude, time). Consequently, the identification of bus bunching which relies solely on AVL data can be regarded as a robust methodology (SUN, 2020).

The spatiotemporal analysis is designed as a retrospective evaluation of sample data to uncover patterns. Particularly, categorizing bus operations into predefined levels of service relies on the coefficient of variation of headway deviations. To obtain the required quotient, the standard deviation of headways is divided by the average headway (CAMPS AND ROMEU, 2016).

The calculation of the coefficient of variation c_{vh} follows equation (1):

$$c_{vh} = \frac{sd(h_A)}{\underline{h}_A} \quad (1)$$

sd : standard deviation
 h_A : actual headway
 \underline{h}_A : average actual headway

The further translation into Levels of Service (LOS) is applied according to the threshold ranges from the PT Capacity and Quality of Service Manual presented in Table 3 (TURNER et al., 2010).

Table 3: Levels of Service of Schedule Adherence based on Headway Deviations

LOS	c_{vh}	$P(\text{abs}(h_i - h) > 0.5 * h)$	Passenger and Operator Perspective
A	0.00 - 0.21	$\leq 2\%$	Service provided like clockwork
B	0.22 – 0.30	$\leq 10\%$	Vehicles slightly off headway
C	0.31 – 0.39	$\leq 20\%$	Vehicles often off headway
D	0.40 – 0.52	$\leq 33\%$	Irregular headway, with some bunching
E	0.53 – 0.74	$\leq 50\%$	Frequent bunching
F	≥ 0.75	$> 50\%$	Most vehicles bunched

There are few studies in which schedule adherence is ascribed to minor importance. This is reasoned by the high utility of bus lines with short headways of less than ten minutes. Riders are assumed to travel spontaneously, meaning they do not check the upcoming departure times of their bus services (BARTHOLDI AND EISENSTEIN, 2012). However, a valuable contribution from the spatiotemporal analysis is to spot the locations at which bus bunching occurs regularly (LI et al., 2013).

Preliminary choice of method for data analysis

The prescribed methodology involves a suitable calculation method. Table 4 below describes the six steps carried out within that final methodological step. Because of the low traffic volumes during the night times, only the hour bands which are relevant to grasp the phenomenon are studied thoroughly. Data from weekdays in February are chosen as these do not interfere with public holidays or other strong seasonal influences.

Table 4: Six-step Heuristic as Final Methodological Step

Step	Description
1) Data cleaning	Elimination of faulty (e.g. double) or missing records.
2) Data sorting	Sort records by scheduled / actual arrival time for each stop for scheduled and actual headway calculation, respectively.
3) Headway calculation and bus identification	Headway can be easily obtained by subtracting two consecutive arrival times for each stop
4) Bus bunching identification	The set headway threshold for bus bunching identification is set to be $0.25 * h_{sched}$ (scheduled headway).
5) Bus bunching distribution and further KPI calculation	Count the number of identified bus bunching records for each stop in each hour for all days of the same type of day (weekday). Calculate the coefficient of variation for each stop in each hour and return the corresponding LOS.
6) Data aggregation and plotting	Aggregate data sets of same day type and plot results.

Supplementary analysis of delay propagation

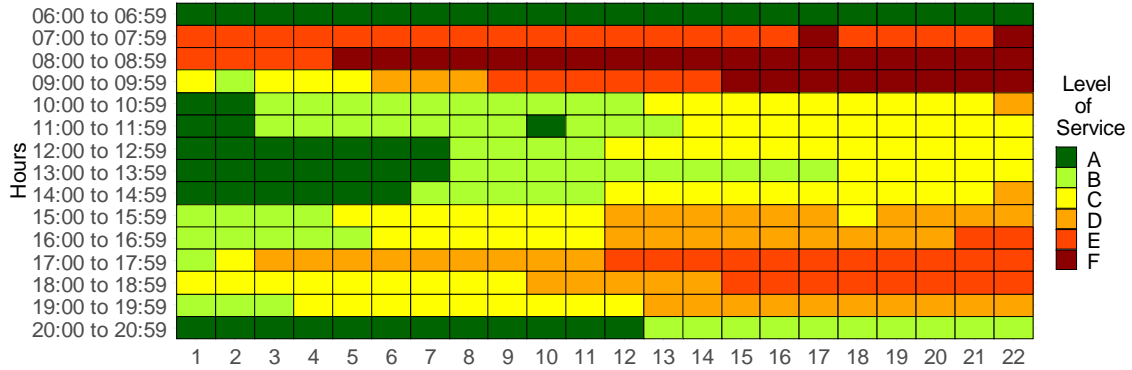
In classic scheduled services, buffer and turnaround times are scheduled at the terminals of scheduled routes. However, despite these preventive measures, it may happen that delayed buses from one direction do not re-enter the line route in the opposite direction on time. This delay propagation is closely linked to the phenomenon of bus bunching, as poor schedule adherence applies in both cases. The six-step heuristic for bus bunching analysis is perpetuated to conclude a direct association between late departures of buses induced by delay propagation from one direction and increased bunching occurrence in the opposite direction. By doing so, each trip is marked by a flag regarding its deviation from the scheduled departure at the start-stop. Thus, trips are divided into three categories - trips that depart more than a minute before their scheduled departure, trips that depart between a minute early and a minute late (one-minute tolerance), and trips that are more than a minute late. The latter serve as an indicator of the relationship between delay propagation and the occurrence of bus bunching events. Subsequently, for the three aforementioned categories of trips, bunching events are identified and additionally, the number of bunching events per trip is determined. The stop on the route at which the bunching event occurs has secondary importance – nevertheless, a trip can inherit more than one bunching event (PARK, 2020).

3. RESULTS AND DISCUSSION

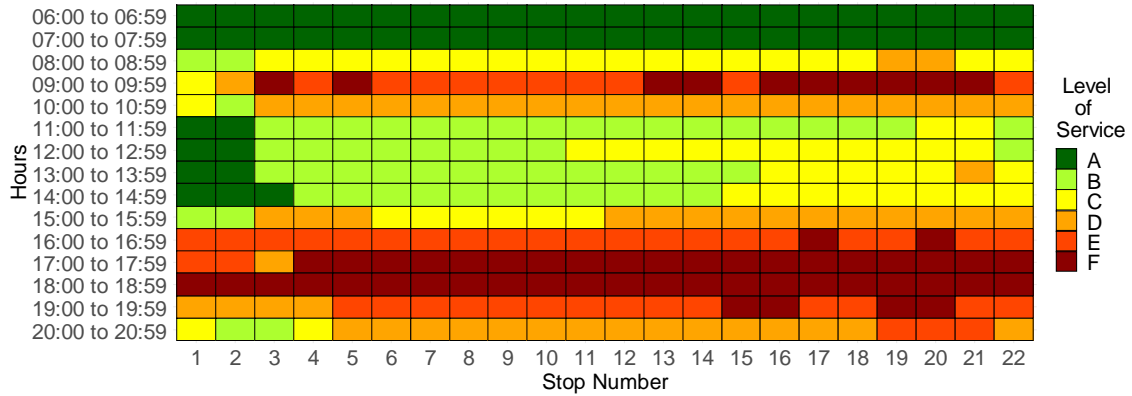
Low LOS resulting from high coefficients of variation of the calculated headway deviations indicate bus bunching. The following figure shows the LOS according to Table 3 for each stop of line 304 in both directions. For instance, the examined line 304 runs in the north-south direction in and out of Sydney's highly demanded central business district (CBD) on a ten-to-twelve-minute frequency during normal weekday hours. During peak hours, its headway is shortened to six minutes for the major commuting direction. On parts of the route close to the CBD, the headway is even lowered to three minutes. In contrast to the inbound results, Line 304 in the outbound direction reveals lower LOS in the afternoon.

Inbound

Calculation based on observations per Hourband aggregated from 20 weekday days
 (Indication of Bus Bunching using Levels of Service relying on Coefficients of Variation of Headway Deviations)



Outbound



Data acquired from: TfNSW's Bus Opal Assignment Model

Figure 1: Line 304 - Schedule Adherence Index in February 2020

During the week, most bunching events occur in the rush hour between 8 and 9 a.m. Weekends feature most bunching events around midday. According to the major commuting direction, the evening peak is significantly more affected by bunching than the morning, which is evident not only during the week but also on weekends as shown in Table 5 below.

Table 5: Line 304 – Number of Bus Bunching Events in February 2020

<u>Inbound</u>										
daytype	mon	tue	wed	thu	fri	sat	sun	weekdays	weekends	total
06:00 to 06:59	1	11	9	0	15	0	0	36	0	36
07:00 to 07:59	120	117	138	103	91	0	0	569	0	569
08:00 to 08:59	146	110	111	156	128	0	0	651	0	651
09:00 to 09:59	88	88	70	104	48	0	0	398	0	398
10:00 to 10:59	13	26	33	48	1	0	0	121	0	121
11:00 to 11:59	11	5	16	6	19	6	0	57	6	63
12:00 to 12:59	14	9	3	11	9	23	9	46	32	78
13:00 to 13:59	4	34	8	21	12	44	1	79	45	124
14:00 to 14:59	19	7	9	12	6	26	2	53	28	81
15:00 to 15:59	43	20	53	34	50	7	5	200	12	212
16:00 to 16:59	71	41	32	51	40	2	0	235	2	237
17:00 to 17:59	19	6	29	20	35	12	2	109	14	123
18:00 to 18:59	29	61	22	78	47	17	2	237	19	256
19:00 to 19:59	0	19	17	15	38	3	0	89	3	92
20:00 to 20:59	0	1	0	20	5	0	0	26	0	26
total	578	555	550	679	544	140	21	2906	161	3067
mean	145	139	138	170	136	28	5,25	145,3	17,89	105,76

<u>Outbound</u>										
daytype	mon	tue	wed	thu	fri	sat	sun	weekdays	weekends	total
06:00 to 06:59	0	0	0	0	0	0	0	0	0	0
07:00 to 07:59	1	0	2	0	0	0	0	3	0	3
08:00 to 08:59	14	7	9	32	1	0	0	63	0	63
09:00 to 09:59	55	39	38	55	24	0	0	211	0	211
10:00 to 10:59	49	49	49	83	21	0	0	251	0	251
11:00 to 11:59	10	11	0	23	0	1	0	44	1	45
12:00 to 12:59	7	14	3	4	20	34	0	48	34	82
13:00 to 13:59	18	18	7	29	1	22	0	73	22	95
14:00 to 14:59	7	23	7	42	11	54	1	90	55	145
15:00 to 15:59	53	23	6	28	50	22	0	160	22	182
16:00 to 16:59	67	57	110	53	139	20	0	426	20	446
17:00 to 17:59	106	97	135	136	108	1	0	582	1	583
18:00 to 18:59	99	133	160	110	121	47	0	623	47	670
19:00 to 19:59	17	50	65	103	180	24	0	415	24	439
20:00 to 20:59	0	46	9	55	92	0	0	202	0	202
total	503	567	600	753	768	225	1	3191	226	3417
mean	126	142	150	188	192	45	0,25	159,55	25,11	117,83

To conclude this section, it can be said that in both directions of line 304, the counted bunching events fit the heat maps of the schedule adherence index (Figure 1). Both the peak hours, as well as the major commuting direction, are apparent.

Validating identified bunching events using records from GTFS-real-time feed

Table 6: Identified Bunching Event for Exemplary Validation

Route	Trip	Stop	Sched- uled headway	Scheduled arrival time	Actual headway	Actual arrival time	Schedule deviation
304	21	8	540 s	18:14:00	807 s	18:14:04	-4 s
304	22	8	180 s	18:17:00	25 s	18:14:29	151 s

Table 6 depicts an identified bunching event as two consecutive arrivals of line 304 at stop 8 (trips 21 and 22) are recorded only 25 seconds after another. To validate whether a bus bunching event has occurred, the matching GTFS-real-time feed records at 18:14 needs to be considered (see Table 7). For the validation, the two trip IDs are to be checked for bunching at stop 8 (despite an insignificant five-second delay of the following vehicle's record).

Table 7: Matching GTFS records for Exemplary Validation

Route	Trip	Start time	Lati- tude	Longi- tude	Timestamp	Vehicle	Direction
304	21	17:59:00	-33.885	151.214	18:14:40	1339858	Inbound
304	22	18:02:00	-33.881	151.214	18:14:45	1340083	Inbound

Finally, a comparison of the latitude and longitude coordinates using the statistical software R computes the distance of the allegedly bunching buses. R yields a distance of only 515 meters, which unambiguously indicates bus bunching.

Dependency of bunching occurrence from delay propagation

Beyond the presented suitability of the methodology to study bus bunching and the spatiotemporal dimensions of the phenomenon, the relation between delay propagation on trips and bunching occurrences in the opposite direction is discovered. Table 8 notes bunching events according to one of three predefined categories concerning the start delay of the respective trips. Although the number of trips per category varies greatly, it appears that the dispersion of bunching events in the case of trips that suffer from delay propagation of more than one minute reaches higher value ranges.

Table 8: Bunching Events per Trip categorised by Schedule Adherence at Stop 1 (Line 304 Inbound –Weekdays in February 2020)

Schedule adherence at first stop (Category)	Number of trips	Total Bunching Events	Mean (Bunching events / trip)
Earlier than 1 min before schedule	24	74	3.08
Within 1 min deviation from the schedule	1239	1744	1.41
More than 1 min late	490	1002	2.04
Total	1753	2820	1.61

The 1239 trips recorded between one minute before and one minute after the schedule at the start-stop show a considerably lower mean of only 1.41 bunching events per trip (1744 bunching events

recorded). At the same time, indicating the negative effect of the delay propagation, 1002 bunching events are counted among 490 trips that are recorded for the category of more than one minute late at the start-stop. The corresponding average of 2.04 is clearly above the overall average (1.61) of all trips. The category of trips that start more than one minute early on the line route is very rare (only 24 occasions) and the high average value of 3.08 bunching events per trip most likely results from considerable irregularities in the operation of the vehicles.

4. CONCLUSIONS

Although various paradigms and algorithms have already sufficiently addressed the topic of spatiotemporal analysis of bus bunching, the selected measuring instruments allow a transparent view of this phenomenon. The uniqueness of the methodology is the type of data used. The data's prevalence as well as its scope and format are globally distinctive, which caters to a high transferability of the methodology. It shows that bus bunching can be analysed with publicly available PT control data. Typically, punctuality is the focus of analysis, but on-time performance is often not influenceable due to prevailing external factors. However, service regularity is a more promising indicator to assess the service quality of a line. PT agencies that record actual values of buses' arrivals or departures along the route can use the methodology presented here to better understand the occurrence of bus bunching in their network.

Following this work, the influence of short turns or buffer times on bus bunching events gives room for further investigation. These are the simplest tool for transport operators and can mitigate the proven delay propagation and associated bunching occurrence. Further influencing factors such as weather conditions, and temporal dimensions like the day of the week, time of day, and season could be additionally differentiated. Overall, bunching analysis and drawing the right conclusions from it could bridge the time until automated mitigation actions might be implemented in the onboard computers of buses.

REFERENCES

- Barbeau, S. (2018): Quality control-lessons learned from the deployment and evaluation of GTFS-realtime feeds. *97th Annual Meeting of the Transportation Research Board, Washington, DC*. 2018.
- Bartholdi, J. J.; Eisenstein, D. D. (2012): A self-coordinating bus route to resist bus bunching, *Transportation Research Part B: Methodological*. vol. 46, no. 4.
- Camps, J. M.; Romeu, M. E. (2016): Headway Adherence. Detection and Reduction of the Bus Bunching Effect. *European Transport Conference 2016*. Association for European Transport (AET).
- Iliopoulou, C.; Vlahogianni, E. I.; Kepaptsoglou, K. (2020): Understanding the factors that affect the bus bunching events' duration. *IEEE, 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1-6.
- Li, F.; Yang, D.; Ma, K. (2013): Bus Rapid Transit (BRT) Bunching Analysis with Massive GPS Data, *American Society of Civil Engineers. ICTE 2013: Safety, Speediness, Intelligence, Low-Carbon, Innovation*. 2013. 959-965.
- Moreira-Matias, L.; Gama, J.; Mendes-Moreira, J.; Freire de Sousa, J. (2014): An Incremental Probabilistic Model to Predict Bus Bunching in Real-Time *Advances in Intelligent Data Analysis*

XIII: 13th International Symposium, IDA 2014, Leuven, Belgium, October 30–November 1, 2014. Proceedings 13. Springer International Publishing, 2014.

Park, Y.; Mount, J.; Liu, L.; Xiao, N.; Miller, H. J. (2020): Assessing public transit performance using real-time data: spatiotemporal patterns of bus operation delays in Columbus, Ohio, USA, *International Journal of Geographical Information Science*.

Sun, W. (2020): *Bus Bunching Prediction and Transit Route Demand Estimation Using Automatic Vehicle Location Data*. Dissertation

Turner, S.; Robinson, B.; Koorey, G. (2010): Technical Note - Transportation Research Board (TRB) Annual Meeting (Conference) 2009 & 2010: Highlights, University of Canterbury. Civil and Natural Resources Engineering.

Xin, Q.; Fu, R.; Yu, S.; Ukkusuri, S.; Jiang, R. (2021): Modeling bus bunching and anti-bunching control accounting for signal control and passenger swapping behavior, University of South Florida Libraries, *Journal of Public Transportation*. vol. 23, no. 1.

Yu, H.; Chen, D.; Wu, Z.; Ma, X.; Wang, Y. (2016): Headway-based bus bunching prediction using transit smart card data, *Transportation Research Part C: Emerging Technologies*. vol. 72, pp. 45–59.

Explaining Walking in Cities – a Machine Learning Approach

Rasha Bowirrat*¹, Karel Martens², Yoram Shiftan³

¹ Architect and Urban Planner, Faculty of Architecture and Town Planning, Technion, Israel

² Professor, Faculty of Architecture and Town Planning, Technion, Israel

³ Professor, Faculty of Civil and Environmental Engineering, Technion, Israel

SHORT SUMMARY

A large body of research has developed on walking and walkability, in part in response to increasing concerns over people's health, climate change, livability, and social cohesion. Literature shows that some built environment and socio-demographic characteristics influence walking rates more than others.

Different approaches and methods have been used to study the relationship between the built environment characteristics, socio-demographic variables and walking patterns. Yet, so far very few studies have applied machine learning tools to study and explore these relationships. This research aims to start filling this void.

The study draws on a dataset contains details about trips made by over 37,000 respondents in the Tel-Aviv metropolitan area. The detailed data allow us to differentiate between walk-only trips and walk trips that are combined with other modes of transport. Our results show that the built environment shapes walk-only trips more than walking as an access or egress mode.

Keywords: Data analysis, Machine learning, Mobility, Urban planning, Walkability, Walking.

1. INTRODUCTION

Walking is a mode of transport that enables getting from one place to another, and it is the most prevalent form of physical activity. Walking is a fundamental constituent of nearly all trips, as it enables physical access to different kinds of facilities. Transportation means such as trains, buses, and private transport, require walking both for access and egress (Wigan, 1995).

Walking is not shaped solely by dedicated infrastructure (e.g., pavements and crossings) but is also highly dependent on other features of the built environment, as these can promote or constrain walking (Forsyth & Krizek, 2010; Lee & Moudon, 2004). The act of walking is shaped by the city, infrastructures, its built environment characteristics, and the sociodemographic variables of people.

Different approaches and tools are used to study walking behavior and investigate its relationship with personal and sociodemographic variables and built environment variables. Very few studies have applied data analysis and machine learning tools to study and explore the (non-linear) relationship between the different variables.

This study aims to disentangle the potential of the built environment effects on walking in urban areas and to determine the relative importance of built environment and socio-demographic variables in shaping walking patterns employing a machine learning and data analysis approach. This research is conducted among a diverse population in terms of their characteristics using a large data set that includes nearly 37,100 participants. The research question is: “What type of variables most strongly shape walking patterns?”.

2. METHODOLOGY

This research applies a Random Forest (RF) multiclass classification algorithm to identify the set of (walkability) parameters that most strongly shape walking in urban areas. The two compared groups of parameters include the sociodemographic and the built environment variables. RF algorithm can easily handle a large number of variables as it weighs the contribution of each variable according to how dependent it is on other variables (Breiman, 1996; T. Shi & Horvath, 2006).

We developed a model where we distinguish between four possible trip types that compile for the dependent variables in the model: walk only trips, walk trips in combination with public transport, walk trips in combination with car use, and trips that do not include walking at all (Table 1). We distinguish between these trip types, because we expect that effect of the built environment and of people’s socio-demographic characteristics may vary between the trips. We may expect that built environment factors may particularly shape walk only trips, while being less important for the other two trip types that include a walking leg. By distinguishing between the trip types, we can test this expectation, which has not yet been done in the literature.

We hypothesize that it is more likely that the built environment shapes walk-only trips, since walking as part of public transport or car trip is largely unavoidable. Yet, we also hypothesize that the built environment is more likely to shape the choice of walking integrated with public transport than the choice of walking and car use, as literature shows that the decision to use public transport is partly shaped by built environment factors (An et al., 2022). The dataset contains 288,555 trips. These trips are described by 40 different parameters including built environment variables (Table 2), and socio-demographic variables (Table 3).

Table 1. Dependent Variables; Key characteristics of the four Trips Types Distinguished in the Research

Class index	Class	Description	Percentage of trips of all walk trips	Mean length of travel distance	Mean travel time of trips
1	Walk-only trips	Trips that consist solely of one or more walk legs	24%	0.79 km	9.34 min
2	Walk + public transportation trip	Public transport means that were taken into consideration in this category are: bus, taxi, train, and organized shared transit Every trip with at least one walk leg and one PT leg, irrespective of whether the entire trip chain also includes other modes of transport for some trip legs (e.g. car, bicycle) are included 3. Trips that include public transport, but for which no walk trip was reported, were also added to this category, since we hypothesized that to get to a public transport stop a walk would be necessary in virtually all cases.	9%	8.09 km	35.78 min
3	Walk + car trip	Every trip with at least one walk leg and one leg by car (as driver or passenger) or motorbike, unless the trip chain includes PT for one or more trip legs (and irrespective whether the trip included yet other modes of transport for some trip legs (e.g., bicycle).	2%	7.74 km	32.96 min
4	Trips without a walk leg	All trips that do not belong to one of the categories mentioned above (bicycle trips are included).	65%	3.66 km	12.89 min

Table 2. Built Environment Variables in the Dataset & included in RF model

	Variable	Description	Share	Mean	SD (Standard deviation)	
Built Environment variables (Zonal level)	Residential density (HHdens)	Total number of households in zone divided by zone surface area (number/m ²)		3.8	4.0	
	Population density (popdens)	Total population divided by zone surface area (persons/m ²)		10.1	10.1	
	Land use mix in terms of jobs employment	Number of jobs in zone divided by population (jobs/persons)		0.4	0.1	
	Employees	Number of employees in area		1389.7	1146.1	
	Parking capacity	Parking capacity in area		3119.2	5883.1	
	Urban area type (share of zones that belong to each category)	Metropolitan CBD		3.36%		
		Urban Residential - Low-density: zones under 5,000 inhabitants per sq. km.		12.80%		
		Urban Residential - High-density: zones over 5,000 inhabitants per sq. km.		57.76%		
		Major public institutions: educational / legal / hospitals.		0.40%		
		Commercial: include city centers, shopping centers and markets.		8.48%		
Major Employment centers: employment centers over 4,000 employees.			4.72%			
Medium Employment centers: employment centers under 4,000 employees.		2.00%				

		Mixed Use Areas: areas with a mix of residential, commerce and employment.	3.60%		
		Major Transport facilities: airports, ports and bus stations.	0.24%		
		Sports and Tourism: areas with concentrations of hotels, beaches and sport facilities.	2.56%		
		Rural Areas: rural settlements (like moshavim and kibutzim), agricultural land and isolated developments.	3.12%		
		Open Areas: empty or non-built areas with no special use.	0.16%		
		Military Areas: zones used by the army.	0		
		Cemetery	0.48%		
		Small Settlements: isolated development of urban residential uses outside the urban core.	0		
	Students	Number of students studying in zone		110.79	1105.88
	Socio economic status	Socio-economic level of zone		11.30	4.76
	Parking availability at employment place (EmpPark)	EmpPark_1: Parking available for free for workers	22.2%		
		EmpPark_2: Parking available only near workplace	0.8%		
		EmpPark_3: Unavailable parking spaces	69.3%		

Table 3: Sociodemographic Variables in the Dataset & included in RF model

	Variable	Description	Share	Mean	SD (Standard deviation)
Socio-demographic variables	Age	Respondent's age		33.2	22.9
	Gender	Male respondents	51.5%		
		Female respondents	48.5%		
	Sector	Secular Jew	70.2%		
		Religious Jew	12.6%		
		Orthodox Jew	14.8%		
		Arab	2.3%		
	Education level	Highly educated respondent – undergraduate & graduate studies	26.0%		
		Medium educated respondent – high school certificate	24.8%		
		Low educated respondent – adult without high school certificate	19.0%		
		School students and other	29.9%		
	Employment status	Employed (full/part time job)	46.8%		
		Unemployed	23.3%		
		Other (unknown, irrelevant)	28.8%		
	Car license holding (Clic)	Clic_1: The respondent holds a driving license	62.9%		
		Clic_2: The respondent does not hold a driving license	13.7%		
		Clic_9: Unknown whether the respondent holds a driving license	0.003%		
		Clic_99: Irrelevant	23.3%		
	Household size (HHsize)	Person\household		2.9	1.76
	Children under age 8 in household	Per household		0.4	0.9

	Car ownership (HHVeh)	Households with at least one car in their ownership	76.7%			
	Bicycle ownership	Households with at least one bicycle in their ownership	31.4%			
	Household composition (HHType)	HHType_1: Households with one individual person		22.0%		
		HHType_2: Households for a couple		24.2%		
		HHType_3: Households for parents and their children		51.8%		
		HHType_4: Households for disabled person and assistant		0.1%		
HHType_5: Shared Households (between partners)		1.6%				

Techniques for dealing with the imbalanced dataset

As can be seen in Table 1, the data is imbalanced since the data set has skewed class proportions. The trip type with the least observation is the Walk+car class, as it accounts for only 2% of all trips. In contrast, Trips without any walk leg make up the majority class with 65% of the data. In this case, the RF algorithm will mainly relate to the majority class and treat the minority class features as noise in the data and ignore them. SMOTE technique was used in order to overcome this issue.

Hyperparameters Tuning

Tuning is the task of finding optimal hyperparameters for a RF model for a given dataset (Probst et al., 2018), thus optimizing the model in terms of its performance and running time. RF models works reasonably well with the default values of the hyperparameters specified in software packages. Nevertheless, tuning the hyperparameters can improve the performance of RF. The technique that was used in this research in order to overcome the imbalanced dataset was fitted to the hyperparameters that were accepted after 30 iterations (Table 4).

Oversampling

SMOTE (Synthetic Minority Oversampling Technique) was applied on the training set, alongside with hyper-tuning the model, which is an oversampling technique where synthetic samples are generated for the minority classes to rebalance the original training set.

After running this technique, an evaluation of the results should be done. Since the number of observations in each class is initially unequal, a so-called confusion matrix is needed to describe the performance of a classification algorithm.

In the confusion matrix, the number of correct and incorrect predictions is described with counted values for each class. Prediction for each class can be described by its precision and recall. Precision measures the share of data cases signaled by the model that are real predictions. Recall measures the share of data cases occurring in the domain that are “captured” by the models (Torgo & Ribeiro, 2009).

F1 score is the harmonic mean of the precision and recall of the model (Equation 1) which delivers the best value at 1 and worst score at 0 (Lipton et al., 2014).

Equation 1: F1 score

$$F1 = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$$

Examining the first model's performance (4 classes model) based on the F1-score shows a relatively high prediction accuracy score for all of the classes apart from the walk + car class. The low F1 score for the walk + car trip class indicates that the model did not predict the true labels in this class, as it predicted correctly only 6% of the labels. The majority of the wrong predictions was in favor to trips without a walk leg. This suggests that the model mistakenly confuses between these two classes, thus resulting in inaccurate feature importance of each one of the parameters in the model regarding the walk + car class (Figure 1.a).

In order to overcome this issue, we developed a second model (3 classes model) where we excluded the trips without a walk leg class from the model. In this case, the F1 score for the walk + car trip class was substantially higher than in the first model (62%) (Table 4).

Table 4. Hyperparameters setting and Models classification reports

Hyperparameters	Parameter / Model		SMOTE Model 1: 4 classes model	SMOTE Model 2: 3 classes model
	Parameter	Description (58)		
	n_estimators	Decision trees number being built in the forest	1000	2000
	min_sample_split	Minimum number of samples required to split an internal node	10	2
	min_sample_leaf	Minimum number of data required in node	1	2
	max_features	Maximum features number used for a node split process	“auto”	“auto”
	max_depth	Maximum depth and levels a decision tree is allowed	80	60
	bootstrap	False value: All data is used for every decision tree; else selected bootstrap samples are used when building decision trees	False	False
F1 score for different Classes	1. Walk only trips		0.78	0.91
	2. Walk trip + PT		0.61	0.75
	3. Walk trip + Car		0.08	0.62
	4. Trips without a walk leg		0.90	-

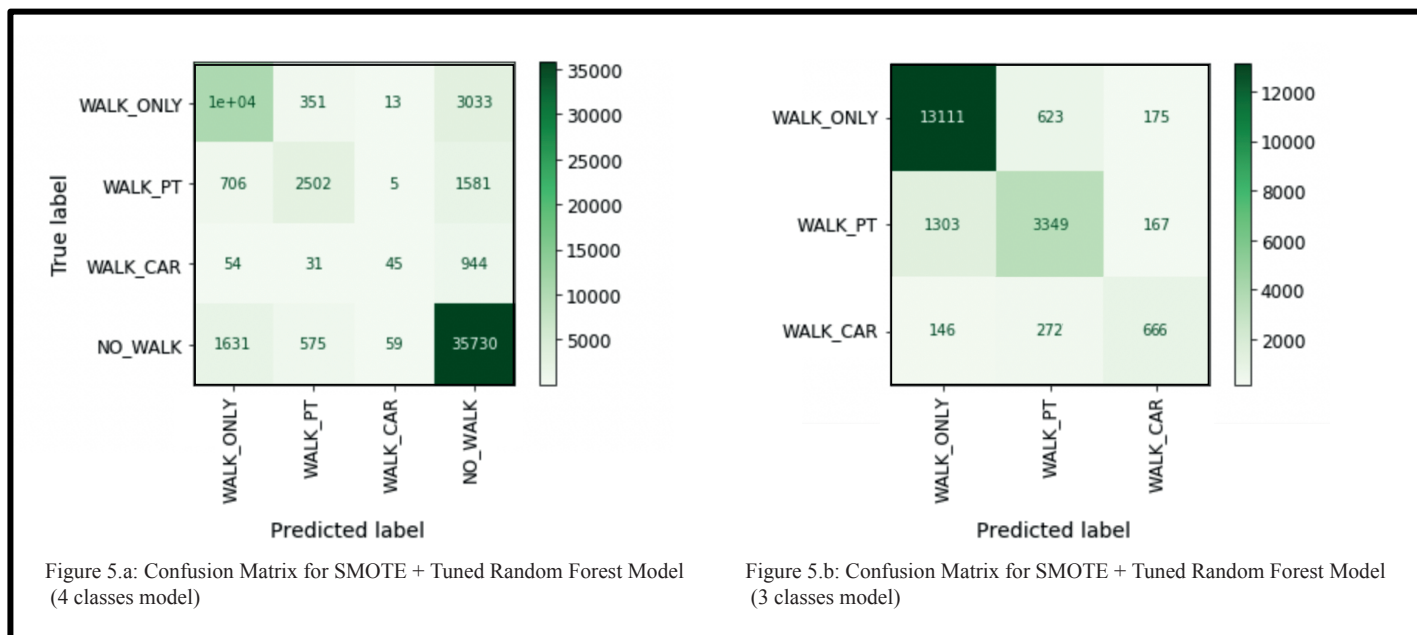


Figure 1. Confusion Matrix for (a) 4 Classes Model, and (b) 3 Classes Model

3. RESULTS AND DISCUSSION

The results of our analysis concern the four trip types that were investigated in this research and depicted in figure 2. For walk-only trips, the likelihood that a person makes a walk-only trip increases with population density at origin, population density at the destination, household density at origin, household density at the destination, no parking available at workplaces, household size, and workers number in origin. In contrast, it decreases with car license holding, household vehicle, and age that have the strongest negative influence.

Walk + PT trips are related to several features. Vehicle license and number of vehicles in the household all shape the number of the walk + PT trips. Additional sociodemographic characteristics that are related to making walk + PT trips are age, gender, sector, employment status, and education level. Males and secular Jews tend to make less walk + PT trips. In terms of the built environment variables, free parking at the workplace leads to a reduction in walk + PT trips. In contrast, the unavailability of parking near the workplace has a positive impact on walk + PT trips. Furthermore, population density at the trip origin increases the choice for walk + PT trips.

For walk + car trips, the likelihood of making this trip type is influenced by holding a car license, age, and the number of cars in the household, among other factors. These three variables have strong importance for a person choosing this trip type. An additional factor is the household size. Part of the most important built environment factors that reduce the likelihood that a person will make walk + car trips is population density at origin and destination, workspace parking availability, and the number of workers at the destination area.

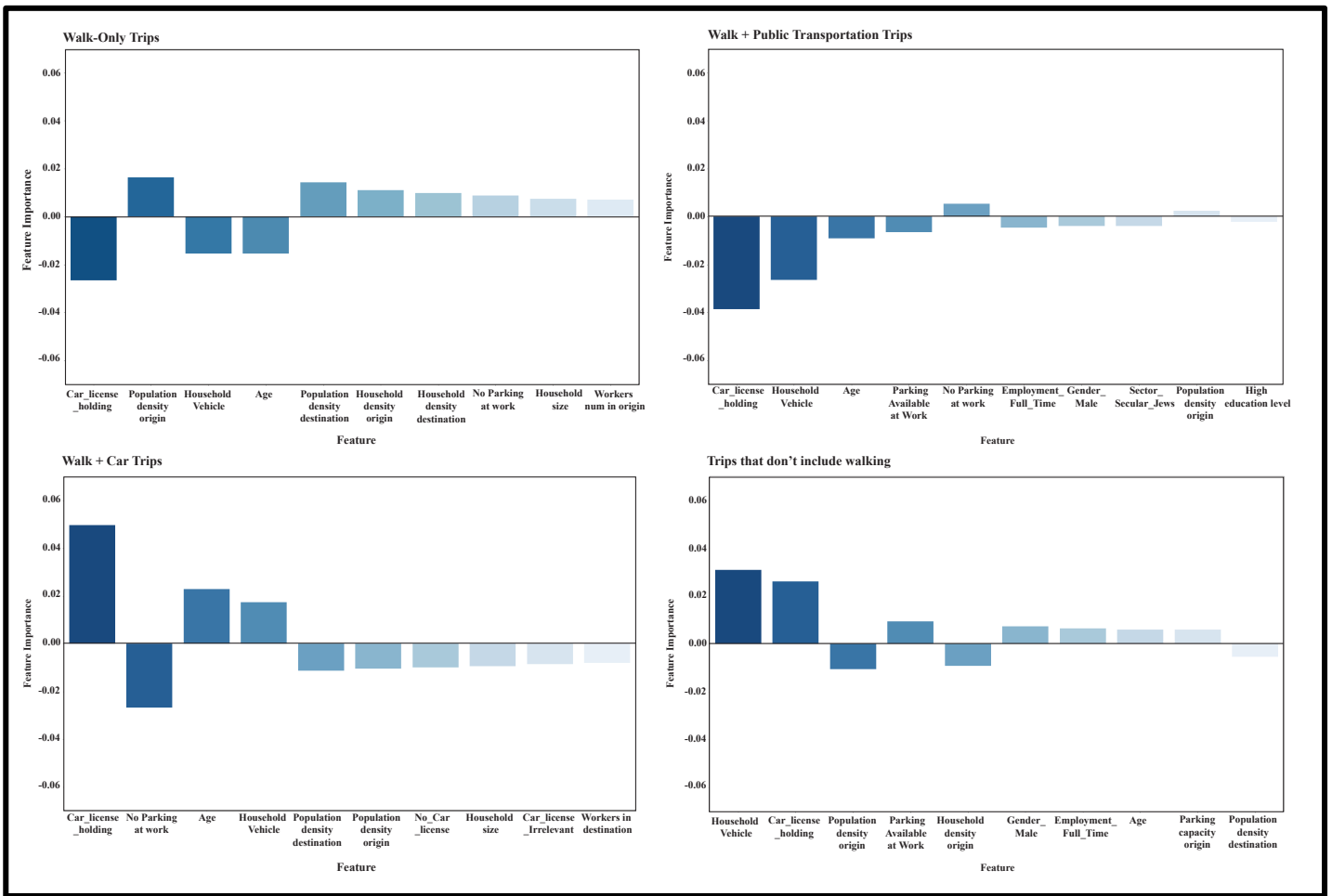


Figure 2. Model results: Feature Importance for Different Trips Types

4. CONCLUSIONS

Our results illustrate the importance of built environment variables in shaping walking-only trips more than other types of trips, with household and population densities having the strongest positive effect on walk-only trips of all built environment characteristics.

Our study shows that machine learning and data science approaches hold promise for the analysis of walking patterns, and for gaining insight into the set of variables that influence walking in cities.

The study suggests that the impacts of minimum parking norms for offices and other employment types have an impact beyond the home-to-work trip. Our findings indicate that parking at the workplace also affects the frequency with people engage in walk-only trips. This underscores the importance of abolishing parking minimums, as is gradually occurring in Israel and elsewhere (Christiansen et al., 2017; Shifan & Burd-Eden, 2001; SimiAeviA et al., 2013).

ACKNOWLEDGEMENTS

This research was supported by the Azrieli Foundation, and the Israeli Smart Transportation Research Center.

REFERENCES

- An, R., Wu, Z., Tong, Z., Qin, S., Zhu, Y., & Liu, Y. (2022). How the built environment promotes public transportation in Wuhan: A multiscale geographically weighted regression analysis. *Travel Behaviour and Society*, 29, 186-199.
- Breiman, L. (1996). Bagging predictors [Article]. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/BF00058655>
- Christiansen, P., Engebretsen, Ø., Fearnley, N., & Usterud Hanssen, J. (2017). Parking facilities and the built environment: Impacts on travel behaviour [Article]. *Transportation Research. Part A, Policy and Practice*, 95, 198–206. <https://doi.org/10.1016/j.tra.2016.10.025>
- Forsyth, A., & Krizek, K. J. (2010). Promoting Walking and Bicycling: Assessing the Evidence to Assist Planners [Article]. *Built Environment (London. 1978)*, 36(4), 429–446. <https://doi.org/10.2148/benv.36.4.429>
- Lee, C., & Moudon, A. V. (2004). Physical Activity and Environment Research in the Health Field: Implications for Urban and Transportation Planning Practice and Research [Article]. *Journal of Planning Literature*, 19(2), 147–181. <https://doi.org/10.1177/0885412204267680>
- Lipton, Z. C., Elkan, C., & Narayanaswamy, B. (2014). *Thresholding Classifiers to Maximize F1 Score*. <http://arxiv.org/abs/1402.1892>
- Probst, P., Wright, M., & Boulesteix, A.-L. (2018). *Hyperparameters and Tuning Strategies for Random Forest*. <https://doi.org/10.1002/widm.1301>
- Shi, Q., & Abdel-Aty, M. (2015). Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways [Article]. *Transportation Research. Part C, Emerging Technologies*, 58, 380–394. <https://doi.org/10.1016/j.trc.2015.02.022>
- Shiftan, Y., & Burd-Eden, R. (2001). Modeling Response to Parking Policy [Article]. *Transportation Research Record*, 1765(1), 27–34. <https://doi.org/10.3141/1765-05>
- SimiAeviA, J., VukanoviA, S., & MilosavljeviA, N. (2013). The effect of parking charges and time limit to car usage and parking behaviour [Article]. *Transport Policy*, 30, 125–131. <https://doi.org/10.1016/j.tranpol.2013.09.007>
- Torgo, L., & Ribeiro, R. (2009). Precision and recall for regression. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5808 LNAI, 332–346. https://doi.org/10.1007/978-3-642-04747-3_26

Wigan, M. (n.d.). Transportation Research Record No. 1487, Nonmotorized Transportation Research, Issues, and Use. In *TRANSPORTATION RESEARCH RECORD*.

Optimal bicycle network expansions with endogenous demand

Mads Paulsen*¹ and Jeppe Rich²

¹Postdoc, Department of Technology, Management and Economics, Technical University of Denmark, Denmark

²Professor, Department of Technology, Management and Economics, Technical University of Denmark, Denmark

SHORT SUMMARY

The challenge of identifying the ideal spatial and temporal prioritization for long-term expansions of bicycle networks is a complex undertaking. Our objective in this research is to determine the most beneficial expansions of bicycle networks for society, while considering the impact of level-of-service effects and induced demand throughout the evaluation period. While the effects of constant demand can be approximated through a sequence of linear binary mathematical programs (Paulsen & Rich, 2023), accommodating induced demand necessitates a different optimization approach that accounts for the likelihood of various segments being integrated in the infrastructure in future years during the optimization process. We put this approach to the test by applying it to the Greater Copenhagen Cycle Superhighway network. It is demonstrated that the optimized infrastructure render benefit-cost ratios exceeding 10, and that accounting for demand effects, significantly increases the societal return and changes the geographical structure of optimal investments.

Keywords: Bicycle network design; Bicycle traffic; Induced demand; Socioeconomic assessment; Dynamic optimization

1 INTRODUCTION

There is considerable evidence that bicycle demand is impacted by the presence of bicycle infrastructure, as demonstrated in several studies including van Goeverden et al. (2015) and Rich et al. (2021). The implementation of bicycle infrastructure not only affects travel time benefits resulting from route choice substitution as demonstrated in Paulsen & Rich (2023), but also the number of bicycle trips in the network (Hallberg et al., 2021). This is a result of mode substitution effects and potentially induced traffic. The societal value of increasing bicycle mileage is evidenced in Breda et al. (2018) and Martin et al. (2006), who study the external health benefit of one kilometer of cycling. The cost-benefit performance of bicycle infrastructure is studied in Rich et al. (2021), who finds that bicycle infrastructure is highly beneficial.

Optimal design of bicycle networks has been studied within operation research since early work by Smith & Haghani (2012) and Mesbah et al. (2012). The objective functions and constraints vary largely across studies, from approaches who i) minimize investment cost constrained by a minimum level-of service (Duthie & Unnikrishnan, 2014), ii) minimize local detours (Lim et al., 2021), iii) maximize cyclists on links where the stress-level is low (Chan et al., 2022; Ospina et al., 2022), iv) minimize generalized costs (Mauttone et al., 2017; Liu et al., 2019), or v) consider multi-objective costs (Lin & Yu, 2013; Lin & Liao, 2016; Liaw & Lin, 2022) subject to budget constraints. Although the studies have considered a large variety of performance measures, none of the studies calculate societal cost-benefit performance of the resulting bicycle network investment plans. Furthermore, many of these studies embed the route choice of cyclists directly in the optimization model, which becomes computationally intractable when considering large-scale applications with very large networks and many origin-destination pairs.

A recent study (Paulsen & Rich, 2023) shows that the consumer surplus of existing users can be approximated closely through sequences of linear binary mathematical programs at the level of the links. The approach is based on a method where OD-benefits are assigned to the network. The study show – under the assumption of constant demand – how optimal infrastructure expansions can be derived from a series of binary linear programs. However, as the study assumes the demand to be exogenous, it is not able to include health benefits that arise from increased bicycle demand. As demonstrated in previous research (Breda et al., 2018; Rich et al., 2021), this is the single

most important factor when calculating the societal net present value of large bicycle network expansions.

This study extends [Paulsen & Rich \(2023\)](#), and the literature by and large, by determining the societal optimal expansions with endogenous demand. Methodologically, the approach is based on a dynamic optimization framework, within which the expected level-of-service and induced demand are approximated forward in time, to then approximate the expected accumulated benefit of selecting a given segment at a given time. By applying the algorithm to the entire evaluation period, we identify a bicycle infrastructure plan that render a solution worth 16 billion DKK in net present value terms. This correspond to a solution that is 81%–417% better than the considered reference strategies and with a benefit-cost ratio exceeding 10.

2 METHODOLOGY

Approximation of net present value

The overall aim is to provide a reasonable and computationally feasible approximation to the net present value at time t of any given existing network configuration \mathbf{u}^{t-1} and investment action $\Delta\mathbf{u}^t$. Here, \mathbf{u}^{t-1} is a binary vector, which has value 1 for *link segments* that were constructed at or before $t-1$, and $\Delta\mathbf{u}^t$ being a binary vector, which is 1 for the segments being constructed exactly at time t . Link segments are natural bundles of links (chunks of routes) along the same corridor, see [Figure 3](#) for an example.

In [Paulsen & Rich \(2023\)](#) it is shown that the travel time function X_ω can be approximated very accurately for any configuration of \mathbf{u}^{t-1} by assigning the OD-level travel time savings back onto the network, and by taking into account the travel time savings. That is,

$$X_\omega(\mathbf{u}^{t-1}) \simeq x_\omega^0 - (\Delta\mathbf{x}_\omega)^\top \mathbf{u}^{t-1}, \quad (1)$$

$\omega \in \Omega$ represents combinations of OD and traveler type, and x_ω^0 is the baseline travel time for ω without any network upgrades. $\Delta\mathbf{x}_\omega = [\Delta x_{1,\omega} \quad \Delta x_{2,\omega} \quad \cdots \quad \Delta x_{|\mathcal{B}|,\omega}]^\top$ is the approximated vector of linear travel reductions from [Paulsen & Rich \(2023\)](#),

$$\Delta\mathbf{x}_{b,\omega} = \frac{\sum_{l \in b} L_l}{\sum_{l \in q_\omega^{\mathcal{L}} \cap \mathcal{L}} L_l} \left(\sum_{l \in q_\omega^0} \tau_{l,\omega} - \sum_{l \in q_\omega^{\mathcal{L}}} \hat{\tau}_{l,\omega} \right), \quad b \in \mathcal{B}, \omega \in \Omega. \quad (2)$$

Here, L_l is the length of link l , $\tau_{l,\omega}$ is the non-upgraded travel time on link l for the traveler type associated with ω , and $\hat{\tau}_{l,\omega}$ is the corresponding upgraded travel time. $q_{\omega,0}$ and $q_{\omega,\mathcal{L}}$ denotes the shortest paths for OD and traveler type ω in the non-upgraded network and the fully upgraded network, respectively.

Equality is guaranteed in [Eq. \(1\)](#) at the two extrema $\mathbf{u}^{t-1} = \mathbf{0}$ and $\mathbf{u}^{t-1} = \mathbf{1}$. For this study, analogously, we introduce the approximated vector of linear travel distance extension $\Delta\lambda_\omega = [\Delta\lambda_{1,\omega} \quad \Delta\lambda_{2,\omega} \quad \cdots \quad \Delta\lambda_{|\mathcal{B}|,\omega}]^\top$, which allows approximating the traveled distance for each ω ,

$$\Lambda_\omega(\mathbf{u}^{t-1}) \simeq \lambda_\omega^0 + (\Delta\lambda_\omega)^\top \mathbf{u}^{t-1}. \quad (3)$$

Again with guarantee for equality for $\mathbf{u}^{t-1} \in \{\mathbf{0}, \mathbf{1}\}$. Here $\Delta\lambda_\omega$ has elements,

$$\Delta\lambda_{b,\omega} = \frac{\sum_{l \in b} L_l}{\sum_{l \in q_\omega^{\mathcal{L}} \cap \mathcal{L}} L_l} \left(\sum_{l \in q_\omega^{\mathcal{L}}} L_l - \sum_{l \in q_\omega^0} L_l \right), \quad b \in \mathcal{B}, \omega \in \Omega. \quad (4)$$

Finally, the demand function D_ω does not require searching through the network, why it can be evaluated sufficiently quickly to be used as it is. In the study we use a simple logit-based mode choice model based on parameters from [Hallberg et al. \(2021\)](#). More advanced models can easily be considered within the framework we propose, for instance models that include destination choice or forecasts demand according the development in GDP.

All in all, this suggest that the net present value can be approximated reasonably and computationally efficient by $\text{NPV}^t(\Delta\mathbf{u}^t; \mathbf{u}^{t-1})$ in the equation below,

$$\begin{aligned}
\text{NPV}^t(\Delta \mathbf{u}^t; \mathbf{u}^{t-1}) = & \overbrace{\kappa^t \sum_{\omega \in \Omega} \zeta_{\omega} \frac{D_{\omega}(\mathbf{u}^{t-1}) + d_{\omega}^0}{2} (x_{\omega}^0 - X_{\omega}(\mathbf{u}^{t-1}))}^{\text{Consumer surplus}} + \\
& \overbrace{\kappa^t \sum_{\omega \in \Omega} \xi_{\omega} (D_{\omega}(\mathbf{u}^{t-1}) \Lambda_{\omega}(\mathbf{u}^{t-1}) - d_{\omega}^0 \lambda_{\omega}^0)}^{\text{Health benefits}} + \\
& \overbrace{\kappa^{|\mathcal{T}|} \sum_{b \in \mathcal{B}} c_b \Delta u_b^t}^{\text{Scrap value}} - \overbrace{\kappa^t \sum_{b \in \mathcal{B}} c_b \Delta u_b^t}^{\text{Construction costs}} - \overbrace{\kappa^t \sum_{b \in \mathcal{B}} m_b u_b^{t-1}}^{\text{Maintenance costs}}. \tag{5}
\end{aligned}$$

The used notation is summarized in Table 1.

$b \in \mathcal{B}$	A link segment b (containing links $l \in b$ along the same corridor), within the set of all link segments \mathcal{B} . \mathcal{B} partitions \mathcal{L} , such that the each link of \mathcal{L} belongs to exactly one link segment $b \in \mathcal{B}$.
c_b	Construction cost of segment b . From Incentive (2018) .
d_{ω}^0	Baseline demand for ω , i.e. $D_{\omega}(X(\mathbf{0}))$.
D_{ω}	Demand function for ω . Parameters adopted from Hallberg et al. (2021) .
ζ_{ω}	Value of time for the traveler type of ω . Value of 91 DKK per hour (Technical University of Denmark, 2022).
κ^t	Discounting factor for time t . From Technical University of Denmark (2022) .
λ_{ω}^0	Baseline travel distance of ω , i.e. $\Lambda_{\omega}(\mathbf{0})$.
Λ_{ω}	Travel distance function for ω .
m_b	Annual maintenance cost of segment b . From Incentive (2018) .
$\Delta \mathbf{u}^t$	The decision variable vector at time t , $[\Delta u_1^t \ \Delta u_2^t \ \dots \ \Delta u_{ \mathcal{B} }^t]^{\top}$, which is 1 for segments being chosen at time t , and 0 otherwise.
\mathbf{u}^{t-1}	The vector $[u_1^{t-1} \ u_2^{t-1} \ \dots \ u_{ \mathcal{B} }^{t-1}]^{\top}$ containing ones for all segments that have been selected at time $t-1$ or before, and zeroes elsewhere. That is, $\mathbf{u}^{t-1} = \sum_{k \leq t-1} \Delta \mathbf{u}^k$.
x_{ω}^0	Baseline travel time for ω , i.e. $X_{\omega}(\mathbf{0})$
X_{ω}	Travel time function for ω . From Hallberg et al. (2021) .
ξ_{ω}	Health benefit factor per km for ω (subtracted the corresponding accident factor). Value of 7.11 DKK per km (Technical University of Denmark, 2022).
$\omega \in \Omega$	Considered OD-pair and traveler type combinations. From Hallberg et al. (2021) .

Table 1: Notation overview for the net present value calculation (Eq. (5))

Optimization framework

The idea is then to embed this expression into an optimization scheme that optimizes $\Delta \mathbf{u}^t$ for all t in the 50 year evaluation period \mathcal{T} , as stated in Problem 1.

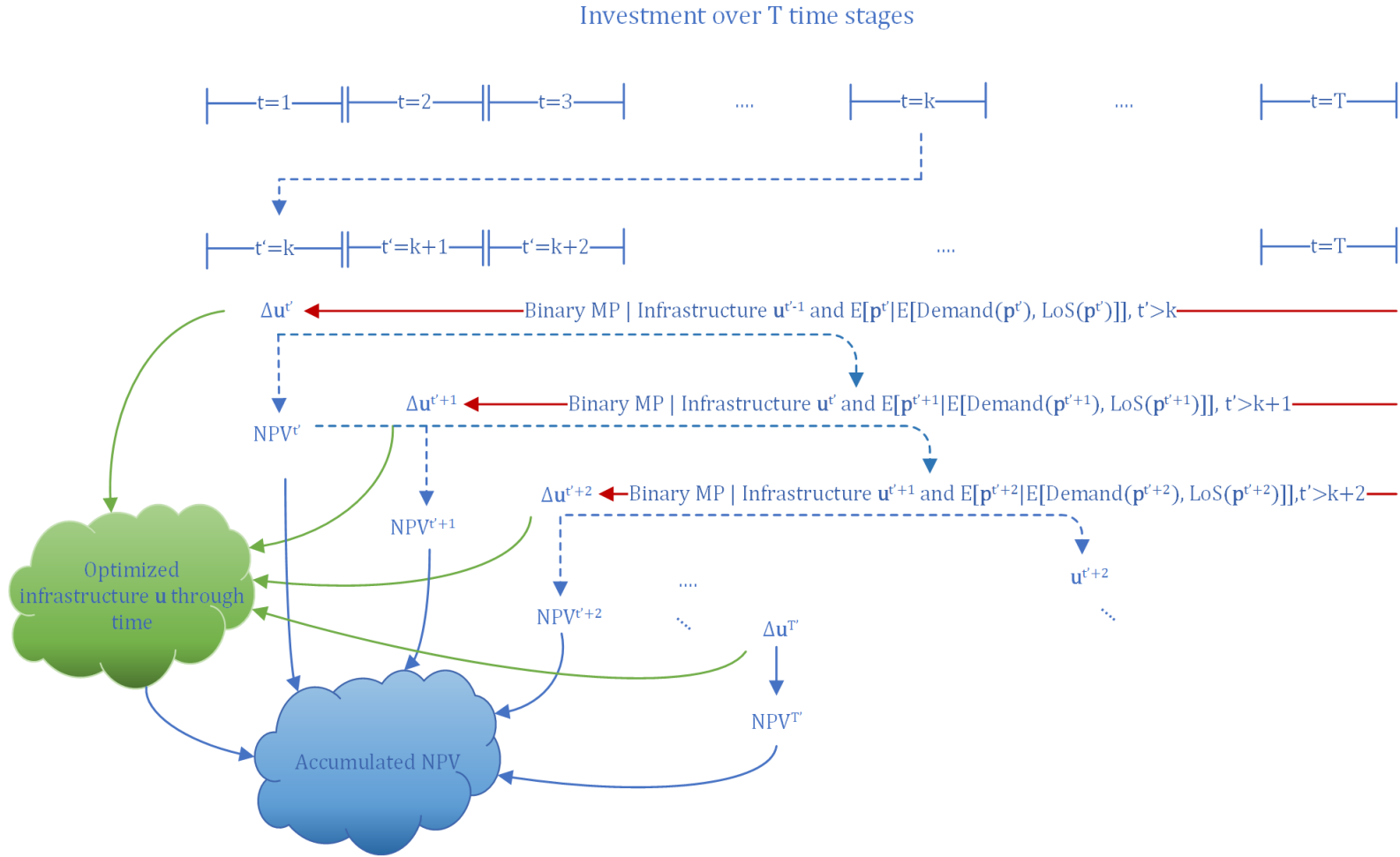


Figure 1: Flow chart of the proposed optimization scheme. $\Delta u^{t'}$ are the optimal strategies at each time step, whereas $\mathbf{p}^{t'}$ represents the expected infrastructure composition at future time steps. The binary mathematical program (Binary MP) is Problem 2. $E[\mathbf{p}^{t'} | E[Demand(\mathbf{p}^{t'}), LoS(\mathbf{p}^{t'})]] = E[\mathbf{p}^{t'} | E[D(\mathbf{p}^{t'}), X(\mathbf{p}^{t'})]]$ forms a fixed point problem across all future $t' > k$, which leads to the vector \mathbf{S}^t . The calculation of \mathbf{S}^t is further detailed in Figure 2, and constitutes the coefficients for the linear objective function of Binary MP used to determine $\Delta u^{t'}$.

$$\max_{\Delta \mathbf{u}^t} Z = \sum_{t \in \mathcal{T}} \text{NPV}^t(\Delta \mathbf{u}^t, \mathbf{u}^{t-1}) \quad s.t. \quad (\text{P1a})$$

$$B^t \geq \sum_{k \in \mathcal{T}: k \leq t} \kappa^k \left(\mathbf{c}^\top \Delta \mathbf{u}^k + \mathbf{m}^\top \mathbf{u}^{k-1} \right), \quad \forall t \in \mathcal{T} \quad (\text{P1b})$$

$$u_b^t \geq u_b^{t-1}, \quad \forall b \in \mathcal{B}, \forall t \in \mathcal{T} \quad (\text{P1c})$$

$$\Delta u_b^t \in \{0, 1\}, \quad \forall b \in \mathcal{B}, \forall t \in \mathcal{T} \quad (\text{P1d})$$

Problem 1: NPV-model with flexible demand

Here, B^t is the cumulative budget for time t , whereas \mathbf{c} and \mathbf{m} are construction costs and maintenance costs vectors defined by $\mathbf{c} = [c_1 \ c_2 \ \dots \ c_{\mathcal{B}}]^\top$ and $\mathbf{m} = [m_1 \ m_2 \ \dots \ m_{\mathcal{B}}]^\top$, respectively. A sketch of the overall idea is outlined in Figure 1.

So far, the approach is more or less similar to that of Paulsen & Rich (2023). That is, given previous decisions \mathbf{u}^{t-1} , we determine the optimal composition of the segments to select at time t ($\Delta \mathbf{u}^t$) subject to budget constraints, such that the expected future net present value is maximized. Once the optimal $\Delta \mathbf{u}^t$ has been found, \mathbf{u}^t can be updated to $\mathbf{u}^t \leftarrow \mathbf{u}^{t-1} + \Delta \mathbf{u}^t$, and we can consider the next choice situation at time $t \leftarrow t + 1$. The basic idea is that at any given decision time t , we can assume that we already know what has happened in the past, i.e. $\Delta \mathbf{u}^{t'}, \forall t' < t$.

However, the optimization problem is complicated by the presence of endogenous demand. A very precise approximation could be made concerning future net present values without taking into account the expectations of future investment in Paulsen & Rich (2023) under the assumption of constant demand. In that case, the effect of each segment could be linearized, fully ignoring their interaction without any notable loss in net present value precision. This would clearly be inappropriate when taking endogenous demand into account.

Thus, instead we develop a vector \mathbf{S}^t that also takes the expectations of future investments $\mathbf{p}^{t'}$ into account for $t' > t$. It gives an approximation of how each, so far un-selected segment, contributes to the expected accumulated net present value. We use this as the coefficients for our objective function in Problem 2.

$$\max_{\Delta \mathbf{u}^t} Z = (\mathbf{S}^t)^\top \Delta \mathbf{u}^t \quad s.t. \quad (\text{P2a})$$

$$B^t \geq \sum_{k \in \mathcal{T}: k \leq t} \kappa^k \left(\mathbf{c}^\top \Delta \mathbf{u}^k + \mathbf{m}^\top \mathbf{u}^{k-1} \right), \quad \forall t \in \mathcal{T} \quad (\text{P2b})$$

$$u_b^t \geq u_b^{t-1}, \quad \forall b \in \mathcal{B}, \forall t \in \mathcal{T} \quad (\text{P2c})$$

$$\Delta u_b^t \in \{0, 1\}, \quad \forall b \in \mathcal{B}, \forall t \in \mathcal{T} \quad (\text{P2d})$$

Problem 2: The individual binary linear problems

Calculation of \mathbf{S}^t

The calculation \mathbf{S}^t is a tedious and complex task not particularly suited for being explained in detail in an extended abstract. Still, we aim at outlining the key aspects in this section. A flow chart of the process of calculating \mathbf{S}^t for a single $t \in \mathcal{T}$ is found in Figure 2.

The calculation of \mathbf{S}^t considers future time stages $t' \in \mathcal{T} : t' > t$, in which we do not yet know which segments will be chosen. We accommodate this by loosening the restriction of binary decisions, and instead introduce a cumulative probability vector $\mathbf{p}^{t'}$ for future $t' > t$ for all segments that remain unchosen at time t , denoted by $\mathcal{B}^t = \{b \in \mathcal{B} : u_b^{t-1} = 0\}$. Likewise, we introduce the instantaneous probability vector $\Delta \mathbf{p}^{t'} = \mathbf{p}^{t'} - \mathbf{p}^{t'-1}$. Initially, we do not differentiate between the probability of various segments, i.e. assign uniform probabilities $p_1^{t'} = p_2^{t'} \dots = p_{|\mathcal{B}^t|}^{t'}, \forall t' > t$ across segments, but we will later set up a fixed point problem where the probabilities feed into an expected future net present value for each $b \in \mathcal{B}^t$, and the future net present values affect the probabilities.

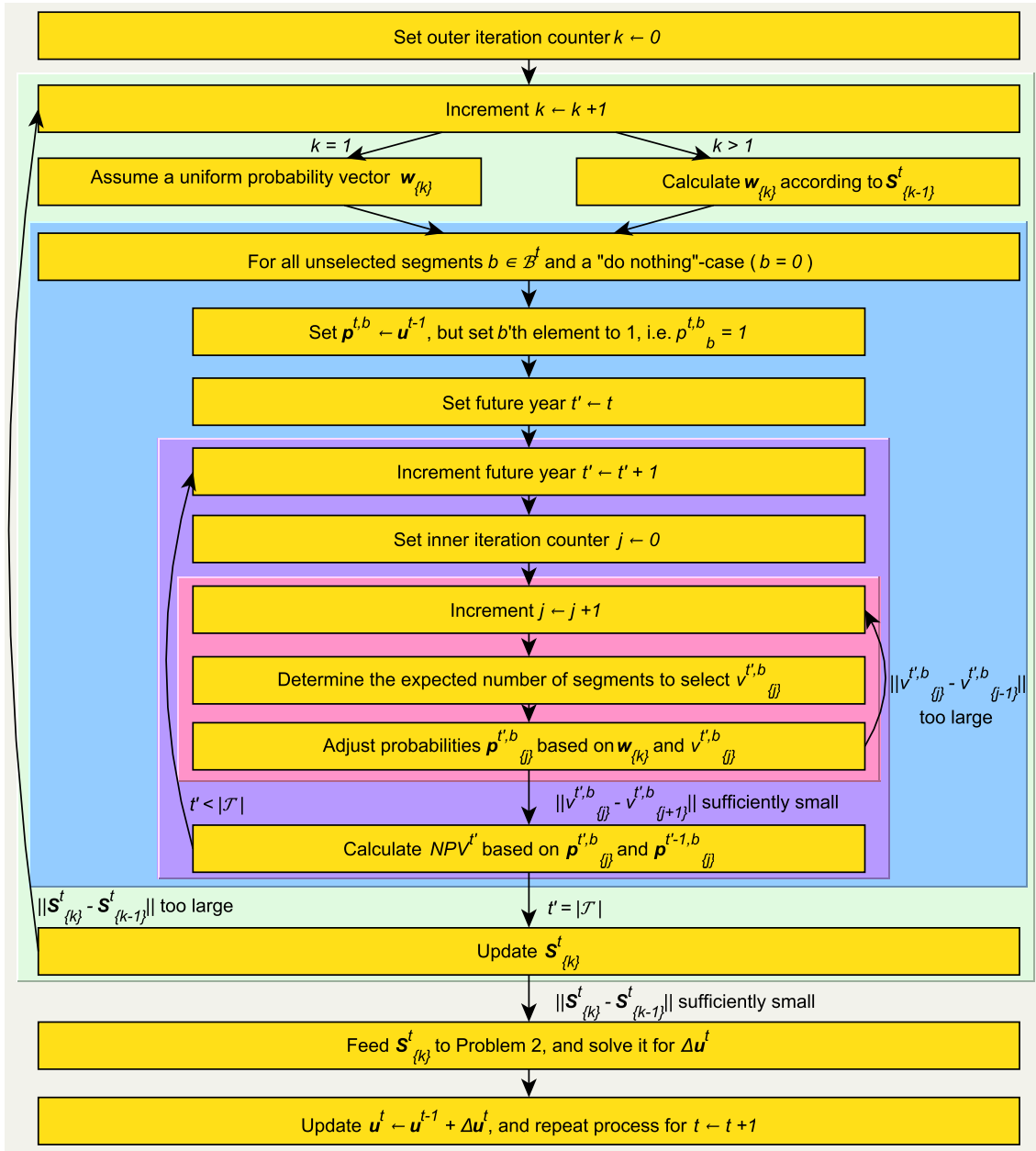


Figure 2: Flow chart of the process of calculating \mathbf{S}^t for a single $t \in \mathcal{T}$. The calculation contains two fixed point problems. One for determining $\nu^{t',b'}$ for every future time steps $t' > t$ for every unselected segment $b \in \mathcal{B}^t$ with intermediate solutions indexed by $j \in \mathbb{N}^+$, and an overall fixed point problem for \mathbf{S}^t with solutions indexed by $k \in \mathbb{N}^+$.

At time t we aim at evaluating the approximate effect of choosing each of the unselected segments $b \in \mathcal{B}^t$, and compare it to a situation where no action is taken. The difference between the two (Eq. (6)),

$$S_{b\{k+1\}}^t = \begin{cases} \sum_{t' \geq t} NPV_{t'} \left(\Delta \mathbf{p}_{\{k\}}^{t',b,t}, \mathbf{p}_{\{k\}}^{t'-1,b,t} \right) - \sum_{t' \geq t} NPV_{t'} \left(\Delta \mathbf{p}_{\{k\}}^{t',0,t}, \mathbf{p}_{\{k\}}^{t'-1,0,t} \right), & b \in \mathcal{B}^t \\ 0, & \text{otherwise} \end{cases}, \quad (6)$$

is calculated using Eq. (5) using continuous rather than binary vectors as input, and reflects an approximation of the added value of choosing segment b at time t .

However, in order to calculate $S_{b\{k+1\}}^t$ we need the probability vectors $\mathbf{p}_{\{k\}}^{t',b,t}$. As we will see shortly, these are mutually dependent and form a fixed point problem which is solved iteratively across an iteration counter k . An exception to this is for $t' = t$, for which we have the evaluated action at

time t , that is

$$p_i^{t,b,t} = \begin{cases} 1, & i = b \\ u_i^{t-1} & \text{otherwise} \end{cases}, \forall b \in \mathcal{B}^t. \quad (7)$$

Before we can determine the actual probability vector for $t' > t$, it turns out to be relevant to determine the expected number of segments chosen in a given future timestep $t' > t$. We denote this number by $\nu^{t'}$. As it (also) forms a fixed point problem with the probability vector in all future time stages, we index it by j . For a given $t' \geq t$ and possibly an intermediate probability vector $\mathbf{p}_{\{j\}}^{\nu_{\{j\}}^{t'}, b, t}$, $\nu_{\{j\}}^{t'}$ can be determined by dividing the (expected) remaining budget $R_{t'}$ at time t' with the probability weighted average construction costs of the remaining segments $\bar{c}_{\{j\}}^{t', b, t}$,

$$\nu_{\{j+1\}}^{t'} = \frac{R_{t'}}{\bar{c}_{\{j\}}^{t', b, t}} = \begin{cases} \frac{B^t - \sum_{k \in \mathcal{T}: k < t} \kappa^k \mathbf{c}^\top \Delta \mathbf{u}^k - \sum_{k \in \mathcal{T}: k \leq t} \kappa^k \mathbf{m}^\top \mathbf{u}^{k-1}}{\kappa^{t'} \frac{1}{|\mathcal{B}^{t', b}|} \sum_{b' \in \mathcal{B}^{t', b}} c_{b'}}, & j+1 = 0 \\ \frac{B^t - \sum_{k \in \mathcal{T}: k < t} \kappa^k \mathbf{c}^\top \Delta \mathbf{u}^k - \sum_{k \in \mathcal{T}: k \leq t} \kappa^k \mathbf{m}^\top \mathbf{u}^{k-1}}{\frac{\kappa^{t'} \sum_{b' \in \mathcal{B}^{t', b}} \Delta p_{b'}^{\{j\}} c_{b'}}{\sum_{b' \in \mathcal{B}^{t', b}} \Delta p_{b'}^{\{j\}}}}, & j+1 \in \mathbb{N}^+. \end{cases} \quad (8)$$

The probability vector does not only depend on the expected number of selected segments, but also on baseline probabilities \mathbf{w}^t which takes into account the expected performance of each segment. As \mathbf{S}^t measures exactly this, it seems reasonable to include \mathbf{S}^t in the determination of the baseline probabilities. Furthermore, since the investments at each time step are limited by the construction costs of the segments, we also adjust the probabilities according to the construction costs, so that expected net present value increase per construction costs is used in the denominator. The suggestion that this is a good performance indicator is supported by [Paulsen & Rich \(2023\)](#) in which a greedy algorithm based on this ratio yields practically identical results as the optimal solution. This, leads to the following baseline probability expression $\mathbf{w}_{\{k\}}^t$ for Iteration k based on a based on a Multinomial Logit ([McFadden, 1973](#)) formulation:

$$w_{b\{k\}}^t = \begin{cases} \frac{1}{|\mathcal{B}^t|}, & k = 0 \\ \frac{\exp\left(\frac{\mu}{M^t} \frac{S_{b\{k\}}^t}{\kappa^t c_b}\right)}{\sum_{i \in \mathcal{B}^t} \exp\left(\frac{\mu}{M^t} \frac{S_{i\{k\}}^t}{\kappa^t c_i}\right)}, & k \in \mathbb{N}^+, \forall b \in \mathcal{B}^t, \forall t \in \mathcal{T}. \end{cases} \quad (9)$$

Here M^t is the range between the best and worst segment in Iteration 1, i.e.

$$M^t = \max_{b \in \mathcal{B}^t} S_{b\{1\}}^t - \min_{b \in \mathcal{B}^t} S_{b\{1\}}^t, \quad (10)$$

and μ is a hyperparameter.

When evaluating the effect of choosing segment b at time t , the baseline probabilities have to be altered accordingly, such that they still sum to 1 when not taking b into account, i.e.

$$w_{b'\{k\}}^{t,b} = \begin{cases} \frac{w_{b'\{k\}}^t}{1 - w_b^t}, & b' \in \mathcal{B}^{t,b} \\ 0, & \text{otherwise} \end{cases}, t \in \mathcal{T}. \quad (11)$$

Assume now that $\nu \in \mathbb{N}^+$ segments are to be selected among \mathcal{B}^t . As selected segments cannot be unselected, it follows that the segment probability is a monotonously non-decreasing function of ν . Based on the baseline probabilities \mathbf{w}^t and ν , we propose the following recursive definition of the probability of being selected within the first ν segments (excluding b),

$$q_{b'}^{\nu,b,t} = \begin{cases} \frac{w_{b'\{k\}}^{t,b} (1 - Q_{b'}^{\nu-1,b,t})}{\sum_{i \in \mathcal{B}^{t,b}} w_{i\{k\}}^{t,b} (1 - Q_i^{\nu-1,b,t})}, & \nu \leq |\mathcal{B}^{t,b}|, \nu \in \mathbb{N}^+, \forall b' \in \mathcal{B}^{t,b}, \forall b \in \mathcal{B}^t, \forall t \in \mathcal{T}, \\ 1, & \text{otherwise} \end{cases} \quad (12)$$

with

$$Q_{b'}^{\nu,b,t} = \begin{cases} 0, & \nu = 0 \\ \min \left\{ 1, \sum_{n=1}^{\nu} q_{b'}^{n,b,t} \right\}, & \nu \in \mathbb{N}^+, \end{cases} \quad \forall b' \in \mathcal{B}^{t,b}, \forall b \in \mathcal{B}^t, \forall t \in \mathcal{T}. \quad (13)$$

This is the regular probability expression, but corrected by the cumulative probability of being selected within the first ν segments. It can be generalized for non-integer ν 's as follows:

$$p_{b'\{j+1\}}^{\nu_{\{j\}},b,t} = \begin{cases} Q_{b'}^{\lfloor \nu_{\{j\}} \rfloor, b,t} + Q_{b'}^{\lfloor \nu_{\{j\}} \rfloor + 1, b,t} \cdot (\nu_{\{j\}} - \lfloor \nu_{\{j\}} \rfloor), & b' \in \mathcal{B}^{t,b}, \nu \in \mathbb{R}^+, \forall b \in \mathcal{B}^t, \forall t \in \mathcal{T}. \\ 1, & \text{otherwise} \end{cases} \quad (14)$$

Since $\nu_{\{j\}}^{t',b,t}$ and $\mathbf{p}^{\nu_{\{j-1\}},b,t'}$ are mutually dependent, the determination of the two forms a fixed point problem across j . Empirically, since the calculations are very fast, the problem converges quickly. Once the fixed point problem has been solved, i.e. when $\left\| \nu_{\{j\}}^{t',b,t} - \nu_{\{j-1\}}^{t',b,t} \right\|_{\infty} < \epsilon_{\nu}$, the resulting $\nu_{\{j\}}^{t'}$ is denoted by $\nu^{t',b,t}$ and we assign $\mathbf{p}^{t',b,t}_{\{k\}} \leftarrow \mathbf{p}^{\nu^{t',b,t},b,t}$.

By doing this for all $t' > t$, the vector $\mathbf{S}_{\{k+1\}}^t$ can be obtained from Eq. (6) for a given k , and $k \leftarrow k + 1$ can be incremented. The $\mathbf{S}_{\{k\}}^t$ is then used to update the baseline probabilities $\mathbf{w}_{\{k\}}$. By applying the Method of Successive Averages (Robbins & Monro, 1951; Sheffi, 1985) on the sequence of $\mathbf{S}_{\{k\}}^t$, the sequence have been found to converge in our application.

When $\left\| \mathbf{S}_{\{k\}}^t - \mathbf{S}_{\{k-1\}}^t \right\|_{\infty} < \epsilon_S$ for some k , the optimal strategy at time t can be determined by assuming linear independence between the elements of $\mathbf{S}_{\{k\}}^t$, and solving Problem 2 using $\mathbf{S}_{\{k\}}^t$ as \mathbf{S}^t .

3 RESULTS AND DISCUSSION

We test our proposed methodology on a large-scale network of Greater Copenhagen, where we consider the expansion of 43 proposed cycle superhighway routes divided into 202 segments (see Figure 3) over a 50 year planning period. Each of the 202 segments has specific construction and maintenance costs from Incentive (2018), and at each $t \in \{1, 2, \dots, 50\}$ the available budget is given by $B^t = 50 \cdot t$ mill. DKK. The set of origin-destination pairs and traveler types Ω are taken from Hallberg et al. (2021) and contains the combinations of 258 origins and destinations and nine traveller types (combinations of speed preference and bicycle technology, see Hallberg et al. (2021) for details), leading to a total of 596,754 entries.

Table 2 summarizes the various costs, benefits and performance measures associated to each of the seven applied solution strategies. **W/ demand effects** is our proposed method, whereas **W/o demand effects** is the solution where demand effects is not taken into account (Paulsen & Rich, 2023). The bottom-five strategies are baseline reference strategies that are not based on optimization.

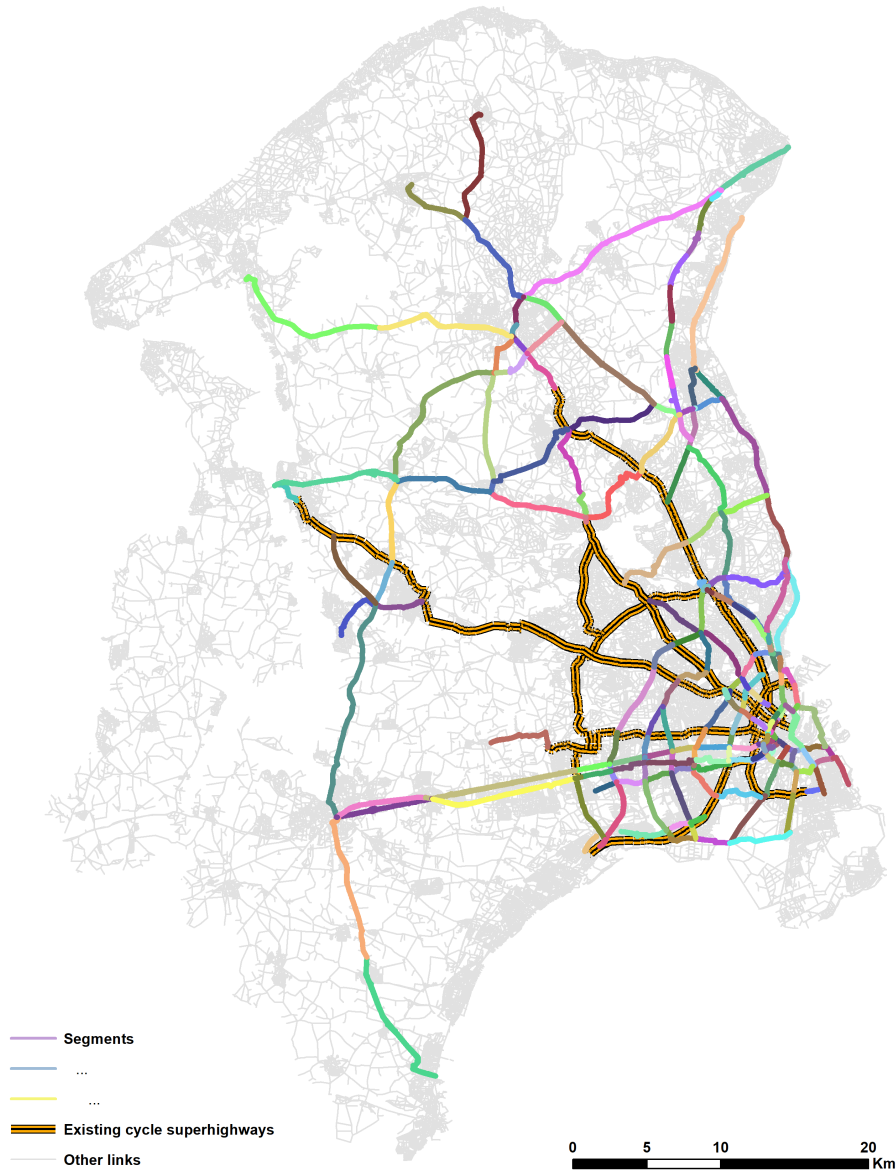


Figure 3: The 202 segments forming the 580km planned future cycle superhighway network extension for the Greater Copenhagen area (Sekretariatet for Supercykelstier, 2019) as well as the existing network.

Strategy	Construction costs	Scrap value	Maintenance costs	Consumer surplus	Health benefits	Net present value	Benefit-cost ratio*
W/ demand effects	715.8	181.9	1,080.5	2,746.3	14,890.5	16,022.4	10.9
W/o demand effects**	321.7	60.4	445.9	2,197.6	11,270.0	12,760.4	19.0
Random order	986.7	358.6	1,406.4	1,153.4	6,526.7	5,645.6	3.77
Shorter segments first	997.2	358.6	1,387.8	1,864.1	9,001.1	8,839.6	5.36
Shorter routes first	998.9	358.6	1,386.9	1,570.3	8,166.4	7,709.5	4.80
Longer segments first	983.6	358.6	1,410.4	694.4	5,187.5	3,846.4	2.89
Longer routes first	978.8	358.6	1,414.7	1,235.7	7,997.7	7,198.5	4.54

Table 2: Investment key-performance indicators [mill. DKK]. *Benefit-cost ratio is dimensionless. ** The solution found with the methodology from Paulsen & Rich (2023).

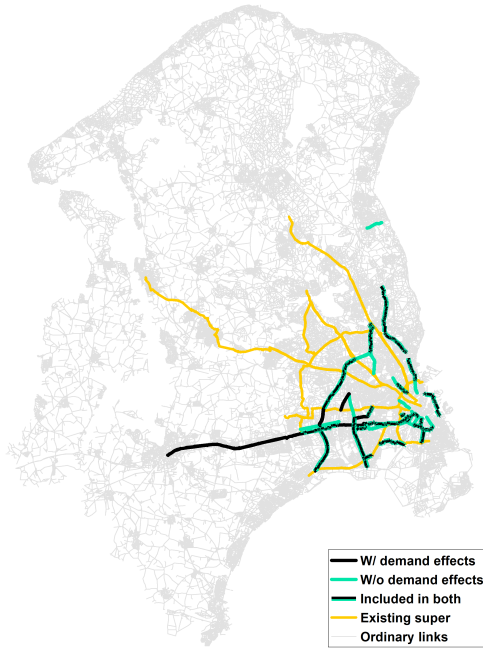
From the results of the baseline strategies it is clearly shown that the overall project portfolio is profitable, leading to net present values between 3.8 and 8.8 billion DKK – largely driven by the health benefits from added bicycle kilometers. The variation in net present value across the baseline reference strategies are substantial, underlining that the order in which segments are implemented have a large effect on the socioeconomic performance. It is also seen that taking a mathematical optimization approach leads to large net present value improvements of at least 3.9 billion DKK when using the method from [Paulsen & Rich \(2023\)](#) (**W/o demand effects**) and 7.2 billion DKK with the approach proposed in this study (**W/ demand effects**), when compared to the best baseline reference strategy (**Shorter segments first**). Thus, the improved methodology leads to a net present value increase that is 83% higher than that of [Paulsen & Rich \(2023\)](#), underlining that taking demand effects into account in the optimization is highly important. Based on the raw net present values of 16.0 billion DKK (**W/ demand effects**) and 12.8 billion DKK (**W/o demand effects**), the relative improvement is 26%.

We note that our optimization routine maximizes an approximation of the net present value, why it is not surprising that the method from [Paulsen & Rich \(2023\)](#) leads to a higher benefit-cost ratio. Especially since that method stops when further expansions are no longer deemed profitable without considering demand effects. When considering these effects, more segments are deemed profitable, leading to a premature stop of investments for the **W/o demand effects** strategy. Only investing in the most profitable segments naturally lead to a high benefit-cost ratio, but fails to achieve the full potential net present value.

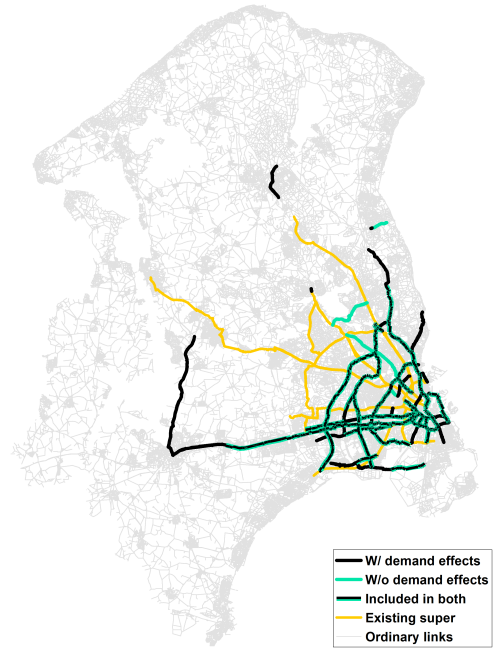
Discussion

We consider the same case study and project portfolio as in [Hallberg et al. \(2021\)](#); [Rich et al. \(2021\)](#), and [Paulsen & Rich \(2023\)](#). Our demand model shares many similarities with [Hallberg et al. \(2021\)](#) and [Rich et al. \(2021\)](#) in that we apply similar level-of-service data and model parameters. However, in the present study we only consider choice of mode and not choice of destination. When upgrading the entire network, we get a relative increase in the number of trips of 3.7%, which compares to an increase of 4.5% in [Hallberg et al. \(2021\)](#). The difference is due to not considering choice of destination. Also, the increase in average cycled trip distance of 8.3% are in line with the 7-8% of [Rich et al. \(2021\)](#). It suggest that our demand sensitivity are largely in line with previous findings.

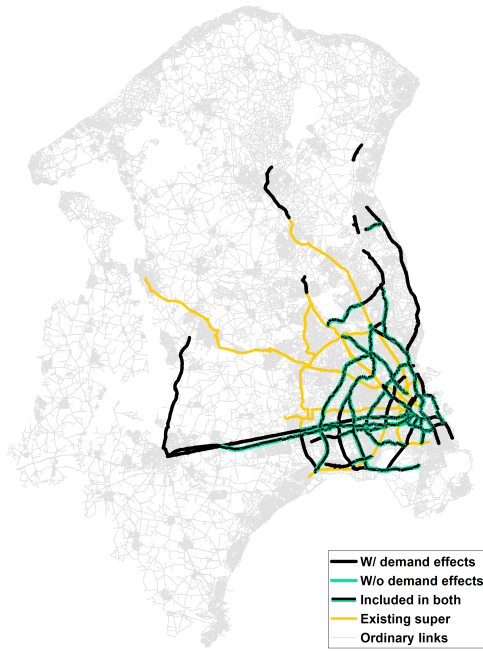
In [Figure 4](#) we compare the solution of our proposed method with that of [Paulsen & Rich \(2023\)](#) that does not incorporate demand effects. Clearly, we see that including such effects encourage building longer routes further away from the city center and cause more segments to be profitable from a socioeconomic point-of-view. Hence, the integration of demand effects implies not only a sizable increase in the welfare contribution, but change the spatial investment pattern as well. The fact that the investment pattern becomes more spatially scattered have some positive indirect implications for the practical implementation of such strategies. Where the solution without demand effect is concentrated mostly the in city center, and hence discourages other municipalities from taking part in the investment scheme, the improved solution actually goes across the geography and makes it highly relevant for municipalities to collaborate when upgrading the infrastructure.



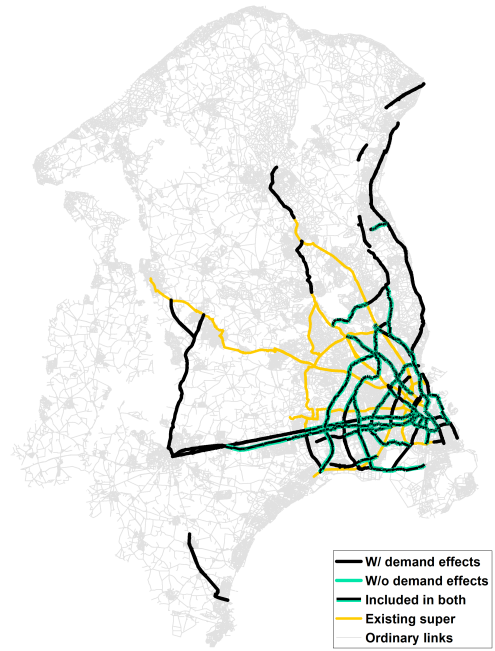
(a) Year 4, cumulative budget of 200 mill. DKK



(b) Year 12, cumulative budget of 600 mill. DKK



(c) Year 20, cumulative budget of 1,000 mill. DKK



(d) Year 28, cumulative budget of 1,400 mill. DKK

Figure 4: Spatial comparison of obtained solutions with (present study) and without (Paulsen & Rich (2023)) demand effects.

4 CONCLUSIONS

With this study we develop and show the large-scale applicability of a methodology for societally optimal expansions of bicycle networks where demand is integrated into the problem. The proposed methodology leads to massive societal benefits with a net present value exceeding 16 billion DKK, in the range of 81%–417% higher than the baseline reference strategies, and 26% higher than the solution found without taking demand effects into account as presented in [Paulsen & Rich \(2023\)](#).

Despite providing a significant contribution to the literature, several research avenues remain open for future research. Methodologically, it is relevant to investigate alternative ways of calculating the expected future net present value contribution of segments (S^t) and compare the performance of the different variations. It will also be relevant to test the effect of a less sensitive demand response, and to incorporate more advanced demand models that allow modeling the composition of regular bicycle users versus electric bicycle users dynamically as network changes occur. As demand effects are even more pronounced for electric bicycle users that travel further ([Hallberg et al., 2021](#)), it is of particular interest to investigate if and how such dynamic modeling of the share of electric bicycle users would alter the optimal infrastructure plans. Future research also includes looking further into regional distribution effects, how to integrate regional budget constraints, and, not least, how these would affect the solution and socio-economic performance measures.

REFERENCES

- Breda, J., Jakovljevic, J., Rathmes, G., Mendes, R., Fontaine, O., Hollmann, S., ... Galea, G. (2018). Promoting health-enhancing physical activity in europe: Current state of surveillance, policy development and implementation. *Health Policy*, 122(5). doi: 10.1016/j.healthpol.2018.01.015
- Chan, T. C. Y., Lin, B., & Saxe, S. (2022, 9). A Machine Learning Approach to Solving Large Bilevel and Stochastic Programs: Application to Cycling Network Design. *arXiv:2209.09404*. Retrieved from <https://arxiv.org/abs/2209.09404> doi: 10.48550/arxiv.2209.09404
- Duthie, J., & Unnikrishnan, A. (2014, 7). Optimization Framework for Bicycle Network Design. *Journal of Transportation Engineering*, 140(7), 04014028. Retrieved from <http://ascelibrary.org/doi/10.1061/%28ASCE%29TE.1943-5436.0000690> doi: 10.1061/(ASCE)TE.1943-5436.0000690
- Hallberg, M., Rasmussen, T. K., & Rich, J. (2021, 7). Modelling the impact of cycle superhighways and electric bicycles. *Transportation Research Part A: Policy and Practice*, 149, 397–418. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0965856421001130> doi: 10.1016/j.tra.2021.04.015
- Incentive. (2018). *Samfunnsøkonomisk analyse af supercykelstierne*. Technical report.
- Liaw, A., & Lin, J.-J. (2022, 7). Bikeway network design model considering utilitarian and recreational bicycling in urban built-up areas. *International Journal of Sustainable Transportation*, 1–14. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/15568318.2022.2101406> doi: 10.1080/15568318.2022.2101406
- Lim, J., Dalmeijer, K., Guhathakurta, S., & Van Hentenryck, P. (2021, 7). *The Bicycle Network Improvement Problem: Optimization Algorithms and A Case Study in Atlanta*. Retrieved from <http://arxiv.org/abs/2107.04451v1> doi: 10.48550/arxiv.2107.04451
- Lin, J.-J., & Liao, R.-Y. (2016, 3). Sustainability SI: Bikeway Network Design Model for Recreational Bicycling in Scenic Areas. *Networks and Spatial Economics*, 16(1), 9–31. Retrieved from <http://link.springer.com/10.1007/s11067-014-9245-7> doi: 10.1007/s11067-014-9245-7
- Lin, J.-J., & Yu, C.-J. (2013, 1). A bikeway network design model for urban areas. *Transportation*, 40(1), 45–68. Retrieved from <http://link.springer.com/10.1007/s11116-012-9409-6> doi: 10.1007/s11116-012-9409-6
- Liu, H., Szeto, W. Y., & Long, J. (2019, 7). Bike network design problem with a path-size logit-based equilibrium constraint: Formulation, global optimization, and matheuristic.

- Transportation Research Part E: Logistics and Transportation Review*, 127, 284–307. doi: 10.1016/J.TRE.2019.05.010
- Martin, B., Kahlmeier, S., Racioppi, F., & et al. (2006). Evidence-based physical activity promotion - hepa europe, the european network for the promotion of health-enhancing physical activity. *J Public Health*, 14. doi: doi.org/10.1007/s10389-006-0029-5
- Mauttone, A., Mercadante, G., Rabaza, M., & Toledo, F. (2017). Bicycle network design: model and solution algorithm. *Transportation Research Procedia*, 27, 969–976. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S2352146517310165> doi: 10.1016/j.trpro.2017.12.119
- McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behaviour. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). New York, NY, USA: Academic Press New York.
- Mesbah, M., Thompson, R., & Moridpour, S. (2012). Bilevel Optimization Approach to Design of Network of Bike Lanes. *Transportation Research Record: Journal of the Transportation Research*, 2284, 21–28. doi: 10.3141/2284-03
- Ospina, J. P., Duque, J. C., Botero-Fernández, V., & Montoya, A. (2022, 5). The maximal covering bicycle network design problem. *Transportation Research Part A: Policy and Practice*, 159, 222–236. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0965856422000325> doi: 10.1016/J.TRA.2022.02.004
- Paulsen, M., & Rich, J. (2023). *Societally optimal expansion of bicycle networks*. Manuscript under review at Transportation Research Part B: Methodological.
- Rich, J., Jensen, A. F., Pilegaard, N., & Hallberg, M. (2021). Cost-benefit of bicycle infrastructure with e-bikes and cycle superhighways. *Case Studies on Transport Policy*, 9(2), 608–615. Retrieved from <https://doi.org/10.1016/j.cstp.2021.02.015> doi: 10.1016/j.cstp.2021.02.015
- Robbins, H., & Monro, S. (1951). A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, 22(3), 400 – 407. Retrieved from <https://doi.org/10.1214/aoms/1177729586> doi: 10.1214/aoms/1177729586
- Sekretariatet for Supercykelstier. (2019). *Tag cyklen på arbejde: Sundt, nemt og sikkert*. Technical report. Retrieved from www.supercykelstier.dk
- Sheffi, Y. (1985). *Urban transportation networks: Equilibrium analysis with mathematical programming methods* (No. 1). Englewood Cliffs, NJ, USA: Prentice-Hall. Retrieved from http://web.mit.edu/sheffi/www/selectedMedia/sheffi_urban_trans_networks.pdf
- Smith, H. L., & Haghani, A. (2012). A Mathematical Optimization Model for a Bicycle Network Design Considering Bicycle Level of Service. In *Transportation research board 91st annual meeting, washington dc*. Washington DC. Retrieved from <https://trid.trb.org/view/1130217>
- Technical University of Denmark. (2022). *Transport Economic Unit Prices vers. 2.0*. Center for Transport Analytics for the Danish Ministry of Transport. Retrieved from <https://www.cta.man.dtu.dk/modelbibliotek/teresa/transportoekonomiske-enhedspriser>
- van Goeverden, K., Nielsen, T. S., Harder, H., & van Nes, R. (2015). Interventions in Bicycle Infrastructure, Lessons from Dutch and Danish Cases. *Transportation Research Procedia*, 10(10), 403–412. Retrieved from www.sciencedirect.comhttps://linkinghub.elsevier.com/retrieve/pii/S235214651500277X doi: 10.1016/j.trpro.2015.09.090

Lane change behavior on motorways based on naturalistic trajectory data

Felix Hofinger*¹ and Martin Fellendorf¹

¹Institute of Highway Engineering and Transport Planning, Graz University of Technology,
Rechbauerstraße 12/II, 8010 Graz, Austria

SHORT SUMMARY

Lane change behavior is a major aspect in traffic flow modeling. Since only a few empirical analyzes are available, validation of lane change models is frequently limited to macroscopic characteristics (e.g. number of lane changes). In addition, many existing lane change models do not reflect lane change behavior in complex situations realistic enough. Therefore, we investigate naturalistic trajectory data from German motorways and analyze gap acceptance behavior focusing on varying discretionary lane change objectives, especially on cooperative lane changes. We propose a methodology to classify different lane change objectives and analyze the critical time gap using the Raff's method. The results show differences in gap acceptance between varying discretionary lane changes classes. Moreover, we found that drivers who perform a cooperative lane change accept rather low time gaps. The analyzes should provide a basis for validating existing and developing new lane change models.

Keywords: cooperative lane changing, gap acceptance, lane change analyzes, naturalistic trajectory data, traffic flow theory

1 INTRODUCTION

The effects of lane changing on traffic flow characteristics and traffic safety on motorways have been widely studied in the past. According to Zheng et al. (2010) lane changes can lead to traffic oscillations, which have a negative impact on traffic safety and traffic efficiency. To analyze these phenomena microscopic traffic flow simulation (TFS) is often used. Although, in the past new data-driven lane change models have been proposed, human driving behavior - and lane changing in particular - is still not fully understood. Besides the uncertainties in driver perception and anticipation (Endsley,1995; Calvert et al., 2020) it is difficult to classify different driving behavior. Most lane change models integrated into TFS software allow the user to define multiple parameter sets that should reflect different driver behavior groups. In addition, due to differences in gap acceptance behavior, mandatory and discretionary lane changes are typically distinguished. Mandatory lane changes need to be performed in order to follow a path, which is why drivers tend to accept lower time gaps compared to discretionary lane changes, which are usually performed to improve the own driving conditions. However, different incentives why a discretionary lane change is performed exist. Especially, in on-ramp areas and weaving sections drivers often try to assist neighboring merging vehicles by performing a cooperative lane change. Regarding lane changing in general there are still limited empirical analyzes available (Knoop et al., 2012). Nowadays, vehicle trajectory data is frequently used to get insights in this field of research (van Beinum et al., 2018; Sharma et al., 2020; Chauhan et al., 2022) . However, to the best of our knowledge no extensive analyzes regarding lane changing, focusing on gap acceptance properties of cooperative lane changes has been carried out.

This paper fills this gap by analyzing two naturalistic trajectory data-sets (highD & exiD dataset) from German motorways (Krajewski et al.,2018; Moers et al., 2022). For this study, we developed and applied a methodology to designate the lane change start. In a next step we analyzed the relevance of surrounding vehicles during the lane change decision period. Within that period we extracted multiple time steps and applied a probabilistic approach to get an indicator, if a preceding and/or right-preceding was relevant for the lane change decision or not. This was done utilizing the basic principles of the action-based Wiedemann car-following model (Wiedemann, 1974). As a last step, we analyzed the critical gap utilizing the method described in Raff (1950) by

investigating accepted as well as rejected time gaps on the target lane. In addition, we analyzed lane change duration, lane change length as well as the number of lane changes under given traffic flow characteristics.

2 METHODOLOGY OF LANE CHANGE ANALYZES

In this section we first give an overview over the used datasets within this study and explain the methodology to identify the lane change start using trajectory data. Then we show the classification framework of different lane change categories. This sections ends with a description of the gap acceptance analyzes.

Trajectory Dataset

For the presented analyzes we used the highD and exiD dataset, which are naturalistic trajectory datasets recorded by drones on German motorway segments (Krajewski et al.,2018; Moers et al., 2022). The highD dataset consists of 16.5 hours of trajectory data (around 110.500 vehicles, fidelity=25Hz) on six different locations on the open road. Within the exiD dataset there is trajectory data from six different locations (on-ramps, off-ramps, weaving segments) included, recorded within a time period of 16 hours. Although the number of driven kilometers is only about 60% in the exiD dataset compared to the highD dataset, the number of lane changes is much higher. We decided to analyze both datasets in order to be able to compare possible differences between on-/off-ramp areas and open road segments. To visualize the trajectory datasets we used the TraViA tool, which is advantageous for validating results and to analyze specific scenarios (Siebinga, 2021).

Lane change detection

According to Sharma et al. (2020) a vehicles lane-changing process starts when it begins to drift laterally and ends when it stabilizes its lateral position on the target lane. Human drivers drift laterally even in lane-keeping situations, which is why is not a trivial task to mark the starting time instances of a lane change process. We investigated the distance between a vehicles center-point and its left/right lane marking (depending on the lane change direction) since this attribute is not dependent on the road geometry, which is highly relevant in on-ramp areas. Within this study we calculated the exponential growth rate and applied an exponentially weighted moving average (EWMA) filter to smooth the noise in the data. This method is shown in Tang et al. (2020), the formula can be seen below:

$$\bar{x}_\alpha(t_i) = \frac{\sum_{k=i-D}^{I+D} x_\alpha(t_k) e^{-\frac{|i-k|}{\Delta}}}{\sum_{k=i-D}^{I+D} e^{-\frac{|i-k|}{\Delta}}} \quad (1)$$

where $x_\alpha(t_i)$ is the state value after filtering at t_i , D denotes the size of the sliding window and Δ represents the average of the sliding time window. The EWMA filter method leads to an exponentially decreasing weighting coefficient over time.

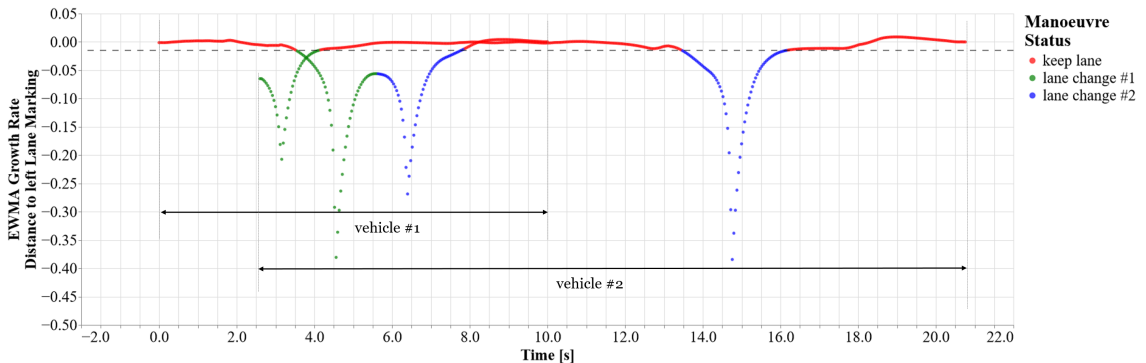


Figure 1: Classification of maneuver status

After analyzing descriptive statistics and visualizing the time series data we set a threshold of the exponential growth rate to $-0.015m$, which marks the time instances, when a lane change begins/ends. The threshold is represented by the dashed horizontal line in figure 1, in which the exponential growth rate of the distance to a vehicle's left/right lane marking of two vehicles are shown. The first vehicle (0.0-10.0s) is performing two lane changes immediately after each other, whereas the second vehicle (2.6-20.7s) does have a lane-keeping phase between the two lane changes. Moreover, vehicle 2 was already performing a lane change when it entered the field of view of the drone. These lane changes were not further investigated as a classification of the lane change objective was not possible.

Classification of lane change objective

In order to classify the drivers lane change objective, we applied a probabilistic approach using the basic principles of the action-based Wiedemann car-following model (Wiedemann, 1974). Using the time instance of the lane change begin as starting point, we extracted multiple time steps (*e.g.* 4, 3, 2, 1, 0s before lane change begin) within the lane change decision period, which we set to maximum 4 seconds, according to the work in Guo et al. (2021). For each of these time instances we calculated a datapoint considering the relative velocity and the distance headway, given that there was a preceding vehicle present. As a next step we computed for all datapoints the difference to a simple linear threshold utilizing the basic principles of the Wiedemann car-following model. A datapoint below the yellow dashed threshold in figure 2a, indicates that a preceding/right-preceding might have been relevant during the lane change decision period. After calculating a difference-value for all datapoints considering all lane changes we computed a distribution (see figure 2b), which allows us to get an indicator regarding the likelihood, that a surrounding vehicle was relevant for the lane change decision.

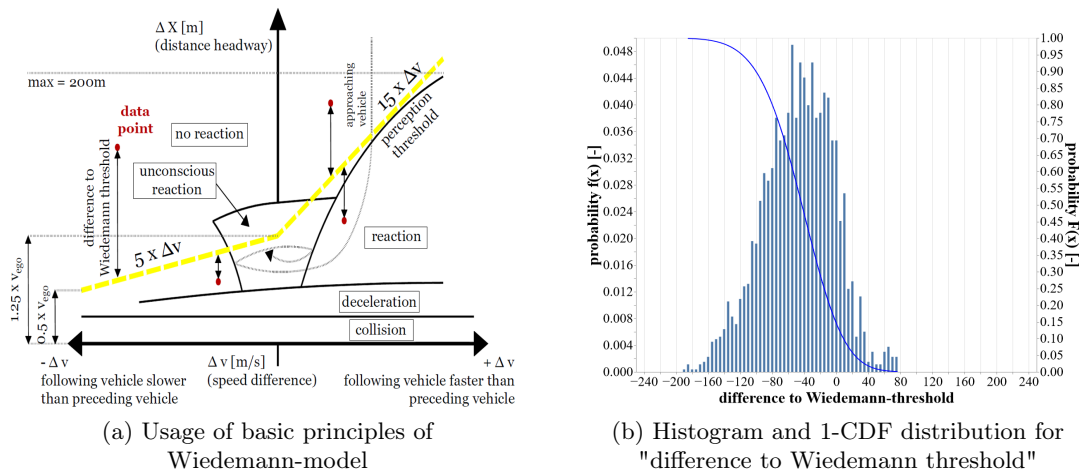


Figure 2: Probabilistic approach to classify relevant surrounding vehicles during lane change decision

Besides the probability value, whether a preceding/right-preceding vehicle was relevant during a certain time instance during the lane change decision, we analyzed a tendency how the probability-value evolves over time. Using these two indicators we defined a framework in order to classify whether a preceding/right-preceding vehicle was relevant or not, which can be seen in table 1. We used this classification in order to categorize different discretionary lane change objectives, which are displayed in table 2. If a vehicle was performing a lane change from the first lane to the second lane and there was no relevant preceding vehicle ($value=-1$) detected but a relevant right-preceding vehicle ($value=1$), then the lane change was classified as a *cooperative* lane change. On the other hand, lane changes were marked as *overtaking*, if there was no relevant right-preceding vehicle present and the drivers intention was to change lane in order to overtake the preceding vehicle. If there was no relevant surrounding vehicle detected in the trajectory data we categorized the lane change objective as *unclassified*. In addition, we introduced a *mixture* class, which represents lane changes, in which both, a relevant preceding and right-preceding vehicle were present. For mandatory lane changes as well as for lane changes to the right-hand side we did not consider the impact of surrounding vehicles.

Table 1: Classification framework of relevant/not-relevant surrounding vehicles

preceding vehicle	probability (p)		tendency	classification
	right-preceding vehicle			
$p > 0.2$	$p > 0.5$		-	1/ relevant
$0.1 < p < 0.2$	$0.2 < p < 0.5$		1/ increasing	1/ relevant
$0.1 < p < 0.2$	$0.2 < p < 0.5$		-1/ decreasing	0/ unclear
$0.025 < p < 0.1$	$0.0 < p < 0.2$		1/ increasing	0/ unclear
$0.025 < p < 0.1$	$0.0 < p < 0.2$		-1/ decreasing	-1/ not relevant
$p < 0.025$	$p < 0.0$		-	-1/ not relevant

Table 2: Classification framework of a drivers lane change objective

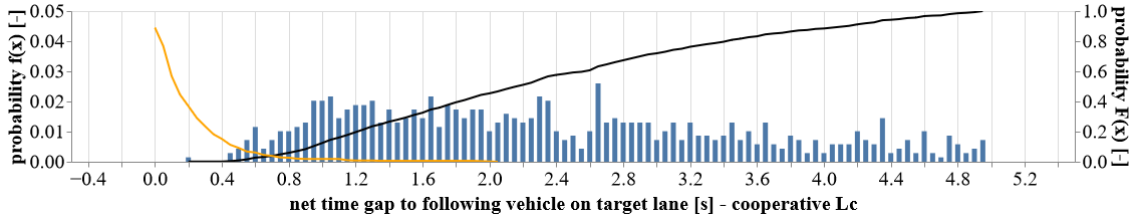
origin lane	target lane	relevance classification		lane change objective classification
		right-preceding vehicles	preceding vehicle	
0	1	-	-	merging
1	0	-	-	diverging
0	2	-	-	merging/overtaking
2 3	1 2	-	-	right
1 2	2 3	1	-1	cooperative
1 2	2 3	1	0	cooperative
1 2	2 3	1	1	mixture
1 2	2 3	0	-1	cooperative
1 2	2 3	0	0	unclassified
1 2	2 3	0	1	overtaking
1 2	2 3	-1	-1	unclassified
1 2	2 3	-1	0	overtaking
1 2	2 3	-1	1	overtaking

Gap acceptance analyzes

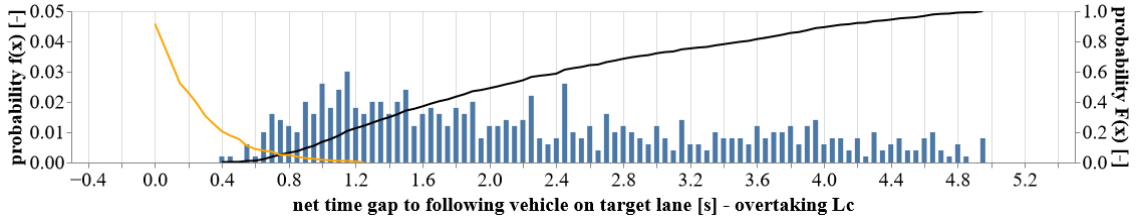
To investigate gap acceptance characteristics, it is necessary to identify the relevant vehicles on the target lane, especially in dense traffic situation in which multiple vehicles are changing lane at the same time. Therefore, we extracted the vehicle IDs on the target lane multiple times during the lane change decision and the lane change execution phase. By doing so we identified situations, in which the vehicle IDs of the preceding and following vehicle on the target lane changed over time, which enabled us to investigate rejected gaps too. If a temporary left following vehicle overtook during the lane change decision period and thus was the preceding vehicle after the lane change, we indicated the time gaps during the lane change decision period as rejected. This approach allows to extract the critical time gap, as it is shown in Raff (1950). Regarding gap acceptance we analyzed the net time gap, net distance headway, relative velocity and time-to-collision (TTC) to the following vehicle on the target lane and for varying lane change objectives. Moreover, we investigated the number of lane changes under given traffic characteristics like traffic volume. Due to space limitations, we will focus on the accepted time gaps within the result discussion in section 3.

3 RESULTS AND DISCUSSION

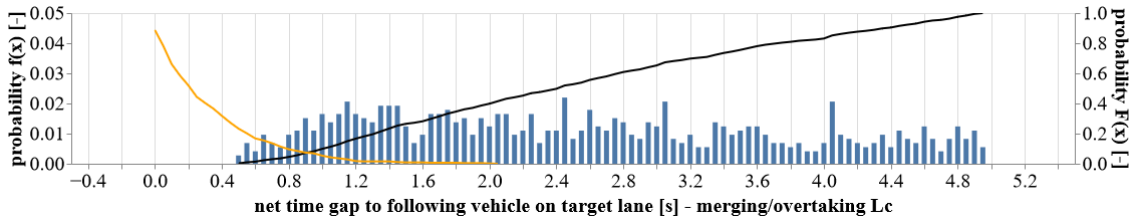
By applying the methodology described in section 2 we computed a distribution for varying lane change objectives considering the net time gap to the following vehicle on the target lane. Relatively high time gaps to the following vehicle have a rather low impact on gap acceptance behavior, which is why we excluded all time gaps bigger than 5s. In figure 3 the gap acceptance distributions for different discretionary lane change objective classes can be seen. The histogram and the corresponding black cumulative distribution function show the accepted time gaps in the instance of the lane change start. The yellow function reflects the cumulative distribution function of all rejected gaps. The intersection between the distributions is defined as the critical gap (Raff, 1950). Comparing the distribution functions shows, that there is a minor difference between cooperative lane changes (3a) and lane changes, which are performed in order to overtake a preceding vehicle (3b). The accepted gaps as well as the critical gap of 0.68s for cooperative lane changes is rather low. This could be due to the fact, that altruistic behavior is mainly observed among experienced drivers, since they are able to anticipate surrounding vehicles pretty good. Figure 3c shows a critical gap of 0.9s for lane changes, which are performed immediately after conducting a merging lane change from an on-ramp. Hence, these complex lane change manoeuvres are mainly performed if the time gap on the target lane is sufficient big. This can be seen in the black distribution function as well, which shows less time gaps below 1s. In figure 3d the distribution of accepted time gaps for lane changes to the right-hand side is displayed. The following vehicles on the target lane are usually slower, resulting in lower accepted time gaps compared to all other lane change classes. We did not display a distribution for rejected time gaps, as not sufficient situations could be observed.



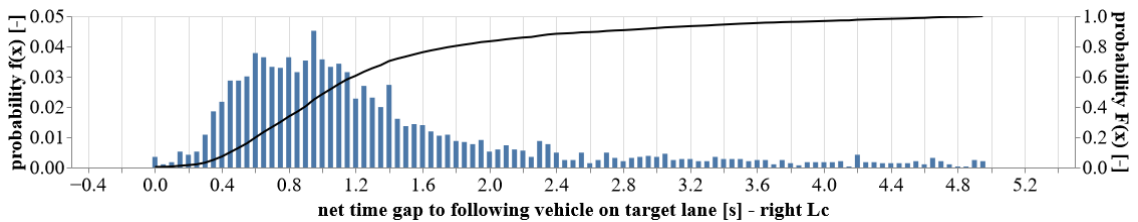
(a) analyzing critical gap - cooperative lane changes



(b) analyzing critical gap - overtaking lane changes



(c) analyzing critical gap - merging/overtaking lane changes



(d) Gap acceptance - right lane changes

Figure 3: Gap acceptance analyzes for different lane change objectives

4 CONCLUSIONS AND OUTLOOK

This paper presents empirical analyzes regarding human lane change behavior by investigating naturalistic trajectory data recorded at German motorway segments. We showed a methodology to detect the time instance, when a vehicle starts to drift laterally and we proposed a probabilistic rule-based classification of different lane change objectives. Moreover, we analyzed gap acceptance characteristics for different discretionary lane change objectives. Surprisingly, drivers who perform a cooperative lane change accept rather low time gaps to their following vehicle on the target lane. This might be the case as these drivers are usually quite experienced and are able to anticipate the surrounding vehicles quite good. However, to underline these assumptions further investigations also with data from driver simulator studies should be focused on. In addition, we plan to train a Long-short-term-memory (LSTM) Neural Network for classifying different lane change objectives, which might lead to improved intermediate results. The presented analyzes gave new insights regarding human lane change behavior and they can serve as a good starting point for further validation of existing lane change models or even for the development of new models.

ACKNOWLEDGEMENTS

This work received funding from FFG, the Austrian Research Promotion Agency (Grant-No.: 891143) in course of the TRIDENT-project. Additionally, we gratefully acknowledge the support from Prof. Hans van Lint & and Dr. Simeon Calvert (TU Delft) during a study visit.

REFERENCES

- Calvert, S. C., Schakel, W. J., & van Lint, J. (2020). A generic multi-scale framework for microscopic traffic simulation part ii – anticipation reliance as compensation mechanism for potential task overload. *Transportation Research Part B: Methodological*, 140, 42–63. doi: 10.1016/j.trb.2020.07.011
- Chauhan, P., Kanagaraj, V., & Asaithambi, G. (2022). Understanding the mechanism of lane changing process and dynamics using microscopic traffic data. *Physica A: Statistical Mechanics and its Applications*, 593, 126981. doi: 10.1016/j.physa.2022.126981
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64. doi: 10.1518/001872095779049543
- Guo, Y., Zhang, H., Wang, C., Sun, Q., & Li, W. (2021). Driver lane change intention recognition in the connected environment. *Physica A: Statistical Mechanics and its Applications*, 575, 126057. doi: 10.1016/j.physa.2021.126057
- Knoop, V. L., Hoogendoorn, S. P., Shiomi, Y., & Buisson, C. (2012). Quantifying the number of lane changes in traffic. *Transportation Research Record: Journal of the Transportation Research Board*, 2278(1), 31–41. doi: 10.3141/2278-04
- Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *2018 21st international conference on intelligent transportation systems (itsc)* (p. 2118-2125). doi: 10.1109/ITSC.2018.8569552
- Moers, T., Vater, L., Krajewski, R., Bock, J., Zlocki, A., & Eckstein, L. (2022). The exid dataset: A real-world trajectory dataset of highly interactive highway scenarios in germany. In *2022 ieee intelligent vehicles symposium (iv)* (p. 958-964). doi: 10.1109/IV51971.2022.9827305
- Raff, M. S. (1950). *A volume warrant for urban stop signs*. Retrieved from <https://rosap.ntl.bts.gov/view/dot/16265>
- Sharma, S., Snelder, M., Tavasszy, L., & van Lint, H. (2020). Categorizing merging and diverging strategies of truck drivers at motorway ramps and weaving sections using a trajectory dataset. *Transportation Research Record: Journal of the Transportation Research Board*, 2674(9), 855–866. doi: 10.1177/0361198120932568

- Siebinga, O. (2021). Travia: a traffic data visualization and annotation tool in python. *Journal of Open Source Software*, 6(65), 3607. doi: 10.21105/joss.03607
- Tang, L., Wang, H., Zhang, W., Mei, Z., & Li, L. (2020). Driver lane change intention recognition of intelligent vehicle based on long short-term memory network. *IEEE Access*, 8, 136898–136905. doi: 10.1109/ACCESS.2020.3011550
- van Beinum, A., Farah, H., Wegman, F., & Hoogendoorn, S. (2018). Driving behaviour at motorway ramps and weaving segments based on empirical trajectory data. *Transportation Research Part C: Emerging Technologies*, 92, 426–441. doi: 10.1016/j.trc.2018.05.018
- Wiedemann, R. (1974). Simulation des straßenverkehrsflusses. *Schriftenreihe des Instituts für Verkehrswesen der Universität Karlsruhe*(8).
- Zheng, Z., Ahn, S., & Monsere, C. M. (2010). Impact of traffic oscillations on freeway crash occurrences. *Accident; analysis and prevention*, 42(2), 626–636. doi: 10.1016/j.aap.2009.10.009

Do price reductions attract customers in urban public transport? A synthetic control approach

Hannes Wallimann¹, Kevin Blättler², Widar von Arx²

¹ PhD, Institute of Tourism and Mobility, University of Applied Sciences and Arts Lucerne, Switzerland

² Institute of Tourism and Mobility, University of Applied Sciences and Arts Lucerne, Switzerland

SHORT SUMMARY

In this paper, we assess the demand effects of lower public transport fares in Geneva, an urban area in Switzerland. Considering a unique sample based on transport companies' annual reports, we find that, when reducing the costs of annual season tickets, day tickets, and hourly tickets (by up to 29%, 6%, and 20%, respectively), demand increases over five years by about 10.6%. To the best of our knowledge, we are the first to show how the synthetic control method can be used to assess such (for policy-makers) important price reduction effects in urban public transport. Furthermore, we propose an aggregate metric that inherits changes in public transport supply (e.g., frequency increases) to assess these demand effects, namely passenger trips per vehicle kilometre. This metric helps us to isolate the impact of price reductions by ensuring that companies' supply changes do not affect estimators of interest.

Keywords: policy evaluation, price reduction, urban public transport pricing, synthetic control method

1. INTRODUCTION

The transport sector is a pivotal contributor to air pollution. Globally, approximately 27% of CO₂ emissions and energy consumption are caused by the transport sector; in the European Union, the figure amounts to about a third (Batty et al., 2015). Therefore, the transport sector also causes negative externalities, which means a situation in which the action of a person imposes a cost on another person who is not a party to the transaction. Private car use will lead to even greater levels of such negative externalities, which a shift in transport mode towards public transport could help to reduce. Lower fares are a frequently discussed tool to motivate individuals to use public transport (see, e.g., Redman et al., 2013).

However, policy-makers must know how existing and potential customers respond to such lower fares. In reality, it is generally challenging to identify the causal effect of lower fares on public transport demand as transport supply changes over time. Therefore, we propose an aggregate metric that inherits a transport company's supply in public transport demand. The metric is composed of passenger trips per vehicle kilometre. Moreover, considering CO₂ emissions, an increase in the metric points to an average emission decrease by each passenger.

In our comparative case study,¹ we use this metric as the outcome variable to analyze lower fares empirically in the case of Geneva, an urban area in Switzerland. There, the electorate decided to

¹ Note that this paper is based on a more extended paper, see Wallimann et al. (2023).

reduce the price of state-owned public transport, which Geneva introduced in December 2014. The reduction amounted to up to 29% for annual season tickets, 6% for day tickets, and 20% for tickets valid for one hour. The policy intervention was the largest price reduction in a long time. As a result, the annual season ticket in Geneva now costs 500 Swiss francs for adults (previously 700 Swiss francs) and 400 Swiss francs for seniors and juniors (previously 500 and 450 Swiss francs, respectively). These prices are more than 200 Swiss francs less than those charged by other Swiss cities. For instance, annual season tickets in Lausanne, Berne, Basel, and Zurich cost 740, 790, 800, and 782 Swiss francs, respectively. The same is the case for single-fare tickets amounting to 3 Swiss francs in Geneva.

To illustrate the price-reduction effect, we analyze the case of TPG, the main operator in the city of Geneva, and its agglomeration belt. To this end, we apply the synthetic control method (Abadie, Diamond, and Hainmueller, 2010, Abadie and Gardeazabal, 2003) to construct a synthetic TPG, a counterfactual that mimics the demand the company would have experienced in the absence of the price reduction. The thing to notice is that the methodology uses a data-driven procedure to create the synthetic TPG from comparable Swiss transport operators.

2. METHODOLOGY

In this section, we outline the synthetic control method used in our empirical analysis. Second, we present the assumptions underlying our analysis.

Methodology and implementation

Let D denote the binary treatment 'price reduction' and Y the outcome 'public transport demand'. The treatment D , the result of the initiative in Geneva, affects one unit (TPG). All the other units (transport companies) in our data are not exposed to the price reduction and thus constitute the control group. We can define the observed outcome of TPG, our unit of interest, as

$$Y_t = Y_t^N + \alpha_t D_t.$$

Y_t denotes the observed outcome, Y_t^N the outcome without the treatment, and α_t the treatment effect at time t . It is important to note that the treatment D takes the value 0 for all units during the period $t < T_0$, with T_0 indicating the introduction of the treatment. This is because also TPG was not exposed to the price reduction during the pre-treatment period. Only looking at the post-treatment period permits to define the treatment effect as

$$\alpha_t = Y_t - Y_t^N.$$

As we observe Y_t , we merely need to estimate Y_t^N , the public transport demand of TPG without the policy intervention. Using statistical parlance, Y_t^N is a counterfactual. That is the outcome one would expect if the intervention had not been implemented. To determine Y_t^N , we use the synthetic control method of Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010). To construct the synthetic control unit (Y_t^N), the synthetic control method uses a data-driven procedure. In our study, the counterfactual Y_t^N , the synthetic TPG, is created out of already-existing companies of the control group, the so-called 'donor pool'. For this purpose, the methodology assigns a weight to each transport company in the control group. These weights are non-negative and sum up to one. On the one hand, we assign large weights to companies with a sizeable predictive power for TPG. On the other hand, transport companies in the control group with a low predictive power receive a small or a zero weight. The goal is to minimize the difference

between TPG and the synthetic TPG in the period $t < T_0$, the pre-treatment period. To discuss the success of this endeavor, we calculate the mean squared prediction error (MSPE) of the outcome variable between TPG and the synthetic TPG. In implementing the synthetic control method, we use the `synth` and `SCtools` packages for the statistical software R by Hainmueller and Diamond (2015) and Silva (2020), respectively.

To empirically challenge the results, we calculate the corresponding 95% bootstrap confidence intervals to the average treatment effect. Therefore, we randomly draw control units with replacement from our donor pool 2,000 times to arrive at these confidence intervals. In every sample, we construct a synthetic TPG and estimate the average gap between TPG and its counterfactual.

Assumptions

Identification requires statistical procedures, as explained in the previous subsection. However, on the other hand, ensuring that our calculation identifies the effect of the price reduction also relies on assumptions about how the world, here the world of public transportation, works (see, e.g., Huntington-Klein, 2021). Therefore, in the following, we discuss the contextual assumptions underlying our analysis (see also Abadie, 2021).

Assumption 1 (no anticipation) is satisfied when the public transport demand in Geneva did not change due to forward-looking customers reacting in advance to the policy intervention. To this end, the price reduction effect would be biased if TPG's travelers had already used public transport before the intervention because they knew that prices would fall later.

By *Assumption 2 (availability of a comparison group)*, there exists a donor pool. The assumption is satisfied when we have a control group with characteristics that are, by assumption, comparable to the treated unit. That implies that other public transport companies do not sharply lower fares in our natural experiment.

Assumption 3 (convex hull condition) is satisfied when pre-treatment outcomes of the synthetic counterfactual can approximate the outcomes of the treated unit. Using statistical parlance, the pre-treatment outcomes of the treated unit are not 'too extreme' (too high or low) compared to the outcomes of the donor pool.

Assumption 4 (no spillover effects) is fulfilled when the price reduction has no spillover effects, either positive or negative, on other transport companies in the donor pool. An obvious failure of this assumption would be a decrease in public transport demand in other Swiss cities because their residents perceive the ticket costs as too high after the price reduction in Geneva.

Assumption 5 (no external shocks): Applying the synthetic control method, we assume that no shocks occur to the outcome of interest during the study period (see, e.g., Abadie, 2021). In our case, this condition is challenging since public transport companies expand the network from time to time, which typically affects the demand for public transport (see, e.g., Brechan, 2017, Holmgren, 2007). To account for such changes in supply, we propose an aggregate metric that breaks down the demand for public transport per company's supply, which we use as our outcome variable. More precisely, we calculate the ratio of passenger trips per vehicle kilometre, being robust against changes on the supply side. Finally, note that, to our knowledge, no large-scale road or parking policy was introduced in the areas of interest during the study period.

3. RESULTS AND DISCUSSION

We subsequently introduce the data underlying this paper. Then, we present the results of applying the synthetic control method, evaluate their significance and investigate their robustness.

Data

To investigate the effect of the policy intervention in Geneva, we use the annual reports of Swiss transport companies, which the Swiss National Library systematically archives. We systematically gathered the most relevant performance indicators from public transport companies for our dataset. TPG operates mainly in the city of Geneva, the densest and second largest city in Switzerland, and its agglomeration belt. Using the synthetic control method, we must choose each unit in the donor pool judiciously to provide a reasonable control for TPG, the treated unit (see Assumption 2). Therefore, we only consider transport companies that operate trams and buses primarily in cities with more than 50,000 inhabitants. These are Bernmobil (Berne), BVB (Basel), SBW (Winterthur), TL (Lausanne), TPL (Lugano), VB (Biel), VBL (Lucerne), VBSG (St Gallen), and VBZ (Zurich).

First, we collected the number of passenger trips, which are standardized in Switzerland. The number of passenger trips counts how many passengers enter a company's vehicle per year. Passenger trips are essential, as we want to measure the increase in public transportation use. Today, companies mainly count passengers automatically, but this was often done by hand in the past. This change in the counting system happened in Geneva from the years 2015 to 2016. Therefore, we adjust our TPG data from 2016 to 2019 based on the observed growth rate of the passenger trips to have a refined panel dataset. Since 2005, TPG has experienced the highest increase in passenger trips. However, TPG has also experienced a high increase in vehicle kilometres. The increase results from the extension of tram routes. Therefore, to mitigate changes in supply, i.e., external shocks increasing companies' networks (see Assumption 5), we use the previously discussed aggregate metric of passenger trips per vehicle kilometre as the outcome variable.

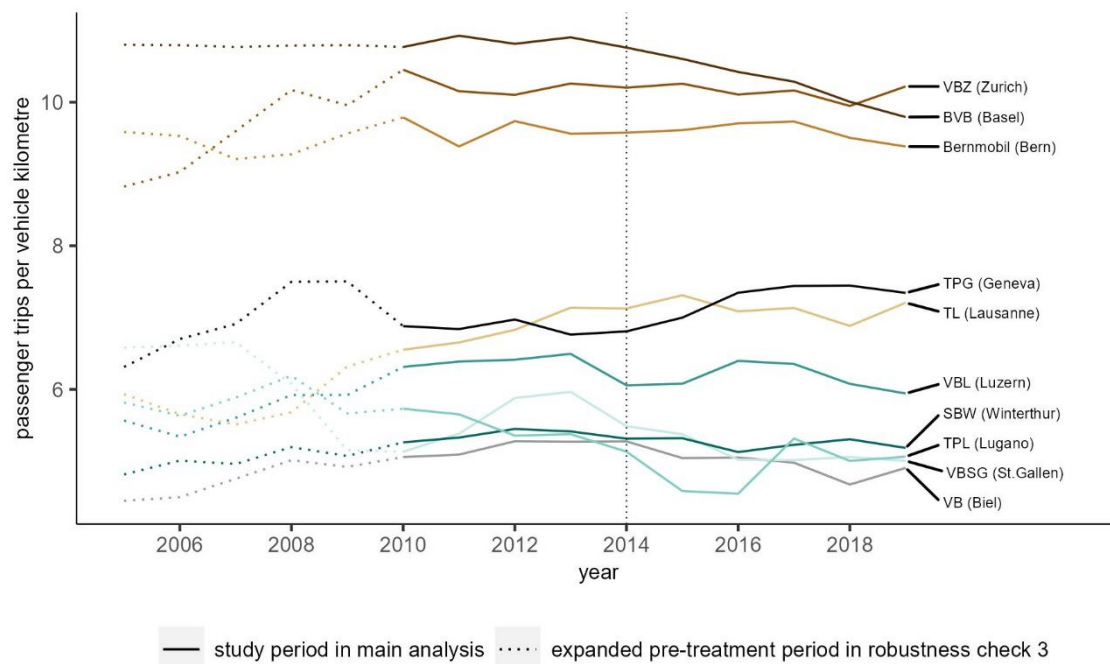


Figure 1 Passenger trips per vehicle kilometre. Note that we restrict our pre-treatment period to 2010 to 2014 (solid lines).

Controlling for supply changes also makes sense, as several studies show considerable effects of vehicle kilometres on demand. For instance, Holmgren (2007) estimates a short-run supply (vehicle kilometres) elasticity of 1.05. Based on this meta-analysis, we assume a considerable supply elasticity of about 1 when applying the ratio. However, and also a thing to notice, due to a substantial increase in vehicle kilometres plied by bus lines in Geneva's agglomeration belt from 2008 to 2010, the ratio in Geneva declined. This is because the aggregate change in TPG's supply occurred in the subarea where public transport is relatively poorly utilized. Therefore, we restrict our pre-treatment period to the years 2010 to 2014. However, collecting several observations on the unit of interest (TPG) and the donor pool is crucial before the price reduction (Abadie, 2021). Therefore, we also perform a robustness check with a more extended pre-treatment period. Moreover, we also oppose our results to estimations without the metric and thus use only passenger trips as the outcome variable. This robustness check is crucial, as unexpected low (or high) supply elasticities could be an alternate explanation of the treatment effect.

The effect of the price reduction

To construct the synthetic TPG, the synthetic control method assigns weights among the control group companies. VB (Biel) receives the highest weight with 0.400, while BVB (Basel) has the second-highest weight with 0.162, and the VBSG (St Gallen) has a zero weight. Figure 2 plots the outcome variable, equal to passenger trips per vehicle kilometre, of TPG and the synthetic TPG from 2010 to 2019. We can easily observe that the two trajectories track each other close in the pre-treatment period, i.e., the pre-price reduction period. Thus, the mean squared prediction error (MSPE) of the outcome variable between TPG and the synthetic TPG amounts to a small figure of 0.009. Therefore, our synthetic TPG is a sensible counterfactual of the outcome we would expect if the intervention had not been implemented. While demand from customers of the synthetic TPG continued its slightly downward trend, the demand for TPG increased. This difference is relatively constant over four years, from 2016 to 2019.

Compared to the synthetic counterfactual, the demand increased by, on average, about 0.72, an increase of 10.6% compared to 2014. Thus, we can infer a positive effect on demand in Geneva due to the price reductions. Randomly drawing nine control units with a replacement from our donor pool leads us to bootstrap confidence intervals. The average estimated effect's corresponding 95% bootstrap confidence interval is [0.423; 0.870].

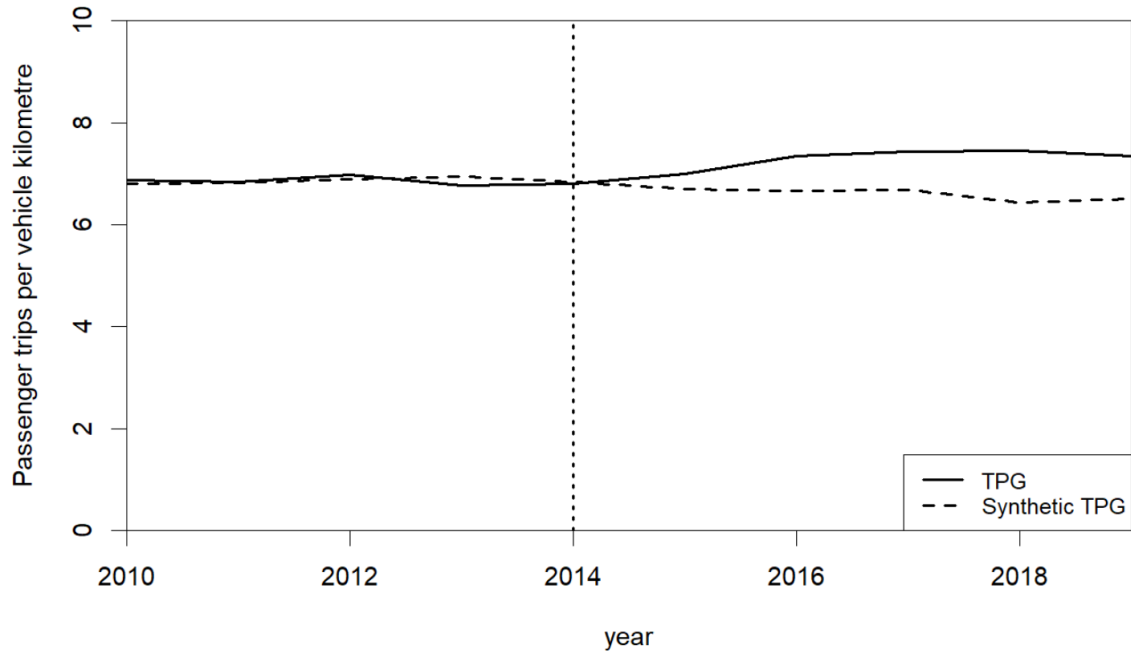


Figure 2: Demand development of TPG and the synthetic TPG

Robustness analysis

Here, we challenge our assumptions and study design by performing robustness investigations. First, as a methodological robustness check, we apply a recent development of the synthetic control method, the synthetic difference in differences approach of Arkhangelsky et al. (2019), to demonstrate the goodness of our results. Second, we expand our pre-treatment period. Third, we expand our donor pool with companies operating in cities with fewer than 50,000 inhabitants. Fourth, we estimate the effect of the lower fares on the number of passengers (and not the number of passengers per vehicle kilometre). Moreover, in the fourth robustness check, the synthetic TPG does not mimic TPG in the pre-treatment period appropriately. Therefore, in a final robustness investigation, we re-estimate the effect on the number of passengers using the synthetic difference in differences approach.

Table 1: Estimates summary of robustness checks

Check	Modification	Price effect
1	Method (SDID)	10.0%
3	More units in the donor pool	10.5%
2	Expanded pre-treatment period	9.0%
4	Passenger trips as outcome variable	Insufficient fit
5	Passenger trips as outcome variable and method (SDID)	3.7%

Note: In robustness checks 1, 2, and 3, we use passenger trips per vehicle kilometres as the outcome variable

In the first, second, and third robustness checks, we use passenger trips per vehicle kilometres as the outcome variable (see Table 1). These robustness checks show that the estimate is robust when we modify the study design, i.e., applying the synthetic difference in differences approach, longer pre-treatment period, or more companies in the donor pool. In the fourth and fifth robustness investigations, we replace our metric with the original number of passenger trips. We consider the synthetic difference in differences methodology more appropriate to analyze the outcome variable passenger trips. However, note that Assumption 5 (no external shocks) is violated when using the outcome variable passenger trips regardless of the methodology.

The results in the fifth robustness check are lower, which shows that our estimate crucially depends on whether we consider the influence of the vehicle kilometres. The demand effect amounts to 3.7% when we apply the synthetic difference in differences methodology. A demand increase of 3.7% is even lower than naively comparing the passenger trips of TPG after and prior to the price discount, amounting to 5.7% additional trips. Moreover, when calculating bootstrap estimates of the effect, we do not get any negative values. The 95% bootstrap confidence interval of the average estimated effect points to an increase of between 2.0% and 12.4%, inclusive. Therefore, we conclude that the effect of 3.7% additional demand is a potential lower bound of the effect.

Elasticities

Considering the revenue shares per ticket category of 2014, we assess an overall price discount of 12.6%. Based on the price discount of 12.6%, we get corresponding point elasticities of demand of -0.84 and -0.29 of our main result and the lower bound, respectively. Therefore, we show that price reductions in urban areas with high-quality public transport attract customers. However, the demand effect is too small to compensate for the loss of revenue due to lower prices. Using the metric passenger trips per vehicle kilometre, we assume high supply elasticities, i.e., about 1, due to findings in the literature (Holmgren, 2007). On the other hand, as a word of caution, high supply elasticity might be lower when public transport quality is high, e.g., Axhausen and Fröhlich (2012). In such a case, increasing or decreasing vehicle kilometres could influence the metric and, therefore, the estimate of interest.

4. CONCLUSIONS

We assess the demand effect of lower urban public transport fares and find that the price reduction in Geneva leads to a demand increase of about 10.6%. To isolate the effect of our mechanism of interest, the price reduction, we propose an aggregate metric inheriting supply changes of public

transport networks. This makes sense as we can block off the effect of increasing and decreasing frequencies as an alternate explanation of demand effects, being in the context of public transport of crucial importance. Moreover, robustness investigations show that the estimate is robust when we modify the study design, i.e., applying the synthetic difference in differences approach, longer pre-treatment period, or more companies in the donor pool. However, the estimate is significantly lower (i.e., 3.7%) when we consider the outcome variable passenger trips and do not isolate the price reduction effect from the supply effects, i.e., representing the lower bound of the effect.

ACKNOWLEDGEMENTS

The authors would like to thank Helen Conradin, Michael Steinle, Silvio Sticher, Lea Oberholzer, Philipp Wegelin, and Martin Huber for their helpful comments. Moreover, we are grateful to the SBB Research Fund for its financial support.

REFERENCES

- Abadie A. 2021. Using synthetic controls: Feasibility, data requirements, and methodological Aspects. *Journal of Economic Literature*, 59(2), pp. 391-425.
- Abadie A., Diamond A., Hainmueller J. 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, 105(490), pp. 493-505.
- Abadie A., Gardeazabal J. 2003. The economic costs of conflict: A case study of the Basque Country. *American economic review*, 93(1), pp. 113-132.
- Arkhangelsky D., Athey S., Hirshberg D. A., Imbens G. W., Wager S. 2019. Synthetic difference in differences. *NBER Working Papers*.
- Axhausen K., Fröhlich P. 2012. Übersicht zu Stated Preference-Studien in der Schweiz und Abschätzungen von Gesamtelastizitäten. Report from the Federal Office for Spatial Development.
- Batty P., Palacin R., Gonzalez-Gil A. 2015. Challenges and opportunities in developing urban modal shift. *Travel Behaviour and Society*, 2(2), pp. 109-123.
- Brechan I. 2017. Effect of price reduction and increased service frequency on public transport travel. *Journal of Public Transportation*, 20(1), pp. 139-156.
- Hainmueller J., Diamond A. 2015. Package Synth. *R Package Version 1.1-5*.
- Holmgren J. 2007. Meta-analysis of public transport demand. *Transportation Research Part A: Policy and Practice*, 41(10), pp. 1021-1035.
- Huntington-Klein N. 2021. *The effect: An introduction to research design and causality*. Chapman and Hall/CRC.
- Redman L. Friman M., Gärling T., Hartig T. 2013. Quality attributes of public transport that attract car users: A research review. *Transport policy*, 25, pp. 119-127.

Silva B. C. 2020. Package SCtools: Extensions to the synthetic controls analyses. *R Package Version 0.3.1*.

Wallimann H., Blättler K., von Arx W. 2023. Do price reductions attract customers in urban public transport? A synthetic control approach. *Transportation Research Part A: Policy and Practice*, 173.

Exploring the Effect of Apps Evolution and Users' Personality on Mobile Apps Adoption and Post-Adoption Pattern Over Time: Evidence from Super-Apps Users in Indonesian Cities

Muhamad Rizki*^{1,2}, Tri Basuki Joewono³, Yusak Susilo⁴

¹ Doctoral Student, Institute for Transport Studies, University of Natural Resources and Life Sciences Vienna, Austria

² Assistant Professor, Department of Civil Engineering, Institut Teknologi Nasional Bandung, Indonesia

³ Professor, Department of Civil Engineering, Parahyangan Catholic University, Indonesia

⁴ Professor, Institute for Transport Studies, University of Natural Resources and Life Sciences Vienna, Austria

SHORT SUMMARY

This study aims to investigate the effect of users' personality as well as apps' transformation to the adoption and post-adoption patterns of multi-functional apps/Super-Apps over time. A questionnaire was distributed to Super-Apps users in four Indonesian cities. Latent Markov model (LMM) was used to investigate the users' Super-Apps usage from 2015 to 2022. Four distinctive states were recognized: the Super User (SU) state; the Transport, Consumption, and Finance (TCF) state; the Food and Beverage (F&B) and Transport (FT) state; and the Less Explored (LE) state. The analysis found that LE users have a higher probability of changing into another state of users, while TCF users tend to be more stable than other groups. A higher number of functions available does not necessarily lead to highest exploration, but it contributes to making the users evolve from LE to other states. Users who more sociable tend to explore the apps more.

Keywords: behaviour changes, users' personality, apps' transformation, latent markov model, Super-Apps

1. INTRODUCTION

Mobile apps (apps) have developed from their beginnings as simple utilities into platforms that fulfill variety of personal needs and desires. The substantial number of apps and the opportunities that they provide have attracted research on apps' adoption in various disciplines (Mehra et al., 2021; Stocchi et al., 2022). However, people's decision process to adopt apps is complex in their consumer journey phases. The phases identified as pre-adoption, which refers to the predisposition of customers before app adoption; adoption, which refers to the result of a positive predisposition towards the app and therefore, downloading and using the app; and post-adoption, which is defined as the continuation of using the app (Stocchi et al., 2022). The customer experience within the customer journey also evolves over time in response to users' characteristics, such as attitude, personality, and demography, as well as changes in the app environment, such as the level of service quality, features, or design (Kim et al., 2016). The latter has been of interest to various technological companies seeking to improve their app environment to increase the level of engagement, including the current evolution of mobile apps that provide multiple functions called Super-Apps.

Most of previous research has separated the investigation of pre-adoption (Stocchi et al., 2019), adoption (Mehra et al., 2021), and post-adoption (Kim et al., 2016). However, research has also shown that the continuation of app use is influenced by users' past behaviour (Kang et al., 2015), suggesting a pattern in mobile app usage. Attempts to integrate adoption and post-adoption by capturing behaviour over time have been done by Horvath et al. (2022) with a three-wave panel survey. Despite the study integrating the attitude of the respondents towards the apps, the study did not accommodate the effect of personality (Stocchi et al., 2022). The importance of personality has been underlined because it also drives people's motivations (Yeh et al., 2021) and consumption characteristics (such as stickiness, willingness to pay, etc.) (Dinsmore et al., 2017). Moreover, the relationship between personality and app adoption is might even more complicated in Super-Apps due to their evolution from single to multiple functions, leading to different usage patterns.

The investigation of Super-Apps use evolution is important from various perspectives. From marketing, investigating the adoption and post-adoption stages can offer strategic recommendations for apps developers aiming to optimize user engagement within apps (Kim et al., 2016; Stocchi et al., 2022). The importance of this inquiry also lies in the transportation sector, specifically urban mobility. The advancement of apps has been recognized to have a substitution, generation, modification, or neutral impact on travel demand (Mokhtarian, 2009). Some of Super-Apps functions have found related to change the travel behaviour such as ride-sourcing, food and beverage delivery, or e-shopping. Therefore, this examination is relevant to reducing negative externalities resulting from Super-Apps usage.

This study aims to explore how the evolution of apps and users' personalities contribute to the users' behaviour of using Super-Apps over time. Specifically, we focus on the adoption and post-adoption concepts of the customer journey (Stocchi et al., 2022) and extend it with temporal investigation in response to the changes in the number of functions available on the apps. Additionally, we integrate the effect of users' personalities, residential location, and socio-demography, which prior studies only considered as a factor in a specific phase (Yeh et al., 2021). Consequently, this research aims not only to comprehend the pattern of adoption and post-adoption of Super-Apps but also how this pattern evolves over time and is influenced by the apps' transformation and users' characteristics.

2. METHODOLOGY

Data Analysis

For examining the effect of the users' personality and apps' evolution to the use of Super-Apps over time, this study uses LMM. The LMM is commonly known as extension of the dynamic logit model for longitudinal data (Bartolucci et al., 2017).

For this research, we observed the use of Super-Apps services (Y_{ij}^t) as the response/indicator variable, for each service j ($j \in \{1, 2, \dots, J\}$), respondent i ($i \in \{1, 2, \dots, n\}$), and time t ($t \in \{1, 2, \dots, T\}$), by the category of 0 = not using the app and 1 = using the app. Let also x_{it} be the vector of respondents' (i) covariates for $t = 1, 2, \dots, T$. The general LMM formulation assumes the existence of latent process, denoted by U_c^t , which affects the distribution of response variables, and the process is assumed to follow a first-order Markov chain with state space ($c \in \{1, 2, \dots, k\}$). Under local independence assumption (Bartolucci et al., 2017), the response vector (Y_{ij}^t) are assumed to be conditionally independent given the latent process (U_c^t). The parameter of measurement model to determine the latent states are the conditional response probabilities:

$$\phi_{y|ux}^t = P(Y_{ij}^t = y | U_i^t = c) \quad (1)$$

The parameters of the latent process include the initial probabilities of (π_c) of each latent state and the transition probability ($\pi_{c|\zeta}$) between states:

$$\pi_{c|x} = P(U_i^1 = c | X_i^1 = x) \quad (2)$$

As the initial probability, where the $c = 1, \dots, k$, and the transition probability:

$$\pi_{c|\zeta x}^{(t)} = P(U_i^t = c | U_i^{t-1} = \zeta, X_i^t = x) \quad (3)$$

Where $t = 2, \dots, T$, $\zeta, c = 1, \dots, k$, x denotes a realization of the covariates X^t , c a realization of U^t , and ζ a realization of U^{t-1} . The initial and transition probability (latent model) are adopting multinomial logit parametrization (Bartolucci et al., 2017). In this study the latent model:

$$\log \frac{P(U_i^1 = c | X_i^1 = x)}{P(U_i^1 = 1 | X_i^1 = x)} = \log \frac{\pi_{c|x}}{\pi_{1|x}} = \beta_{0c} + x^\top \beta_{1c}, c = 2, \dots, k \quad (4)$$

As the initial probability with having first state as reference, and for transition probability:

$$\log \frac{P(U_i^t = c | U_i^{t-1} = \zeta, X_i^t = x)}{P(U_i^t = c | U_i^{t-1} = \zeta, X_i^t = x)} = \log \frac{\pi_{c|\zeta x}}{\pi_{\zeta|\zeta x}} = \gamma_{0\zeta c} + x^\top \gamma_{1\zeta c}, \quad (5)$$

where $t = 2, \dots, T$ and $\zeta, c = 1, \dots, k$, with $\zeta \neq c$. In above expression (4, 5), $\beta_c = (\beta_{0c}, \beta_{1c}^\top)^\top$ and $\gamma_{\zeta c} = (\gamma_{0\zeta c}, \gamma_{1\zeta c}^\top)^\top$ are parameter vectors to be estimated. The illustrative framework of the LMM of Super-App use, indicators and covariates is described in Figure 1.

Data Collection

Data used in this research is part of a 2022 one-week virtual and physical activity diary survey in Indonesian cities that also specifically investigates Super-App use behaviour. Grab and Gojek, launched in 2015, are the representation of transportation-based Super-apps in this study. We used the data of the respondents' socio-demography, personality traits with the Big Five Inventory (BFI), and chronology of Super-Apps use from the survey. The chronology of use section asks about the monthly frequency of each use of Super-App service (e.g., transport, goods, food transport, etc.) from 2015 to 2022 and we transformed this number into the code of whether they use it or not use it for the analysis.

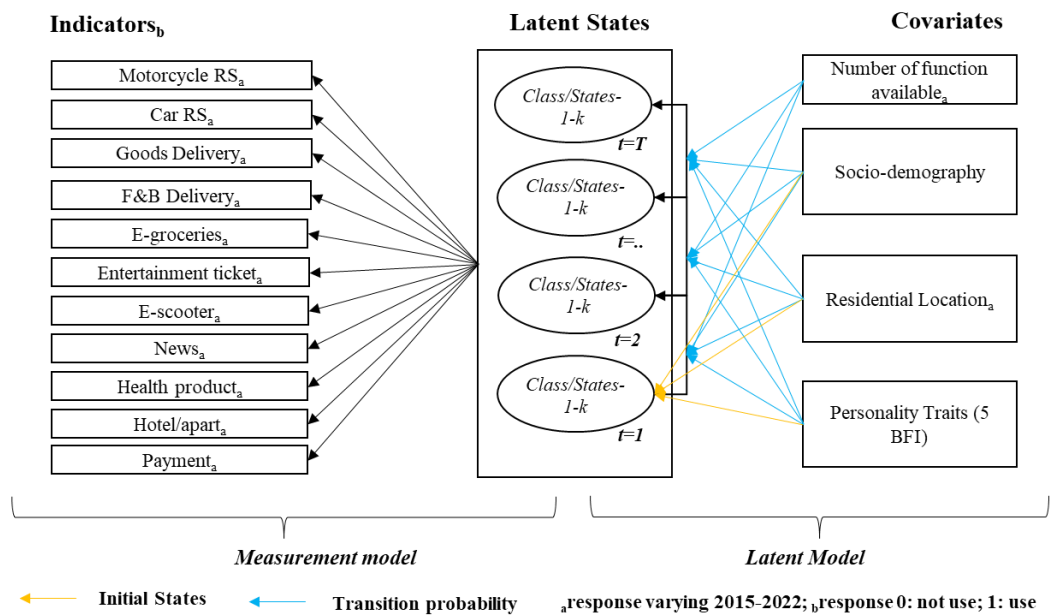


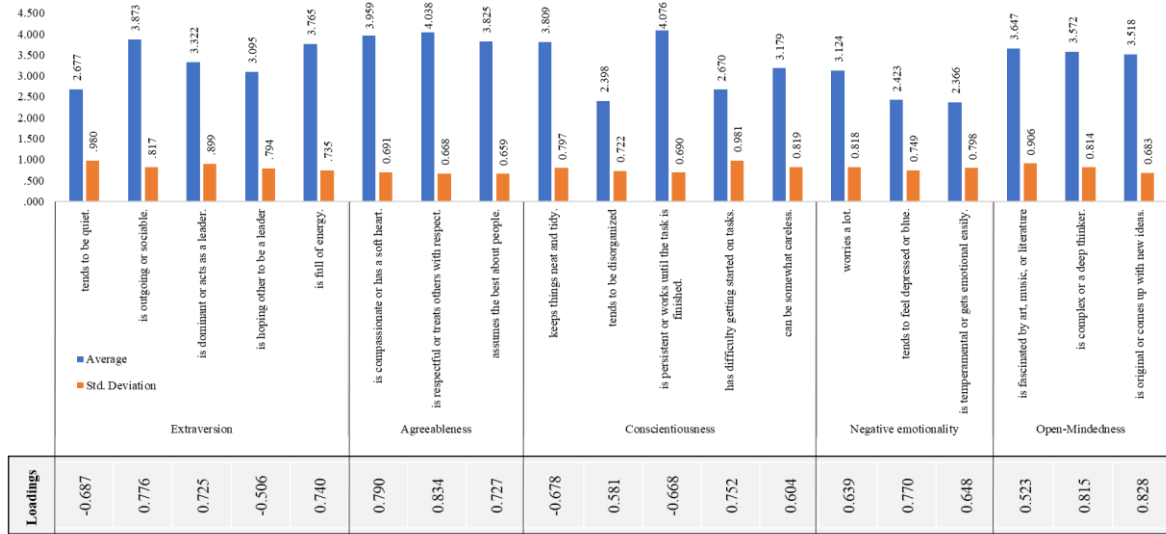
Figure 1. Conceptual Framework of Analysis

Convenience sampling method was used for data collection with surveyors distributed the questionnaire based on the small district within the city. The study location is on four Indonesian Cities that are differ in terms of population size, namely Jakarta, a megapolitan city, Bandung, a metro-politan city, Denpasar, a big city, and Cianjur, a medium city. The survey was conducted between May 2022 and January 2023. 1251 data were valid for the analysis and this study only use on 1051 datasets of Super-Apps users, excluding the respondents that not the users.

3. RESULTS AND DISCUSSION

In this study (N=1051), most of the respondents were male (56.2%), employed (70.4%), and possessed undergraduate degrees (55%). Additionally, most of them had a monthly household income of more than 6 million IDR (394 USD) (39.5%). The respondents were nearly evenly distributed across four locations based on their residential status, with Jakarta having the highest percentage (28.9%), followed by Bandung, Denpasar, and Cianjur. The study also uses secondary data about Super-Apps transformation. In 2015, only five functions were available: motorcycle and car ride-sourcing (RC), goods delivery, food and beverage (F&B) delivery, and e-groceries. In 2016, the function that offered buying entertainment tickets became available, followed by various payments (i.e., electricity, bills, stocks, donation, mobile package) in 2017. E-scooters, news/information/education, health services and products, and hotel booking became available in 2019 until now.

The respondents' personality characteristics are described in Figure 2. The variable that has the highest score is persistency in the conscientiousness dimension, while the lowest one is temperamental in negative emotionality. We used Confirmatory Factor Analysis (CFA) to reduce the personality dimension based on the five-group classification of BFI. The loading factors of each group are shown in Figure 2.



Response 1 (strongly agree) to 5 (strongly disagree); Loadings are calculated based on the CFA of the personality trait factors with Principal component analysis, and Bartlett scores for generated for each case of each factor.

Figure 2. Respondents Personality Characteristics (Question: I am person who...)

Determining the appropriate number of states is a crucial factor in LMM. Based on the recommendations of a prior research, we utilize BIC and AIC criteria to assess the states. The assessment, which is presented in Figure 3, indicates that there is no considerable improvement in AIC and BIC scores beyond four states. Additionally, the interpretability of the profile for four states is satisfactory and therefore, we decide that using four states.

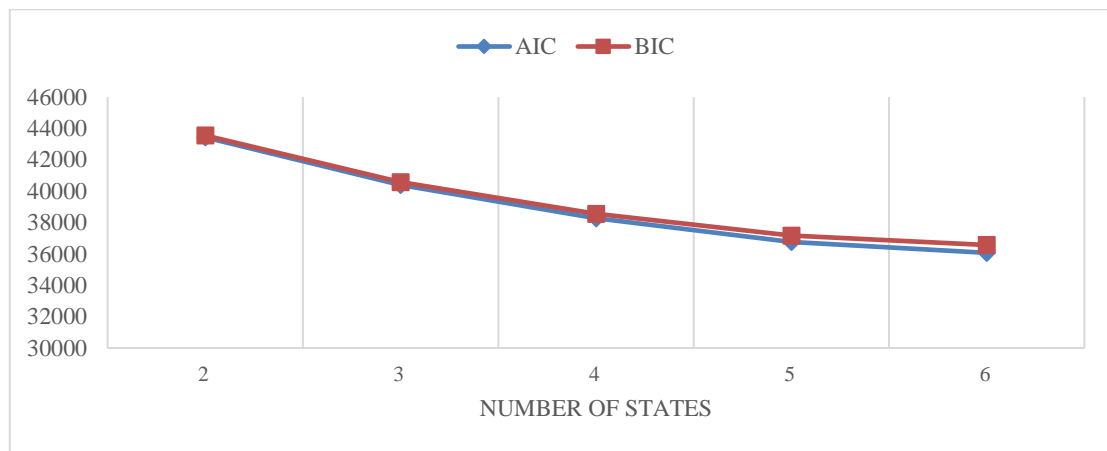


Figure 3. Determine Number of States

The profile of states based on Super-Apps function usage is described in Figure 4. The first state represents the use of motorcycle and car RC, goods delivery, F&B delivery, and payments, which we named the Transport, Consumption, and Finance (TCF) state. The second state is the Super User (SU) state, where users explore most of the Super-Apps functions. The third state represents situations where users only use a limited function, named the Low Exploration (LE) state. The last state represents the use of Super-Apps for motorcycle and car RC, goods delivery, and F&B delivery, or the F&B and Transport (FT) state.

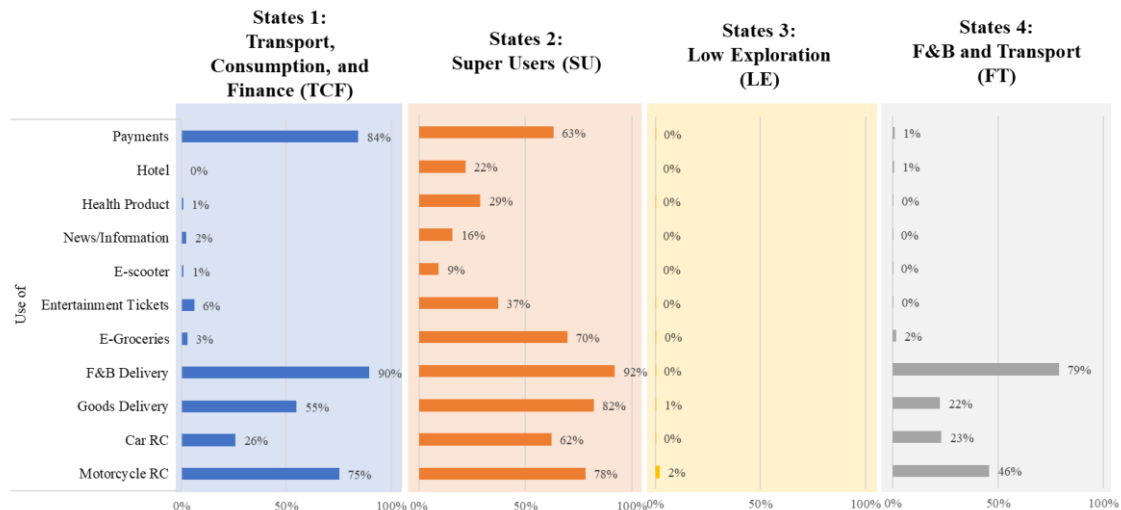


Figure 4. States Profile Based on the Functions Use

Since we are interested in patterns over time, in this study, we focus on estimating the transition probabilities (Table 1). Socio-demographic characteristics were found to influence the transition. It was found that older users are less likely to change from the LE state to another state compared to younger users. Female users are less likely to change from the TCF state to the LE or FT state, but they tend to change from LE states to FT compared to men. Workers tend to change from TCF to SU state compared to non-workers or students. Users with higher education tend to change from SU to TCF state. Differences in city characteristics were also found to shape the transition. All cities have the same negative tendency for transition from TCF to LE states. However, users from Bandung, Jakarta, and Denpasar tend to change from TCF to FT state compared to Cianjur users. Compared to Denpasar, Bandung users tend to change from TCF to SU states. Jakarta users also tend to change to FT from SU state.

The availability of more functions in the apps was found to influence the changes of states when users are in the LE state. Personality was also found to influence the behaviour of use over time. Users who are sociable, respectful, but disorganized tend to change from LE to TCF state. Users who are open-minded are less likely to lower their function use from TCF to FT or LE state. People who are respectful/kind-hearted tend to be less stable but have a positive influence on the transition. While users who are sociable tend to change from TCF to SU or FT state, they are less likely to change from SU to TCF state.

The average transition probability (Figure 4-a) reported that TCF is the most stable state, while the LE/3rd state is the least stable compared to others. Significant changes in LE states (Figure 4-b) were found from 2018 to 2019, five years after Super-Apps launched. Currently, users' states are mostly in TCF and FT.

Table 1. Estimation of Parameter Logit for Transition Probabilities

Transition from	TCF to			SU to			LE to			FT to		
	SU	LE	FT	TCF	LE	FT	TCF	SU	FT	TCF	SU	LE
Variables												
Intercept	-12.2 **	-83.8 **	10.82 **	-46.72 **	-34.87 **	-8.42 *	-10 **	-15.64 **	-7.03 **		-8.29 **	-23.66 **
Age							-0.02 *	-0.07 *	-0.03 **	0.03 *		0.11 *
Female [D]		-10.39 **	-2.4 **		11.05 *	4.11 **			0.38 **			
Household Income	-0.68 **	4.53 *								-0.14 *		
Workers [D]	15.49 **	7.42 **	1.32 **	-11.2 **			0.71 **				0.99 *	
Education		12.47 **	-5.9 **	24.17 **				1.1 *	0.62 **	-0.48 *		-2.77 **
Bandung [D]	3.15 **	-22.99 **	4.97 **	-50.63 **	-74.15 **				1.5 **		3.24 **	-8.54 *
Denpasar [D]	-15.78 **	-25.56 **	10.61 **	-49.15 **	-64.14 **			4.9 **	0.93 *		2.74 **	-9.4 *
Jakarta [D]		-36.45 **	6.24 **	-28 **	-59.62 **	5.59 **		-7.15 **	0.88 *		4.23 **	-17.82 **
Cianjur [D]		-12.93 **	-4.65 **	-31.11 **	-60.17 **			4.58 **	0.88 *		2.13 **	-9.7 *
Number of Functions Available	-0.49 **		-1.27 **		6.44 **		0.69 **	0.75 **	0.43 **	-0.31 **		
Extraversion	0.62 *		0.52 **	-14.72 **			0.46 **				0.38 *	
Agreeableness		5.57 **	1.11 **	6.03 *			0.31 **				0.76 **	
Open-mindedness		-6.53 **	-1.28 **							0.30 *		
Conscientiousness							0.26 **		-0.12 *	0.57 **		

[D]: dummy variables; *significant at 5%; **significant at 1%

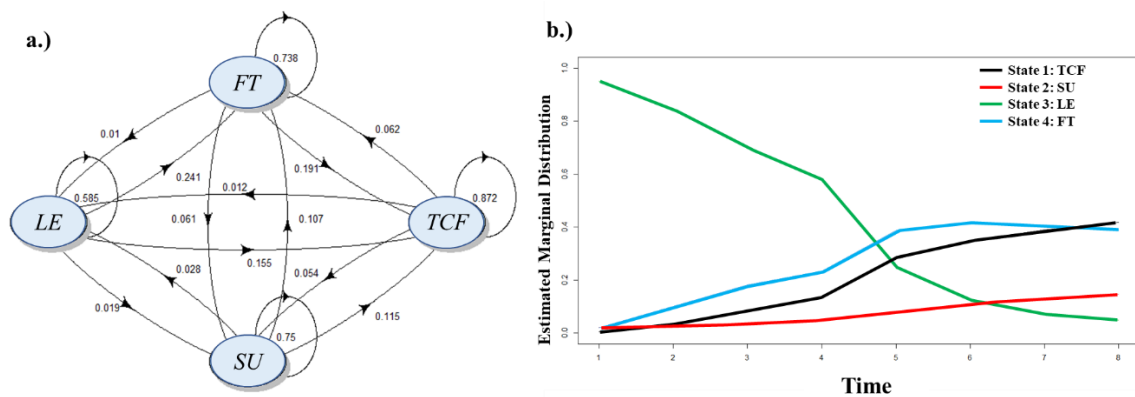


Figure 4. a.) Average Transition Probability, b.) Marginal Distribution Over Time

4. CONCLUSIONS

This study explores the effects of the transformation of apps and users' personality, residential location, and personal characteristics on the pattern of Super-Apps use over time. The study found four states that represent specific use of Super-Apps' functions: the Transport, Consumption, and Finance (TCF) state, Super users (SU) state, Low exploration (LE) state, and F&B and Transport (FT) state. The study also found that adding more functions to the apps influences the adoption of Super-Apps, leading to a transition from LE to other states but not necessarily to the highest exploration (SU state). Personality also influences the transition, with sociable users tending to explore the apps more, suggesting that Super-Apps might facilitate their outgoing personality needs. The economic intensity of a city does not influence greater use exploration, as every city has its own transition into the SU state. However, Jakarta stands out from other cities in its tendency to change from the SU to the FT state.

ACKNOWLEDGEMENTS

This study is funded by the Research Grant from ITENAS Bandung, the WCR Grant from The Ministry of Research, Technology, and Higher Education, the Republic of Indonesia, and the DAVeMoS BMK Endowed Professorship in Digitalisation and Automation in Transport Systems (FFG project number: 862678). The first author also acknowledged the Ernst Mach Grant ASEA-UNINET for the doctoral scholarship. We are also thankful for the insightful comments from Giovanni Circella in the earliest draft of this paper.

REFERENCES

Bartolucci, F., Pandolfi, S., & Pennoni, F. (2017). LMest: An R Package for Latent Markov Models for Longitudinal Categorical Data. *Journal of Statistical Software*, 81, 1–38. <https://doi.org/10.18637/jss.v081.i04>

Dinsmore, J. B., Swani, K., & Dugan, R. G. (2017). To “Free” or Not to “Free”: Trait Predictors of Mobile App Purchasing Tendencies. *Psychology & Marketing*, 34(2), 227–244. <https://doi.org/10.1002/mar.20985>

- Horvath, L., Banducci, S., Blamire, J., Degnen, C., James, O., Jones, A., Stevens, D., & Tyler, K. (2022). Adoption and continued use of mobile contact tracing technology: Multilevel explanations from a three-wave panel survey and linked data. *BMJ Open*, *12*(1), e053327. <https://doi.org/10.1136/bmjopen-2021-053327>
- Kang, J.-Y. M., Mun, J. M., & Johnson, K. K. P. (2015). In-store mobile usage: Downloading and usage intention toward mobile location-based retail apps. *Computers in Human Behaviour*, *46*, 210–217. <https://doi.org/10.1016/j.chb.2015.01.012>
- Kim, S., Baek, T. H., Kim, Y.-K., & Yoo, K. (2016). Factors affecting stickiness and word of mouth in mobile applications. *Journal of Research in Interactive Marketing*, *10*(3), 177–192. <https://doi.org/10.1108/JRIM-06-2015-0046>
- Mehra, A., Paul, J., & Kaurav, R. P. S. (2021). Determinants of mobile apps adoption among young adults: Theoretical extension and analysis. *Journal of Marketing Communications*, *27*(5), 481–509. <https://doi.org/10.1080/13527266.2020.1725780>
- Mokhtarian, P. L. (2009). If telecommunication is such a good substitute for travel, why does congestion continue to get worse?: *Transportation Letters*, *Vol 1*(No 1), 1–17.
- Stocchi, L., Michaelidou, N., & Micevski, M. (2019). Drivers and outcomes of branded mobile app usage intention. *Journal of Product & Brand Management*, *28*(1), 28–49. <https://doi.org/10.1108/JPBM-02-2017-1436>
- Stocchi, L., Pourazad, N., Michaelidou, N., Tanusondjaja, A., & Harrigan, P. (2022). Marketing research on Mobile apps: Past, present and future. *Journal of the Academy of Marketing Science*, *50*(2), 195–225. <https://doi.org/10.1007/s11747-021-00815-w>
- To, P.-L., Liao, C., & Lin, T.-H. (2007). Shopping motivations on Internet: A study based on utilitarian and hedonic value. *Technovation*, *27*(12), 774–787. <https://doi.org/10.1016/j.technovation.2007.01.001>
- Yeh, C.-H., Wang, Y.-S., Wang, Y.-M., & Liao, T.-J. (2021). Drivers of mobile learning app usage: An integrated perspective of personality, readiness, and motivation. *Interactive Learning Environments*, *0*(0), 1–18. <https://doi.org/10.1080/10494820.2021.1937658>

Routing Passengers while Timetabling Based on Promises from Line Planning: A Logic-Based Benders Approach

Fuchs Florian^{*1}, Viera Klasovítá², and Francesco Corman³

¹MSc, Institute for Transport Planning and Systems, ETH Zürich, Switzerland

²MSc, Institute for Transport Planning and Systems, ETH Zürich, Switzerland

³Professor Dr., Institute for Transport Planning and Systems, ETH Zürich, Switzerland

SHORT SUMMARY

Effective line planning and timetabling are critical for enhancing public transport efficiency and passenger satisfaction. We propose a Logic-Based Benders decomposition approach to optimise a timetable for a passenger railway system based on the promises made in earlier planning stages. Our approach ensures that the promised travel times and transfers are available and passenger routes are chosen according to the shortest available path. We test this approach on real-world data from the Rhätische Bahn railway system, demonstrating promising results. The proposed approach has shown to be valuable for optimising transfers, improving efficiency and passenger satisfaction, and reducing travel times. The method has limitations, including the inability to consider multiple connections per origin-destination pair, adaptation time at the origin station, and crowding. Further research can focus on improving and extending the model's performance to include these factors.

Keywords: Logic-Based Benders Decomposition, Integrated Public Transport Planning, Railway Timetabling, Periodic Railway Timetabling.

1 INTRODUCTION

Railways play a vital role in the transport system; however, public transport planning is complex and includes several stages. From these stages, line planning and timetabling are two essential steps in planning public transport systems (Schiewe, 2020). Line planning involves selecting the routes and services to operate on each route while timetabling involves assigning the trips to specific times. While we assume that transfers are possible in the line planning, it might be possible that such a transfer is then not given in the subsequent timetable. This can result in passengers being unable to make their intended connections, leading to longer travel times and reduced passenger satisfaction. Combining the planning stages of line planning and timetabling is a promising step to overcome this issue. It enables finding a schedule that minimises total travel time while ensuring that all passengers have a connection and that the following timetable is feasible. However, the timetabling problem is already complicated, and combining it with line planning further increases the complexity.

Various approaches have been proposed to integrate line planning and timetabling, including integrated optimisation methods (Schiewe, 2020) and combining passenger routing with periodic timetabling (Schmidt & Schöbel, 2015; Borndörfer et al., 2017; Robenek et al., 2016). However, these methods can be computationally intensive and may need improvements to scale better for more extensive networks. In response to these challenges, Polinder et al. (2021) proposed a scalable timetabling approach that considers the passenger perspective and aims to determine a high-quality timetable outline in the strategic planning phase. In addition, the method includes adaptation time, i.e., the waiting time at the origin station, to ensure regular connections between passengers' origins and destinations.

In conclusion, effective line planning and timetabling are critical for enhancing public transport efficiency and passenger satisfaction. While the line planning problem and periodic timetabling have established methods, recent research on timetabling with passenger routing looks promising. Integrated optimisation and passenger routing approaches have shown promising results but require significant computational resources. Therefore, developing efficient and scalable solutions for integrated planning and timetabling is crucial for public transport planning.

Our approach aims to address this gap by creating an optimal timetable for a passenger railway exploiting the fact that the timetable usually stems from a line plan (Bull et al., 2019). Proposed approaches to finding optimal line plans typically minimise total travel time by assuming specific in-vehicle and transfer times. We interpret these assumptions in the line plan as promises made to passengers and use them to formulate an optimisation model. Our model ensures that the promised travel times and transfers are available and that passenger routes are chosen according to the shortest available path. We aim to provide an optimal timetable for passengers by keeping these promises and optimising travel efficiency.

This leads to the following contributions of this work:

1. We propose a logic-based Benders decomposition approach that optimises a timetable for a passenger railway system based on the promises made in earlier planning stages.
2. We test this approach on real-world data from the Rhätische Bahn railway system, demonstrating promising results.

Developing efficient and scalable solutions for integrated line planning and timetabling is crucial for enhancing public transport efficiency, reducing travel times, and improving passenger satisfaction.

2 METHODOLOGY

This section summarises our mixed integer linear programming model before we outline the Logic-Based Benders Decomposition approach.

Baseline model

We propose routing passengers and scheduling trains on a network using mixed integer linear programming (MILP) to minimise travel time while adhering to established schedules. Time windows are added for each commercial train arrival and departure to prevent schedule violations. Passenger expectations are met by enforcing pre-established in-vehicle travel and transfer times based on the assumptions taken during line planning. Time windows also aid in identifying optimal transfer connections. While we aim for all transfers to occur within the time limit, it may not always be possible. We allow multiple time limits per transfer to address this, which relaxes the transfer constraint. Our formulation seeks to optimise connections for all passengers within the constraints to minimise the total travel time.

To model the demand routing through the network, we use the Passenger Graph (PG), a directed graph representing the transportation network and consisting of *trip/dwell/transfer* arcs. We add all *trip/dwell* arcs to the PG for each train specified by the line plan. We then connect the arrival/departure nodes at stations according to the connections given in the line plan and add transfer arcs for each potential *transfer* between two connected stations. We add multiple *transfer* arcs between the exact arrival and departure nodes if a *transfer* has multiple duration limits.

To integrate the timetabling part into the transfer scheduling problem, we use the periodic event scheduling problem (PESP) formulation Serafini & Ukovich (1989), which is suitable for periodic timetabling. Furthermore, we use the Cycle Periodic Formulation (CPF) of Peeters (2003) for the PESP.

Logic-Based Benders decomposition

The MILP is challenging due to the computational complexity of the timetabling part, as shown by (Borndörfer et al., 2020). To address this issue, we propose a Logic-Based Benders Decomposition approach that separates the routing and timetabling problems to improve solution efficiency (Hooker, 2007). Specifically, we use SAT-based approach for the timetabling subproblem, adapted from Großmann (2016), allowing for faster feasibility determination and the generation of combinatorial Benders cuts from unsatisfiable cores (Codato & Fischetti, 2006). Classical Benders Decomposition is a two-stage method that optimises the master problem and checks the feasibility of the subproblem. In contrast, the Branch-and-Cut Benders method combines branch-and-cut algorithms with Benders Decomposition to reduce computational time (Rahmaniani et al., 2017). We implement each method and a hybrid approach that uses both, allowing for information exchange in the form of feasible solutions to improve efficiency further. We employ two strategies to improve the cut generation process, relaxing requirements and expanding conflicting transfer cuts.

3 RESULTS AND DISCUSSION

We evaluate the performance and suitability of our decompositions using real-world data from the Swiss railway company, RhB. We compare four different approaches to solving the problem: **Baseline** (MILP implementation), **Classical** (Classic Logic-Based Benders decomposition), **Modern** (Logic-Based Branch and Benders Cut), and **Hybrid** (**Modern** and **Classical** parallel). We use an instance of the RhB network with today’s line plan of ten lines and a real-life dataset for demand, which includes 1747 origin/destination locations. The line plan is depicted in Figure 1.

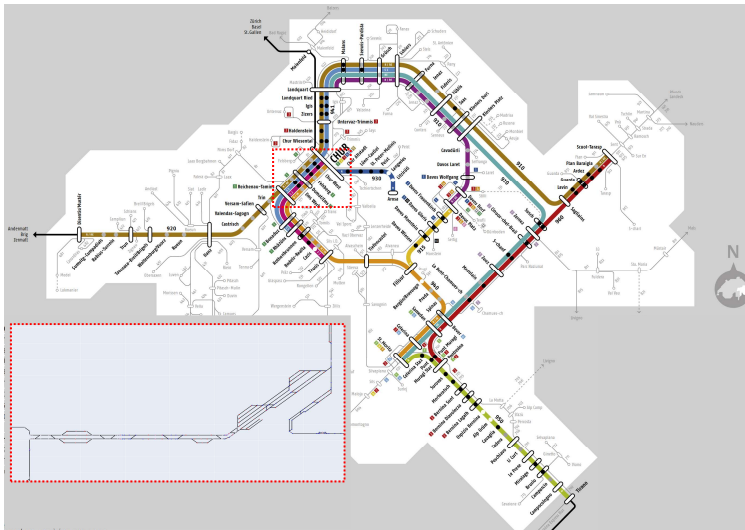


Figure 1: Current line plan of RhB (RhB, 2023).

We run each model with eight different random number seeds in the MILP solver. We report the best solution found in the form of total travel time and the remaining optimality gap expressed as a percentage. We are implementing the approach using Python 3.10 and the OpenBus Toolbox (Fuchs & Corman, 2019) and solving the MILP formulation using GUROBI 10.0.1 (Gurobi Optimization, LLC, 2023). In addition, we are solving the sat instances using the Glucose 4.1 sat solver (Audemard & Simon, 2018) with pySAT (Ignatiev et al., 2018). We allow GUROBI to use up to four threads for each run, while Glucose is single-threaded.

Table 1: Results for a time of three hours. A cell shows the objective (total travel time) and optimality gap.

<i>Seed</i>	Baseline	Hybrid	Modern	Classic
0	-, -	323.62 h, 38.76 %	323.54 h, 91.57 %	323.58 h, 37.09 %
1	-, -	323.55 h, 39.55 %	323.53 h, 79.74 %	324.24 h, 27.34 %
2	-, -	323.58 h, 38.38 %	323.55 h, 89.16 %	324.23 h, 19.57 %
3	-, -	324.23 h, 27.80 %	323.53 h, 79.13 %	324.23 h, 17.84 %
4	-, -	324.24 h, 27.81 %	323.54 h, 75.70 %	324.23 h, 20.64 %
5	-, -	323.56 h, 38.77 %	323.54 h, 90.21 %	323.57 h, 38.58 %
6	-, -	323.53 h, 38.61 %	323.54 h, 82.15 %	323.57 h, 39.09 %
7	-, -	323.57 h, 38.63 %	323.56 h, 76.57 %	324.23 h, 17.40 %

The results presented in Table 1 demonstrate that our proposed decomposition approach outperforms the **Baseline** approach. Notably, all runs of the **Baseline** implementation failed to yield any results, in stark contrast to all the decomposition approaches. Our comparison of the **Classical**, **Modern**, and **Hybrid** approaches revealed that they all produce similar objective function results. However, we observe differences in their ability to close the optimality gap, which remains a challenging problem. In particular, the **Classic** approach exhibited superior performance on a one-hour time limit, while on an extended three-hour time limit, **Classic** and **Hybrid** approaches performed comparably.

4 CONCLUSIONS

The proposed approach is valuable for optimising transfers, improving efficiency and passenger satisfaction, and reducing travel times. The **Classic** approach outperformed the other two decomposition approaches due to its ability to fix infeasible instances using a heuristic. However, the optimality gap remains a challenge. Solving the problem as a MILP, as in the Baseline approach, is not viable. Furthermore, the proposed method has limitations, including the inability to consider multiple connections per origin-destination pair, adaptation time at the origin station, and crowding. Further research can focus on improving and extending the model’s performance to include these factors. Overall, the proposed approach has shown promising results on real-world data from Rhätische Bahn and can be applied to other transportation systems.

ACKNOWLEDGEMENTS

We want to thank RhB for supplying us with data and feedback for our research.

REFERENCES

- Audemard, G., & Simon, L. (2018, 02). On the glucose sat solver. *International Journal on Artificial Intelligence Tools*, 27, 1840001.
- Borndörfer, R., Hoppmann, H., & Karbstein, M. (2017, 7). Passenger routing for periodic timetable optimization. In *Public transport* (Vol. 9, pp. 115–135).
- Borndörfer, R., Lindner, N., & Roth, S. (2020). A concurrent approach to the periodic event scheduling problem. *Journal of Rail Transport Planning and Management*, 15, 140–159.
- Bull, S., Larsen, J., Lusby, R. M., & Rezanova, N. J. (2019). Optimising the travel time of a line plan. *For*, 17(3), 225–259.
- Codato, G., & Fischetti, M. (2006). Combinatorial benders’ cuts for mixed-integer linear programming. *Operations Research*, 54(4), 756–766.
- Fuchs, F., & Corman, F. (2019). An open toolbox for integrated optimization of public transport. In *2019 6th international conference on models and technologies for intelligent transportation systems (mt-its)* (p. 1-7).
- Großmann, P. (2016). *Satisfiability and Optimization in Periodic Traffic Flow Problems* (Unpublished doctoral dissertation). Technische Universität Dresden.
- Gurobi Optimization, LLC. (2023). *Gurobi Optimizer Reference Manual*. Retrieved from <https://www.gurobi.com>
- Hooker, J. N. (2007). Planning and scheduling by logic-based benders decomposition. *Operations research*, 55(3), 588–602.
- Ignatiev, A., Morgado, A., & Marques-Silva, J. (2018). PySAT: A Python toolkit for prototyping with SAT oracles. In *Sat* (pp. 428–437).
- Peeters, L. W. P. (2003). *Cyclic railway timetable optimization* (Vol. 22). Rotterdam: Erasmus University Rotterdam.
- Polinder, G. J., Schmidt, M., & Huisman, D. (2021). Timetabling for strategic passenger railway planning. *Transportation Research Part B: Methodological*, 146, 111–135.
- Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3), 801-817.
- RhB. (2023). Streckennetz [Computer software manual]. Retrieved 2023-02-04, from <https://www.rhb.ch/de/service-souvenirs/streckennetz>
- Robenek, T., Maknoon, Y., Azadeh, S. S., Chen, J., & Bierlaire, M. (2016). Passenger centric train timetabling problem. *Transportation Research Part B: Methodological*, 89, 107–126.

- Schiewe, P. (2020). *Integrated optimization in public transport planning* (Vol. 160). Springer.
- Schmidt, M., & Schöbel, A. (2015). Timetabling with passenger routing. *OR Spectrum*, 37(1), 75–97.
- Serafini, P., & Ukovich, W. (1989). A Mathematical Model for Periodic Scheduling Problems. *SIAM Journal on Discrete Mathematics*, 2(4), 550–581.

Quantification of non-linear effects in agglomeration economies for transport appraisals

Anupriya¹, Daniel J. Graham¹, and Prateek Bansal²

¹Transport Strategy Centre, Department of Civil & Environmental Engineering, Imperial College London, London, SW7 2AZ, UK.

²Department of Civil & Environmental Engineering, National University of Singapore.

SHORT SUMMARY

Agglomeration economies arising from the spatial concentration of economic activity have been known to exist and induce higher productivity for firms. The existing empirical evidence, however, has two key caveats. First, it mostly assumes a pre-specified (mostly log-log) functional form for the relationship between firm productivity and agglomeration. Second, it may lack valid instruments to adjust for potential confounding biases (for instance, from the omission of characteristics of local input and output markets) in the estimation of this relationship. This study adopts a flexible Bayesian Non-Parametric Instrumental Variables based approach to quantify non-linear effects in agglomeration economies. The approach uses innovative external instruments derived from traffic casualty data. We adopt a two-step framework: we first isolate the firm's total factor productivity from a Cobb-Douglas production function and thereafter estimate the non-linear effects of agglomeration on this productivity. Using data from a sample of firms classed into six key industry sectors in England, we present novel evidence that indicates the presence of significant non-linearities in agglomeration elasticities for most industry sectors. Our results provide critical inputs for the appraisal of transport investments.

Keywords: Agglomeration; Cost-benefit Analysis; Wider Economic Impacts; Elasticity; Non-parametric statistics; Bayesian machine learning.

1 INTRODUCTION

Transport investments are frequently aimed at bringing economic and social benefits to the economy. However, given their scale, it is important for policymakers to have a rigorous apriori understanding of the magnitude of potential benefits arising from the investment. Cost Benefit Analysis (CBA) provides a well-established theoretical basis to measure these benefits ex-ante and thus warrant such extensive investments (Graham & Gibbons, 2019; A. Venables et al., 2014; Mackie et al., 2012). CBA recasts the costs and benefits of the investment in monetary terms to estimate the net change in social welfare arising from the transport improvement.

The current appraisal process embraced in the UK's Transport Analysis Guidance (TAG) recognises two broad categories of such welfare impacts: (1) Direct user benefits (DUBs) and (2) Wider economic impacts (WEIs). DUBs consist of impacts on existing and new users of the transport system, generated via changes in the generalised cost of travel (say, via alterations in travel time or quality of service). Economic theory suggests that DUBs can capture all impacts under idealised economic conditions: under perfect competition, constant returns to scale, and in the absence of market failures. DUBs, therefore, formed the crux of the calculations in conventional CBA. Nevertheless, the above-described market conditions are seldom encountered in practice, thus, undermining the ability of DUBs to capture economic impacts in exhaustive detail. The scope of CBA has therefore been extended, primarily in the past decade, to include the economic impacts caused by market imperfections and externalities. Such *wider* group of impacts on the economy are manifested in WEIs. The overarching aim of this paper is to revisit the empirical evidence on WEIs of transport investments emerging via scale economies of agglomeration.

Background

Agglomeration economies are spatial externalities that arise when economic agents (individuals and firms) locate in close proximity to each other, or in other words, locate within agglomerations

of economic mass. Such proximity facilitates greater sharing, matching, and learning interactions between agents, which are known to be the key drivers of agglomeration (Duranton & Puga, 2004). For firms, agglomeration translates into benefits in form of improved labour market interactions, knowledge spill-overs, specialisation, and increased sharing of inputs and outputs (Marshall, 1920). Given these underlying benefits, economic theory predicts a positive impact of agglomeration on the productivity of firms. Theory suggests two forms of agglomeration economies: (i) economies of industry concentration or localisation economies, which includes benefits that occur through enhanced specialisation; and, (ii) economies of urban concentration or urbanisation economies, which result from the scale and diversity of markets (DfT, 2016). The former effects are external to the firm but internal to the industry, while the latter are external to the firm and the industry but internal to the urban area (or city) (Graham & Gibbons, 2019). Firms mostly experience the two forms of agglomeration economies simultaneously. It is, therefore, difficult to disentangle the two effects. The paper focuses on the quantification of urbanisation economies, which are typically considered in appraisal calculations.

The impact of transport investments on agglomeration economies emerge from transport's effect on economic geography (distribution of individuals and firms) and consequent proximity between economic agents (A. J. Venables, 2007). For fixed geography, transport improvements reduce the generalised costs of interaction between agents in a cluster, thereby assisting interactions and changing the effective density of the cluster. These effects are referred to as static agglomeration effects. Transport investments can also bring about changes in economic geography by making some locations more attractive to live and work, resulting in the relocation of individuals and firms. Such movements lead to changes in the physical density of the cluster and facilitate new and different interactions (DfT, 2016). These effects are termed dynamic agglomeration effects. Static effects generally appear as a subset of the overall dynamic effects of a transport provision. Interestingly, dynamic agglomeration effects may have a positive or negative impact on the productivity of firms in a cluster depending on whether the spatial density of the cluster increases or decreases as a consequence of the transport investment. In summary, the above discussion suggests that if increasing effective agglomeration results in productivity gains for firms, and if transport has an underlying role in determining effective agglomeration, then transport investments may induce productivity benefits or disbenefits for firms. Therefore, comprehension of these productivity effects is critical to developing a meticulous assessment of the impact of transport investments.

Consistent with the theory of urban agglomeration, the weight of empirical evidence in the literature supports a significant positive relationship between agglomeration (often represented by city size or the degree of access to economic mass (ATEM)) and productivity (measured by wages or by Total Factor Productivity (TFP)) (see, for instance, Lall et al., 2004; Au & Henderson, 2006; Rosenthal & Strange, 2008; Baldwin et al., 2010; Graham & Dender, 2011; Combes et al., 2012; Morikawa, 2011; Maré & Graham, 2013; Marrocu et al., 2013; Ahlfeldt et al., 2015). Elasticity estimates from the literature range between -0.800 and 0.658, with an unweighted mean of 0.046 and median value of 0.043 (refer to Graham & Gibbons, 2019, for a detailed review). The literature thus indicates that agglomeration economies exist and cause higher productivity for firms and workers. Nevertheless, the estimated magnitude of this effect varies substantially across studies. A meta-analysis of the empirical literature conducted by Melo et al. (2009) indicates that such variation results from contextual factors associated with the study design such as the nature of economies and urban systems and the type of industry sectors under study. Additionally, Graham & Gibbons (2019) highlights that the differences in estimated elasticities also result from the differences in methodological approaches adopted in previous studies. While most studies concurrently assume a log-log relationship between agglomeration and productivity, the approaches adopted to identify this relationship vary substantially. In particular, Graham & Gibbons (2019) note that there are considerable discrepancies in the extent to which these studies attempt to correct for potential confounding biases in the estimation of this relationship.

Relatedly, Graham & Gibbons (2019) identify six potential mechanisms via which such biases may emerge. First, confounding may occur due to the presence of unobserved firm-level sources of productivity that are not only crucial to the firm's choice of inputs, and thereby its TFP (see Van Beveren, 2012, for details), but may be determined by local technology factors such as agglomeration. Second, confounding may also occur due to the absence of knowledge on a firm's market exit decisions (see Akerberg et al., 2006, for details), which may be determined by agglomeration. In particular, firms located in clusters of higher agglomeration may experience more competition, which could result in the exit of less productive firms from the market. Third, confounding biases may emerge via unobserved heterogeneity in output prices of firms, which have a systematic correlation with market competition, and thereby with agglomeration. Fourth, confounding may appear due to spatial sorting or self-selection of firms, which occurs when firms within the same

industry derive unobserved productivity benefits by engaging in different activities across different locations. Such unobserved heterogeneity is often correlated with the level of agglomeration. Fifth, the relationship between agglomeration and productivity may be simultaneously determined. As shown by Graham et al. (2010), higher productivity locations may attract more private investment over time leading to larger agglomeration and a consequent increase in productivity. Failure to account for this reverse causality between productivity and agglomeration may produce biased and inconsistent estimates of agglomeration economies. Finally, additional confounding may emerge from unobserved components of local technology, such as specific characteristics of local input and output markets, that may be determine both agglomeration and productivity. One of the key objectives of this research is to deliver agglomeration elasticities that are robust to confounding, and as such suitable to inform the appraisal of transport investments.

Overview of the Analysis

This paper estimates the relationship between agglomeration and productivity by adopting a two-step approach Combes & Gobillon (2015). The first stage model within this approach estimates TFP from the production function. The predicted values of TFP are then used as the dependent variable in a second-stage regression on agglomeration, which delivers the agglomeration elasticities. However, contrary to previous studies, we exploit the ability of this approach to model a flexible non-parametric relationship between agglomeration and productivity. In particular, we note that existing models in the literature mostly presume a Cobb-Douglas model for the agglomeration-productivity relationship. We argue that while economic theory suggests a positive impact of agglomeration on productivity, it does not necessarily imply that the relationship should be log-log. Furthermore, contextualisation of agglomeration elasticities across studies (as discussed in Section 1) may be indicative of the variation of these elasticities over agglomeration. Hence, we assert that parametric models with a predefined functional form may fail to capture the non-linearities in the agglomeration-productivity relationship, thus delivering estimates of agglomeration elasticities that may be biased.

To address this limitation, in this paper, we empirically estimate the relationship between agglomeration and productivity using a Bayesian non-parametric instrumental variables (NPIV) estimator (Wiesenfarth et al., 2014) that allows us to (i) capture non-linearities in the relationship with a non-parametric (NP) specification that does not require an assumed a-priori functional form, and (ii) adjust for any confounding bias using instrumental variables (IVs). As discussed in Section 1, such biases may emerge from reverse causality or from the omission of important covariates. Critical to the second point, we note that the literature may lack valid IVs for agglomeration, which could hinder the identification of the agglomeration-productivity relationship. To overcome this limitation, we recognise a novel external IV for agglomeration that is derived from traffic casualty records. In particular, we consider the severity of traffic accidents among active mode and motorcycle users during peak hours in a given location and time period as a relevant and exogenous instrument for agglomeration in that location and time.

We apply the proposed approach to a sample of firms in England, divided into six key industrial sectors: Manufacturing; Construction; Wholesale and Distribution; Transport; Information and Communication Technology; and Finance. To measure the TFP of these firms, we make use of an exhaustive panel dataset recorded by the Department of Trade and Industry. The data relate to annualised accounting information provided by all companies registered in the UK. For the purpose of this study, we consider the data between the years 2015 and 2019. Relevant measures of agglomeration for this period are formed using employment records maintained by the Office for National Statistics.

Contributions

The major contributions of this study can be summarised as follows:

1. We derive a novel external instrument from traffic accident data to identify the relationship between agglomeration and firm productivity.
2. Our study delivers a novel comprehension of the non-linearities in agglomeration elasticities across key industry sectors in England. For instance, we find the finance sector to be associated with agglomeration diseconomies at lower levels of agglomeration and agglomeration economies at higher levels of agglomeration. This result indicates the presence of a critical economic mass beyond which productivity benefits set in for firms in this industry sector.

Interestingly, across all sectors, we note that the agglomeration elasticities take more extreme values than those from a linear model. For instance, while the linear model suggests agglomeration economies of magnitude 0.144 for firms in the finance industry, our non-linear model finds the elasticity estimates to go up to a level of 0.710.

3. Quantification of non-linearities in agglomeration economies facilitates a novel understanding of the spatial distribution of the productivity benefits of agglomeration across England. Such a mapping could be instrumental in identifying the potential gainers and losers of productivity benefits arising from a transport improvement.

The rest of this paper is structured as follows. Section 2 describes the data and the methodology used to estimate the agglomeration elasticities. Section 3 presents the results from the empirical study. Conclusions are drawn in the final section.

2 MODEL AND DATA

This section is divided into four subsections. The first subsection elaborates on the measure of agglomeration used in this study. The second subsection discusses the adopted two-step approach to estimate the relationship between agglomeration and productivity. The third subsection briefly describes the Bayesian NPIV method in the context of this study. The final subsection discusses the data used in this analysis.

Measure of Agglomeration

A crucial prerequisite to understanding the WEIs that arise from agglomeration is to develop a suitable measure of agglomeration for each location or geographical zone. In line with the current CBA practice in the UK, we represent agglomeration using ATEM, or in other words, the Mean Effective Density (MED). The MED ρ_j for zone, $j, j = (1, \dots, n)$, is calculated as follows:

$$\rho_j = \frac{1}{n} \sum_{i=1}^n m_i f(d_{ij})$$

where m_j represents a measure of economic activity in each zone j and $f(\cdot)$ denotes the deterrence function, which is a decreasing function of the cost of travelling from origin j to destination k . The measure is designed to capture the effects of the geographic centrality of the zones, their size distribution, and the spatial distribution of economic mass (Graham & Gibbons, 2019). We consider the zonal employment level E_{jt} as the measure of the economic activity of zone j and year t and the inverse Euclidean distance between the centroids of each zone d_{jk}^α for the construction of the deterrence function, where α is the distance decay parameter, generally assumed to take a value of 1.0. The resulting MED for zone j in year t is thus:

$$\rho_{jt} = \frac{1}{n} \sum_{k=1}^n \frac{E_{kt}}{d_{jk}^\alpha}. \quad (1)$$

Agglomeration and Productivity

As mentioned in the Introduction, to quantify the impact of urban agglomeration on productivity, we adopt a two-step approach. The first step involves estimating the TFP of a firm by constructing its production function. The second step comprises regressing the estimated TFP values on the chosen measure of agglomeration (that is, MED) to derive the agglomeration elasticities δ^ρ , given by

$$\delta^\rho = \frac{\partial \log \omega}{\partial \log \rho} \quad (2)$$

We emphasise that we choose TFP over labour productivity measures such as wage rate as the latter has the following disadvantages in the context of appraising WEIs. First, they can be determined by transport improvements via routes other than productivity (for instance, via shifts in labour supply). Second, while TFP is exclusively determined by local technology, output prices and average labour skills; wages are additionally influenced by the relative prices of other factors such as land and housing prices. Such dependability can introduce severe confounding biases in

the estimation of agglomeration elasticities (Combes & Gobillon, 2015). Third, wage-based measures carry the assumption that the wage equals the value of the marginal product in competitive equilibrium. However, the equality assumption seldom holds in practice as wages are typically proportional to labour productivity (Combes & Gobillon, 2015). Lastly, wage-based measures only provide a partial representation of productivity as they are limited to the impacts on the labour input alone. Conversely, TFP provides a more comprehensive measurement of productivity with respect to all inputs, which in the agglomeration context is more critical as agglomeration may affect technology in several ways (Maré & Graham, 2009).

Step 1: Estimating total factor productivity

Consistent with the literature, we assume that the production of outputs Y_{it}^s by a firm i in industry sector s in year t to follow a Cobb Douglas production function structure with inputs; capital K_{it}^s , labour L_{it}^s and materials M_{it}^s ; as covariates:

$$\log Y_{it}^s = \beta_k^s \log K_{it}^s + \beta_l^s \log L_{it}^s + \beta_m^s \log M_{it}^s + \omega_{it}^s + \gamma_t^s + e_{it}^s \quad (3)$$

where β_k^s , β_l^s and β_m^s are constants representing the elasticities of output with respect to the associated factor of production. ω_{it}^s is the unobserved efficiency or productivity of the firm, commonly referred to as its Total Factor Productivity (TFP). TFP represents the efficiency level that remains unobserved by the analyst, but is known to (or predicted by) the firm. γ_t^s are year dummies that capture the year-specific effects on productivity and inflation. e_{it}^s is a normally distributed idiosyncratic error term, or in other words, all random shocks to the outputs. From equation 3, the firm's TFP ω_{it}^s can be estimated as follows:

$$\hat{\omega}_{it}^s = \log Y_{it}^s - \hat{\beta}_k^s \log K_{it}^s - \hat{\beta}_l^s \log L_{it}^s - \hat{\beta}_m^s \log M_{it}^s - \hat{\gamma}_t^s. \quad (4)$$

Note that TFP affects the firm's choice of input factors and market exit decisions, thus rendering the variable factors of production, labour, and materials, endogenous in the model (De Loecker, 2007). Identification of the model parameters and estimation of TFP thus requires careful consideration of the potential confounding biases caused by the endogenous outputs. Following from the review of the literature on TFP estimation by Van Beveren (2012), we make use of a panel control function (CF) approach proposed by Akerberg et al. (2006), which is an extension to Levinsohn & Petrin (2003). Akerberg et al. (2006)'s CF approach uses a function with materials and capital as arguments to proxy for the endogenous unobserved productivity. This function is introduced into the production function (equation 3) as an additional model component to obtain consistent estimates of the model parameters.

Step 2: Estimating the effect of agglomeration on productivity

To estimate the causal impact of agglomeration on productivity, we consider the estimated TFP $\hat{\omega}_{it}^s$ to be a function of the agglomeration measure ρ_{it}^s indicating the MED of the zone j where the firm i is located.

$$\hat{\omega}_{it}^s = S^s(\rho_{it}^s) + \eta_{it}^s + \xi_{it}^s. \quad (5)$$

where η_{it}^s consists of the unobserved characteristics of firm-level productivity. ξ_{it}^s represents an idiosyncratic error term capturing all random shocks to the dependent variable. The exact structural form of how ρ_{it}^s enters the equation is unknown, so we adopt a non-parametric specification $S^s(\cdot)$ in which the shape of the relationship is delivered from the data and regression splines. Note that the percentage change in the estimated $S^s(\cdot)$ with respect to the percentage change in the model covariate at any level of the covariate ρ^s gives the corresponding value of agglomeration elasticity $\delta\rho^s$.

We expect η_{it}^s to be correlated with ρ_{it}^s . This correlation follows from the presence of omitted variables such as specific characteristics of local input and output markets, and functional or occupational differences caused by spatial self-selection by firms (see Section 1 for a detailed discussion). Further, the relationship between ρ_{it}^s and productivity $\hat{\omega}_{it}^s$ may be simultaneously determined as higher productivity locations may attract a greater level of private investment over time leading to larger economic mass, which has a feedback effect on productivity. These estimation issues need to be carefully addressed to ensure that the agglomeration elasticity estimates are, as far as possible, causal rather than being simply associational. Therefore, we adopt a non-parametric instrumental variables (NPIV) regression, which not only enables non-parametric specification of $S^s(\cdot)$ but also addresses potential endogeneity biases.

Bayesian Nonparametric Instrumental Variable Approach

IV-based estimators such as two-staged least square (2SLS) are widely adopted in applied econometrics to estimate parametric models that contain endogenous covariates. However, finite-dimensional parametric models (such as log-log models) for the relationship between agglomeration and productivity, are based on assumptions that are rarely justified by economic theories. The resulting model misspecification may lead to erroneous estimates of agglomeration elasticities. On the other hand, non-parametric methods have the potential to capture the salient features in a data-driven manner without making a priori assumptions on the functional form of the relationship (Horowitz, 2011). Therefore, a fairly growing strand in the econometrics literature proposes different approaches for NPIV regression, but such methods have not been considered in the estimation of the agglomeration-productivity relationship. Extensive reviews can be found in Newey & Powell (2003) and Horowitz (2011).

Classical (frequentist) NPIV regression approaches are popular in theoretical econometrics (Newey & Powell, 2003; Horowitz, 2011; Newey, 2013; Chetverikov & Wilhelm, 2017), but they are challenging to apply in practice due to two main reasons. First, tuning parameters to monitor the flexibility of $S(\cdot)$ are often required to be specified by the analyst. Second, standard errors are generally computed using bootstrap, making these methods computationally prohibitive for large datasets. Therefore, we adopt a scalable Bayesian NPIV approach, proposed by Wiesenfarth et al. (2014), that can produce a consistent estimate of non-parametric $S(\cdot)$, even if the analyst does not observe η_{it}^s . This Bayesian method addresses both challenges of the frequentist estimation because it *learns* tuning parameters related to $S(\cdot)$ during estimation and uncertainty in parameters estimates is inherently captured by credible intervals (analogous to classical confidence intervals). In addition, it also enables nonparametric specification of the unobserved error component ξ_{it}^s , precluding the need for making additional assumptions.

We discuss the adopted Bayesian NPIV approach (Wiesenfarth et al., 2014) for a model with a single endogenous covariate, that is,

$$\hat{\omega} = S(\rho) + \epsilon_2, \quad \rho = h(z) + \epsilon_1 \quad (6)$$

Note that η are encapsulated in ϵ_2 , and z is an instrument for the endogenous regressor ρ . The relationship between ρ and z is represented by an unknown functional form $h(\cdot)$ and ϵ_2 is an idiosyncratic random error term. For notational simplicity, we drop the firm-year subscripts and sector superscripts. Bayesian NPIV is a control function approach, and assumes the following standard identification restrictions:

$$E(\epsilon_1|z) = 0 \quad \text{and} \quad E(\epsilon_2|\epsilon_1, z) = E(\epsilon_2|\epsilon_1), \quad (7)$$

which yields

$$\begin{aligned} E(\hat{\omega}|\rho, z) &= S(\rho) + E(\epsilon_2|\epsilon_1, z) = S(\rho) + E(\epsilon_2|\epsilon_1) \\ &= S(\rho) + \nu(\epsilon_1), \end{aligned} \quad (8)$$

where $\nu(\epsilon_1)$ is a function of the unobserved error term ϵ_1 . This function is known as the control function.

Conditional on the availability of a valid instrument (see Section 2), Bayesian NPIV can correct for confounding bias. To account for the nonlinear effects of continuous covariates, both $S(\cdot)$ and $h(\cdot)$ (refer to equation 6) are specified in terms of additive predictors comprising penalised splines. Each of the functions $S(\cdot)$ and $h(\cdot)$ is approximated by a linear combination of suitable B-spline basis functions. The penalised spline approach uses a large enough number of equidistant knots in combination with a penalty to avoid over-fitting. Moreover, the joint distribution of ϵ_1 and ϵ_2 is specified using nonparametric Gaussian Dirichlet process mixture (DPM), which ensures the robustness of the model relative to extreme observations. Efficient Markov chain Monte Carlo (MCMC) simulation technique is employed for fully Bayesian inference. The resulting posterior samples allow us to construct simultaneous credible bands for the non-parametric effects (i.e., $S(\cdot)$ and $h(\cdot)$). Thereby, the possibility of non-normal error distribution is considered and the complete variability is represented by Bayesian NPIV. We now succinctly discuss specifications of the kernel error distribution in Bayesian NPIV.

To allow for a flexible distribution of error terms, the model considers a Gaussian DPM with

infinite mixture components, c , in the following hierarchy:

$$\begin{aligned}
(\epsilon_{1i}, \epsilon_{2i}) &\sim \sum_{c=1}^{\infty} \pi_c N(\mu_c, \Sigma_c) \\
(\mu_c, \Sigma_c) &\sim G_0 = N(\mu|\mu_0, \tau_{\Sigma}^{-1}\Sigma) \text{IW}(\Sigma|s_{\Sigma}, S_{\Sigma}) \\
\pi_c &= v_c \left(1 - \sum_{j=1}^{c-1} (1 - \pi_j) \right) = v_c \prod_{j=1}^{c-1} (1 - v_j), \\
c &= 1, 2, \dots \\
v_c &\sim \text{Be}(1, \psi).
\end{aligned} \tag{9}$$

where μ_c , Σ_c and π_c denote the component-specific means, variances and mixing proportions. The mixture components are assumed to be independent and identically distributed with the base distribution G_0 of the Dirichlet process (DP), where G_0 is given by a normal-inverse-Wishart distribution. The mixture weights are generated in a stick-breaking manner based on a Beta distribution with concentration parameter $\psi > 0$ of the DP. The concentration parameter ψ determines the strength of belief in the base distribution G_0 .

Estimation Practicalities

We exclude discussion of the Gibbs sampler of Bayesian NPIV for brevity and focus mainly on implementation details and posterior analysis. Interested readers can refer to Wiesenfarth et al. (2014) for the derivation of conditional posterior updates.

We use the *BayesIV* and *DPpackage* in R to estimate the Bayesian NPIV. We consider 50,000 posterior draws in the estimation, exclude the first 15,000 burn-in draws and keep every 10th draw from the remaining draws for the posterior analysis. The point-wise posterior mean is computed by taking the average of 3,500 posterior draws. Bayesian simultaneous credible bands are obtained using quantiles of the posterior draws. A simultaneous credible band is defined as the region I_{θ} such that $P_{S|data}(S \in I_{\theta}) = 1 - \theta$, that is, the posterior probability that the entire true function $S(\cdot)$ is inside the region given the data equals to $1 - \theta$. The Bayesian simultaneous credible bands are constructed using the point-wise credible intervals derived from the $\theta/2$ and $1 - \theta/2$ quantiles of the posterior samples of $S(\cdot)$ from the MCMC output such that $(1 - \theta)100\%$ of the sampled curves are contained in the credible band. A similar process is used to obtain the credible intervals of $h(\cdot)$.

Instrumental Variable

To satisfy the identification restrictions presented in equation 7, we need an instrumental variable (IV) z . The IV should be (i) exogenous, that is, uncorrelated with ϵ_2 ; (ii) relevant, that is, correlated with the endogenous covariate ρ , conditional on other covariates in the model.

We derive valid external instruments from traffic casualty data. We consider the ratio of serious and severe traffic casualties to total casualties among active mode (pedestrians and cyclists) and motorcycle users during morning and afternoon peak hours (that is, 6:30-10:30 hours and 16:00-20:00 hours) in zone j in year t as an IV z_{jt} for the MED (agglomeration) ρ_{jt} in zone j and year t . We argue that as the MED of a city increases, peak-hour road network congestion in the city may also increase, and consequently, the average speed of travel in the network may decrease. As a result of slower vehicular speeds, the proportion of serious and severe traffic casualties to total casualties among active mode and motorcycle users during peak hours may decrease. Our argument follows from the traffic safety literature that suggests that a decrease in congestion may exacerbate the severity of peak-hour traffic casualties amongst active mode users and cyclists (Li et al., 2012; Noland et al., 2008). We thus expect a strong negative correlation between the chosen IV z_{jt} and the endogenous covariate ρ_{jt} . Nevertheless, we argue that the chosen IV is exogenous because we do not expect the IV to scale with city size (population) and affect labour supply, and therefore, not directly determine the response variable (that is, TFP) of any firm i located in zone j and year t . In other words, we do not anticipate the chosen IV z_{jt} to feature in a model for the response variable ω_{jt} .

Data

To gauge the presence of non-linearities in agglomeration elasticities, we consider a sample of firms in England as our case study. We consider the period between 2015 to 2019 as the study

period. For this period, we investigate the causal impact of MED and productivity in six most relevant industry sectors: Manufacturing (MAN), Construction (CON), Wholesale and Distribution (WAD), Transport (TRA), Information and Communication Technology (ICT), and Finance (FIN). As geographical regions or zones, we consider the Middle Layer 2011 Census Super Output Areas (MSOA11) in England, which includes a total of 6,791 units with a mean population of 8185 people.

The data sources for the key variables of interest are detailed in the next two subsections.

Mean Effective Density

We obtain the data on annual employment levels in each MSA11 unit from the Business and Employment Register available at Nomis¹ (official census and labour market statistics), a public repository maintained by the Office for National Statistics (ONS). To calculate the distance between the MSA11 units, we extract the location information on MSA11 units available in the ONS Postcode Directory², that is a detailed location database of all UK postcodes.

Traffic casualty data for the construction of IVs is obtained from the publicly available road safety data, maintained by the Department for Transport³.

Total Factor Productivity

The Department of Trade and Industry records all the accounting information provided by all companies registered in the UK. This information is available via the commercial software package Financial Analysis Made Easy (FAME)⁴, co-hosted by Vistra and Bureau Van Dijk. To estimate the production function in equation 3, we extract the annual data on the following variables for each registered firm:

1. Turnover (output): The net income of the company.
2. Fixed Assets (capital): The depreciated value of buildings, plants and equipment.
3. Current Assets (materials): The current stocks and debt owned by the company.
4. Total Employees (labour): The total number of employees in the company.

To limit potential endogeneity biases emerging from spatial self-selection by firms (Graham & Gibbons, 2019), we remove firms with more than one trading address and those that have a registered office address different from the main trading address. We also filter out firms with international subsidiaries. Additionally, we only focus on small and medium-sized firms with a number of employees between 10 and 249 to reduce endogeneity from spatial self-selection of labour (Graham, 2009). Finally, we class the filtered data into the six industry sectors using their two-digits Standard Industrial Classification 2007 (SIC07). The resulting number of observations for each industry sector is reported in Table 1.

Table 1: Classification of firms into industry sectors.

Industry Sector	SIC07	Firms	Observations
MAN	10-33	842	4210
CON	41-43	368	1840
WAD	45-47	688	3440
TRA	49-56	246	1230
ICT	58-63	357	1785
FIN	64-74	1452	7260

3 RESULTS AND DISCUSSION

This section is divided into four subsections. In the first subsection, we describe our MED estimates for England. In the second subsection, we briefly visit the estimated parameters of the production function for various industry sectors and the estimated TFP values. In the penultimate subsection,

¹Available at <https://www.nomisweb.co.uk/>.

²Available at <https://geoportal.statistics.gov.uk/>.

³Available at <https://www.data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data>.

⁴Available at <https://fame.bvdinfo.com/version-202274/fame/1/Companies/Search>.

we discuss the agglomeration elasticities obtained from the Bayesian NPIV estimation. Using these estimates, we also describe the spatial distribution of agglomeration benefits in England. In the final subsection, we present the estimated kernel error distributions to illustrate the importance of the non-parametric DPM specification. The relevance of our instruments is also demonstrated in this subsection.

Estimated Mean Effective Densities

Table 2 presents the summary statistics for the estimated MED values for each MSOA11 in England. The table indicates that the distribution of MED values in England is positively skewed. While some zones have high levels of MED or agglomeration, most zones have low values. Thus, only a few zones in England show high values of agglomeration. This observation is further supported by Figure 1, which maps MED values constructed using total employment in 2019 as mass. This figure illustrates that whereas regions in and around cities like London, Manchester, and Birmingham correspond to higher levels of agglomeration, they only constitute a small geographical area in England.

Table 2: Summary of estimated MED for England.

Statistic	2015	2016	2017	2018	2019
Mean	3934.64	4008.44	4071.00	4091.90	4154.83
Median	3375.18	3435.73	3480.50	3502.10	3550.63
Std. dev.	2272.62	2328.46	2374.13	2397.93	2450.14
Max	19706.27	20297.82	20745.36	21064.78	21694.94
Min	663.97	675.28	684.24	687.70	697.25
Skewness	2.43	2.44	2.44	2.48	2.50

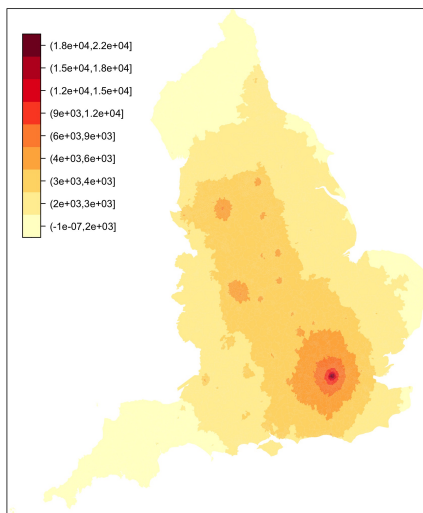


Figure 1: Map of MED values for England with total employment in 2019 as mass.

We complement the above-described statistics with Figure 2, which provides a histogram of the MED levels experienced by firms in each industry sector. We note from this figure that the majority of firms in the CON, MAN, TRA, and WAD sectors tend to locate in areas with MED values less than 6000. From a strategic point of view, firms in these sectors require large factories or warehouses. They may, therefore, prefer to locate these facilities in the periphery of cities where land prices, rents, and other costs are lower. Nonetheless, additional local maxima in their density plots at higher levels of MED also reveals the presence of a small number of firms in city centers, which may choose to locate their offices in central business district (CBD) for an easy commitment and location status. Conversely, firms in the ICT and FIN sectors, primarily tend to spread across the CBD, to avail the above-mentioned advantages. In the rest of this section, we quantify how these location choices translate into productivity benefits for firms.

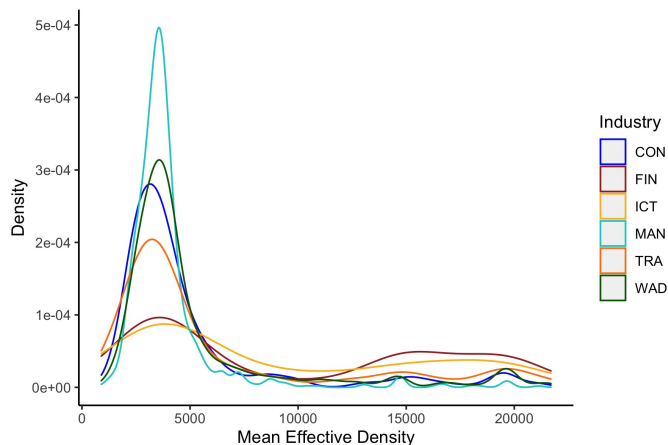


Figure 2: Histogram of agglomeration levels experienced by firms in 2019.

Estimated Total Factor Productivity

The parameter estimates of the production function, given by equation 3, for the six industry sectors are summarised in Table 3. We note that the FIN, ICT, MAN and WAD sectors are associated with returns to scale (RTS) values less than one, that is, decreasing RTS indicating that their output increases by less than the proportional change in all inputs. On the contrary, the CON and TRA sectors are associated with increasing RTS, implying a more than proportional increase in outputs with respect to inputs. These estimates summarise the technological advantages or disadvantages to firms in each industry.

Table 3: Parameter Estimates of the Production Function.

Sector	No. of firms	β_l	β_k	β_m	RTS
CON	1841	0.695 (0.010)	0.467 (0.014)	-0.040 (0.012)	1.123 (0.021)
FIN	7261	0.481 (0.006)	0.477 (0.008)	0.032 (0.010)	0.990 (0.014)
ICT	1786	0.323 (0.008)	0.630 (0.008)	0.006 (0.020)	0.959 (0.023)
MAN	4211	0.332 (0.004)	0.600 (0.012)	0.020 (0.012)	0.953 (0.014)
TRA	1231	0.439 (0.012)	0.539 (0.050)	0.026 (0.016)	1.005 (0.054)
WAD	3441	0.304 (0.002)	0.692 (0.002)	-0.011 (0.002)	0.985 (0.003)

Table 4 presents the summary statistics for the estimated TFP values for each industry sector. The mean and median statistics suggest that firms in the WAD sector are associated with the highest level of unobserved productivity, closely followed by firms in the MAN and ICT sectors. The sectors FIN, TRA, and CON show lower levels of unobserved productivity.

Table 4: Summary of estimated TFP values.

Statistic	CON	FIN	ICT	MAN	TRA	WAD
Min	-2.11	-1.29	2.35	-2.09	1.32	5.37
Median	6.93	7.41	8.02	8.07	7.26	8.80
Mean	6.94	7.45	8.04	8.10	7.19	8.90
Max	10.38	14.19	13.18	12.89	12.48	13.48
Std. dev.	0.90	1.09	1.084	0.74	1.11	0.92

Results from Bayesian NPIV Estimation

Figure 3 presents the estimates of $S(\cdot)$ (see equation 6, second-stage) for the six industry sectors. The plots include the mean estimates and the 95 percent credible bands (shown by the dotted line). The density of tick marks along the X-axis represents the number of observations in the corresponding domain of agglomeration. The figure indicates the presence of significant

non-linearities in the agglomeration-productivity relationship. This observation validates our hypothesis that presuming a log-log functional form may yield biased estimates of this relationship and associated agglomeration elasticities.

In Figure 4, we plot the agglomeration elasticities obtained at different levels of agglomeration and their corresponding credible bands. Note that these estimates are obtained using the 3,500 posterior draws (refer to Section 2) that are available at multiple points along the support of the model covariate, that is, MED. To obtain the elasticity at any point ρ , we identify a small interval $[\rho_1, \rho_2]$ surrounding ρ , where $\rho_1 = \rho - \Delta\rho$ and $\rho_2 = \rho + \Delta\rho$. We extract the 3,500 posterior draws at the two points ρ_1 and ρ_2 and calculate the change in elasticity for each draw. The mean of the resulting 3,500 samples gives the reported elasticity estimate at ρ and the quantiles 0.025 and 0.975 give the corresponding lower and upper limits of the 95-percent credible bands.

For the CON sector, we note that agglomeration elasticities remain positive for MED levels below 8000, with values increasing from 0.15 to 0.72 and back to 0.48, and become statistically insignificant beyond that. For the MAN sector, we observe that the agglomeration elasticities are statistically significant between MED levels of 8000 to 1000 and 13000 to 18000, and statistically insignificant otherwise. The estimated elasticities increase from 0.19 to 0.78 until a MED level of 16000 and drop back to 0.17 before becoming statistically insignificant. Our estimates for the WAD sector suggest that the agglomeration elasticities fall from a value of 0.16 at MED level 2000 to a value of -0.20 at MED level 8000, while remaining positive (and statistically significant) between MED levels of 2000 to 5000, become negative (and statistically significant) between MED levels of 7000 to 10000. The agglomeration elasticities become positive (and statistically significant) again at a MED level of 11000, followed by a steep increase to the value of 0.75 at a MED level of 13000. The agglomeration elasticities, thereafter, fall to a value of 0.22 at a MED level of 15000, beyond which they become statistically insignificant. The estimated agglomeration elasticities for the TRA sector remain positive and statistically significant between MED levels of 6000 to 12000, and range between 0 to 0.50, the maximum being achieved at a MED level of 9000. Higher levels of agglomeration of the order of 18000 to 20000 MED are associated with negative (and statistically significant) values of agglomeration elasticities ranging between -0.75 to -0.50. Firms in the ICT sector are found to be associated with positive agglomeration elasticities in the interval $[0.12, 0.22]$ at agglomeration levels between 7000 to 15000 MED, the maximum being observed at a MED level of 13000. The agglomeration elasticities remain statistically insignificant otherwise. Finally, for the FIN sector, we first observe negative and statistically significant agglomeration elasticities between MED levels of 3000 to 5000. The estimated elasticities are of the order of -0.1. Nonetheless, the elasticities remain positive (and statistically significant) at MED levels between 7000 to 9000, 11000 to 1500, and also beyond MED levels of 18000. The positive values range in the interval $[0.18, 0.60]$, with the maximum occurring at a level of 13000 MED. Overall, Figure 4 indicates that barring the CON sector, the productivity benefits in all sectors comment into effect beyond a critical mass of agglomeration. This critical mass varies across industries.

Table 5 summarises the estimated agglomeration elasticities. The final column in the table reports the estimates from a one-step procedure where MED enters as a covariate in the production function (equation 3). The values in the final column are fairly consistent with the literature in which elasticity estimates have been derived by assuming the productivity-agglomeration relationship to be log-log (see Graham & Gibbons (2019) for a summary of 47 international empirical studies on the effects of agglomeration on productivity). Our results suggest that our non-linear agglomeration elasticity estimates take more extreme values compared to their log-log counterparts.

Next, we map the estimated agglomeration elasticities to the different zones (that is, MSOA11 units) in England using their MED values. This mapping allows us to understand the spatial distribution of the agglomeration impacts in England. Figure 5 shows these distributions for each industry sector. Note that we adopt the same colour key for each map in Figure 5 to allow comparison of agglomeration impacts across industries.

From Figure 5a, we note that the highest levels of agglomeration benefits (elasticities ranging in the interval $[0.4, 0.9]$) in the CON sector can be observed in the peripheral regions of the Greater London Area (GLA) (for instance, Slough, Watford, and Loughton) and within cities of Manchester and Birmingham. Interestingly, the areas within the GLA are associated with statistically insignificant productivity effects of agglomeration. All other MSOA11 units are associated with agglomeration elasticities ranging between 0.1 and 0.3. Figure 5b suggests that the significant agglomeration benefits in the MAN sector remain confined within the GLA, while remaining the highest in the regions immediately surrounding the CBD of London (which includes the City of London, City of Westminster and Kensington and Chelsea, among others). Figure 5c indicates the firms in the WAD sector avail the highest benefits of agglomeration on their productivity by locating in the outskirts of the CBD of London. The impacts within the CBD of London and cities such as

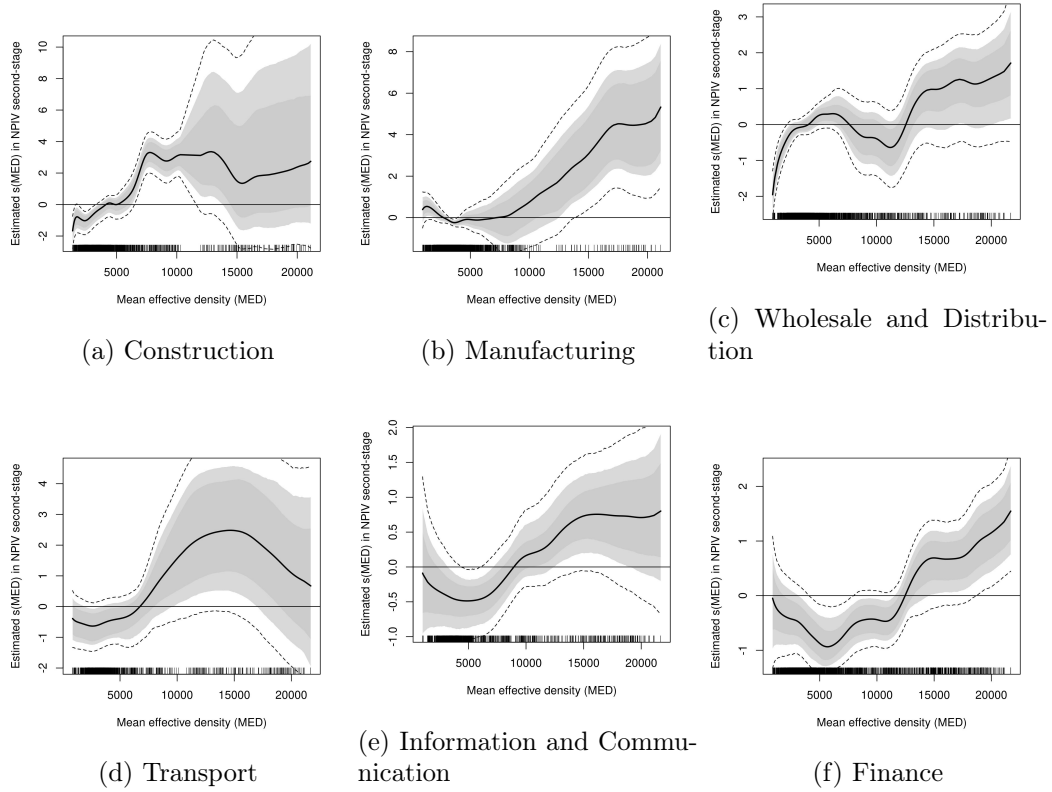


Figure 3: Estimated relationships between agglomeration and firm productivity.

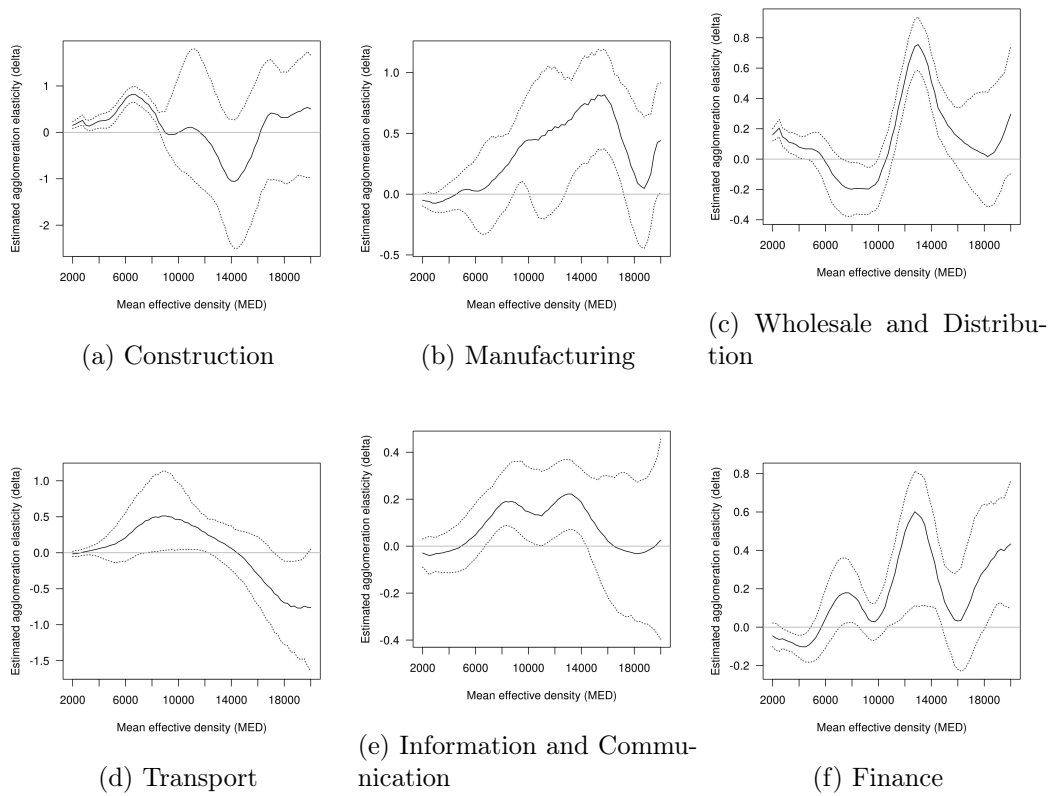


Figure 4: Estimated agglomeration elasticities.

Table 5: Summary of estimated agglomeration elasticities.

Sectors	Mean Effective Density (MED)									
	2000	4000	6000	8000	10000	12000	14000	16000	18000	One-step estimate
CON	0.15 (0.04)	0.24 (0.07)	0.72 (0.09)	0.48 (0.09)	0 (0.48)	-0.1 (0.69)	-1.05 (0.68)	-0.17 (0.68)	0.32 (0.57)	0.045 (0.009)
MAN	-0.05 (0.02)	-0.03 (0.06)	0.03 (0.15)	0.19 (0.16)	0.45 (0.21)	0.56 (0.32)	0.73 (0.22)	0.78 (0.2)	0.17 (0.3)	-0.002 (0.005)
WAD	0.16 (0.02)	0.08 (0.04)	-0.02 (0.08)	-0.2 (0.1)	-0.14 (0.08)	0.52 (0.08)	0.51 (0.09)	0.15 (0.1)	0.03 (0.19)	-0.067 (0.002)
TRA	-0.02 (0.02)	0.06 (0.05)	0.21 (0.17)	0.49 (0.28)	0.47 (0.23)	0.28 (0.12)	0.06 (0.15)	-0.3 (0.23)	-0.7 (0.3)	0.115 (0.015)
ICT	-0.03 (0.03)	-0.02 (0.04)	0.05 (0.06)	0.18 (0.07)	0.15 (0.07)	0.18 (0.08)	0.19 (0.07)	0.02 (0.13)	-0.03 (0.15)	0.072 (0.009)
FIN	-0.05 (0.03)	-0.1 (0.03)	0.04 (0.08)	0.17 (0.08)	0.05 (0.05)	0.5 (0.15)	0.4 (0.13)	0.03 (0.13)	0.3 (0.16)	0.144 (0.003)

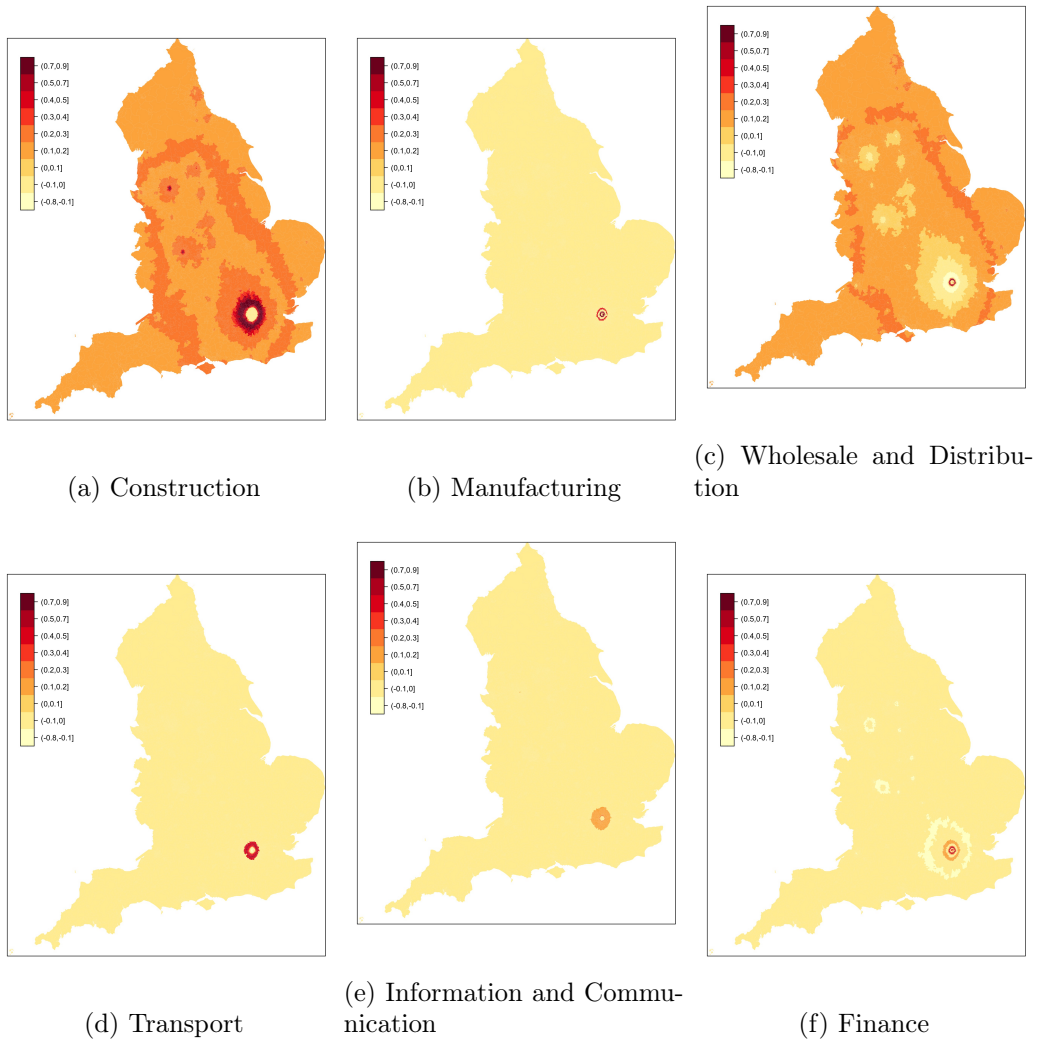


Figure 5: Spatial distribution of the agglomeration impacts in England in 2019.

Birmingham and Manchester either remain negative or statistically insignificant. Other regions show positive impacts of varying degrees as represented by the color key. Similar to the MAN sector, firms in TRA and ICT sectors (Figures 5d and 5e) observe agglomeration benefits in zones surrounding the CBD of London. From Figure 5f, we note that firms in the FIN sector derive the highest agglomeration benefit by locating within the CBD of London. Additionally, while most areas within the GLA observe positive productivity impacts of agglomeration, the peripheral areas of the GLA, Birmingham, and Manchester are associated with diseconomies of agglomeration. The effect in other areas remains statistically insignificant.

4 CONCLUSIONS

Transport accounted for 5 to 8 percent of the total public expenditure on services in the UK through the years 2017 to 2022⁵. In a typical year, the United Kingdom spends about £45 billion pounds on construction and maintenance of transport infrastructure. Understanding the economic and social benefits arising from investments of this order is thus important for policymakers. This paper contributes to the growing strand in the urban economic literature that focuses on measuring the Wider economic impacts (WEIs) of transport investments arising via scale economies of agglomeration.

We make two advances in this agenda. First, we develop a causal statistical framework to quantify the non-linearities in the relationship between agglomeration (represented by Mean Effective Density (MED)) and productivity (measured as Total Factor Productivity (TFP)). The estimated relationships, for the first time, provide a quantification of how agglomeration elasticities vary over different levels of agglomeration. Second, we determine a novel external instrument derived from traffic casualty data to identify the agglomeration-productivity relationship. Our study suggests the use of the severity of traffic casualties among active mode users and motorcyclists during peak hours as a relevant and exogenous instrument for agglomeration.

Our investigation of agglomeration elasticities in six key industry sectors in England suggests that agglomeration elasticities vary significantly over agglomeration levels. For the Construction Sector, we observe positive agglomeration elasticities only at low and mid levels of agglomeration. For the other five sectors which include, Manufacturing, Wholesale and Distribution, Transport, Information and Communication Technology, and Finance, we note the presence of a critical mass of agglomeration beyond which the positive benefits of agglomeration on productivity can be observed. Below this critical level, the agglomeration elasticities either remain negative (but statistically significant) or statistically insignificant at the 95-percent confidence level. Additionally, we note that agglomeration elasticities in the Transport sector become negative at extremely high levels of agglomeration, while very low levels of agglomeration show positive agglomeration elasticities for the Wholesale and Distribution sector, but of lower magnitude. Interestingly, the estimated agglomeration elasticities in this study take more extreme values than ones derived from a log-log model of productivity and agglomeration as adopted in the literature. Our estimates thus have crucial implications for the appraisal of transport investments.

Further, our exploration of the spatial distribution of the agglomeration impacts in England reveals that the highest levels of agglomeration benefits in the Construction sector are observed in the regions surrounding the Greater London Area (GLA) and within Manchester and Birmingham. For the Manufacturing, Wholesale and Distribution, Transport and Information and Communication Technology sectors, the largest productivity benefits of agglomeration are confined within the GLA, particularly along the fringes of its central business district (CBD). For the Finance Sector, the highest positive agglomeration elasticities are associated with the regions in the CBD of the GLA, while the outskirts of the GLA, Manchester and Birmingham see diseconomies of agglomeration. Our findings are unsurprising: these spatial patterns are consistent with the sector-wise preferences for office locations by firms. Refining this investigation with more data, particularly from other years or from other countries, is an important topic for further research and can provide an empirical basis for targetting transport investments in a manner that can spread productivity benefits more evenly.

ACRONYMS

2SLS two-staged least square

ATEM access to economic mass

CBA Cost Benefit Analysis

CBD central business district

CF control function

CON Construction

DP Dirichlet process

DPM Dirichlet process mixture

⁵<https://www.gov.uk/government/statistics/public-expenditure-statistical-analyses-2022>

DUBs Direct user benefits
FAME Financial Analysis Made Easy
FIN Finance
GLA Greater London Area
ICT Information and Communication Technology
IV instrumental variable
IVs instrumental variables
MAN Manufacturing
MCMC Markov chain Monte Carlo
MED Mean Effective Density
MSOA11 Middle Layer 2011 Census Super Output Areas
NP non-parametric
NPIV non-parametric instrumental variables
ONS Office for National Statistics
RTS returns to scale
SIC07 Standard Industrial Classification 2007
TAG Transport Analysis Guidance
TFP Total Factor Productivity
TRA Transport
WAD Wholesale and Distribution
WEIs Wider economic impacts

REFERENCES

- Akerberg, D., Caves, K., & Frazer, G. (2006). Structural identification of production functions.
- Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the Berlin Wall. *Econometrica*, *83*(6), 2127–2189.
- Au, C.-C., & Henderson, J. V. (2006). Are Chinese cities too small? *The Review of Economic Studies*, *73*(3), 549–576.
- Baldwin, J. R., Brown, W. M., & Rigby, D. L. (2010). Agglomeration economies: Microdata panel estimates from Canadian manufacturing. *Journal of Regional Science*, *50*(5), 915–934.
- Chetverikov, D., & Wilhelm, D. (2017). Nonparametric Instrumental Variables Estimation under Monotonicity. *Econometrica*, *85*(4), 1303–1320.
- Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2012). Sorting and local wage and skill distributions in France. *Regional Science and Urban Economics*, *42*(6), 913–930.
- Combes, P.-P., & Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics* (Vol. 5, pp. 247–348). Elsevier.
- De Loecker, J. (2007). Do exports generate higher productivity? Evidence from Slovenia. *Journal of international economics*, *73*(1), 69–98.
- DfT. (2016). TAG UNIT A2.4: Appraisal of Productivity Impacts.
- Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics* (Vol. 4, pp. 2063–2117). Elsevier.

- Graham, D. J. (2009). Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science*, 88(1), 63–84.
- Graham, D. J., & Dender, K. V. (2011). Estimating the agglomeration benefits of transport investments: some tests for stability. *Transportation*, 38(3), 409–426.
- Graham, D. J., & Gibbons, S. (2019). Quantifying Wider Economic Impacts of agglomeration for transport appraisal: Existing evidence and future directions. *Economics of Transportation*, 19, 100121.
- Graham, D. J., Melo, P. S., Jiwattanakupaisarn, P., & Noland, R. B. (2010). Testing for causality between productivity and agglomeration economies. *Journal of regional Science*, 50(5), 935–951.
- Horowitz, J. L. (2011). Applied Nonparametric Instrumental Variables Estimation. *Econometrica*, 79(2), 347–394.
- Lall, S. V., Shalizi, Z., & Deichmann, U. (2004). Agglomeration economies and productivity in Indian industry. *Journal of development economics*, 73(2), 643–673.
- Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2), 317–341.
- Li, H., Graham, D. J., & Majumdar, A. (2012). The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation. *Accident Analysis & Prevention*, 49, 366–377.
- Mackie, P., Graham, D., & Laird, D. (2012). Direct and wider economic benefits in transport appraisal. In *Handbook in transport economics. Edward Elgar, London* (pp. 501–526).
- Maré, D. C., & Graham, D. J. (2009). Agglomeration elasticities in New Zealand. *Motu Working Paper No. 09-06*.
- Maré, D. C., & Graham, D. J. (2013). Agglomeration elasticities and firm heterogeneity. *Journal of Urban Economics*, 75, 44–56.
- Marrocu, E., Paci, R., & Usai, S. (2013). Productivity growth in the old and new Europe: the role of agglomeration externalities. *Journal of Regional Science*, 53(3), 418–442.
- Marshall, A. (1920). *Principles of Political Economy* (8th ed.). London, Macmillan.
- Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics*, 39(3), 332–342.
- Morikawa, M. (2011). Economies of density and productivity in service industries: An analysis of personal service industries based on establishment-level data. *The Review of Economics and Statistics*, 93(1), 179–192.
- Newey, W. K. (2013). Nonparametric Instrumental Variables Estimation. *American Economic Review*, 103(3), 550–556.
- Newey, W. K., & Powell, J. L. (2003). Instrumental Variables Estimation of Nonparametric Models. *Econometrica*, 71(5), 1565–1578.
- Noland, R. B., Quddus, M. A., & Ochieng, W. Y. (2008). The effect of the London congestion charge on road casualties: an intervention analysis. *Transportation*, 35(1), 73–91.
- Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2), 373–389.
- Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of economic surveys*, 26(1), 98–128.
- Venables, A., Laird, J. J., & Overman, H. G. (2014). Transport investment and economic performance: Implications for project appraisal.

- Venables, A. J. (2007). Evaluating urban transport improvements: cost-benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy (JTEP)*, 41(2), 173–188.
- Wiesenfarth, M., Hisgen, C. M., Kneib, T., & Cadarso-Suarez, C. (2014). Bayesian nonparametric instrumental variables regression based on penalized splines and dirichlet process mixtures. *Journal of Business & Economic Statistics*, 32(3), 468–482.

Acceptance of new technologies affecting safety on electric bicycles: evidence from five European countries

Georgios Kapousizis*¹, Rumana Sarker², Baran Ulak³, Karst Geurs⁴

¹ PhD candidate, Department of Civil Engineering, University of Twente, Netherlands

² Research Fellow, Department of Civil Engineering, Monash University, Australia

³ Assistance Professor, Department of Civil Engineering, University of Twente, Netherlands

⁴ Professor, Department of Civil Engineering, University of Twente, Netherlands

SHORT SUMMARY

Electric bicycles (e-bikes) are one of the main solutions towards mitigating transport externalities, such as traffic congestion and emission, and have thus been promoted in many countries. Despite the advantages of e-bikes, users are prone to be involved in crashes, usually due to the high speed. Leveraging new technologies could help reduce such crashes; however, e-bike users' willingness to accept new technologies still needs to be investigated. Hence, this study explores e-bike users' motivation to use smart e-bikes by adopting the extended Unified Theory of Acceptance and Use of Technology (UTAUT2). A cross-national survey was administered in five European countries- Austria, Belgium, Germany, Greece and the Netherlands, differing in sizes and cycling culture. The survey yielded 1116 responses, and the structural equation model (SEM) results indicate that 'performance expectancy', 'hedonic motivation' and 'perceived safety' are the strongest predictors of users' acceptance of new technologies on e-bikes to increase safety and comfort.

Keywords: Cycling safety, E-bikes, UTAUT2, User acceptance, SEM

1. INTRODUCTION

The Covid-19 pandemic and the recent energy crisis have pushed a significant number of people to switch to more active transport modes such as cycling (Nikitas et al., 2021; Shimano, 2022). Despite the numerous benefits of cycling, there are also certain barriers such as low fitness levels, topographical difficulties, and established habits, preventing more people embrace cycling as their everyday transportation (Fishman & Cherry, 2016; Plazier, 2022). E-bikes can help overcome some of these barriers in front of widespread adoption of cycling. Since with an e-bike people can travel faster and longer distances compared to conventional bicycles, several governments worldwide have lately been promoting e-bikes as one of the main measures to mitigate negative transport externalities such as congestion and emission.

Many European countries subsidise the purchase of e-bikes (ECF, 2023) and as a result, there is an increase in the number of e-bikes being sold in Europe in recent years. In 2021, around five million new e-bikes were sold in Europe (Sutton, 2022), which is the highest numbers of bicycles sold in a decade (Statista, 2020). However, this increase in bicycles lead to major safety concerns in many countries which have inadequate cycling infrastructure. Furthermore, e-bikes potentially lead to more severe crashes as they are usually faster than regular bicycles and aging people can also use them more easily (Gadsby & Watkins, 2020; Panwinkler & Holz-Rau, 2021; J. P. Schepers et al., 2014). That is why countries like the Netherlands, with one of the best and well design bicycle network (P. Schepers et al., 2017), still experience many e-bike crashes (Statistics Netherlands (CBS), 2021). One of the ways to address this increasing safety concern is the

adoption of new technologies such as sensors on bicycles and Internet of Things (IoT) to prevent crashes and reduce severities. Such systems can be the future of a sustainable cycling environment as they could positively influence and increase cycling safety (Boronat et al., 2021; Oliveira et al., 2021). Kapousizis et al. (2022) showed that in the last decade, there is a plethora of published studies about new technologies that can increase cycling safety, and proposed a classification for the ‘bicycle smartness levels’ (BSLs) consisting of 6 levels. This study focused on technologies at third level to investigate user acceptance of specific functionalities in their e-bike, considering that the Level 3 involves the most feasible and readily available technology given the highest Technology Readiness. The level 3 consists of surrounding detection, collision avoidance, speed warnings and post-accident notifications.

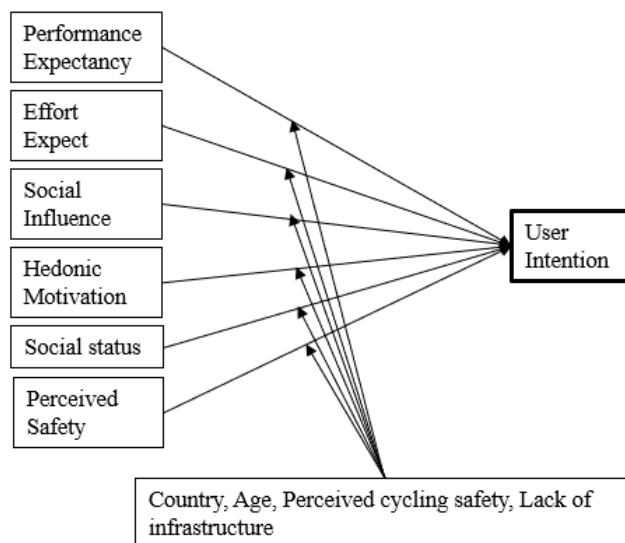
While new bicycle technologies are shown to positively affect cyclists’ safety and comfort, less is known about users’ acceptance and intention to use these features. There is still scarce literature investigating users’ intention to adopt new bicycle technologies to increase safety and comfort. To cover this gap, this study aims to investigate users’ intention to accept new technologies on e-bikes by collecting data and comparing factors across different countries.

2. METHODOLOGY

In this study, the framework of the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) was adopted (Venkatesh et al., 2012). We used the UTAUT2 as a baseline and adjusted it with most appropriate constructs that fit this study. The conceptual model with the constructs is presented in Figure 1. Adjustments in the UTAUT2 framework are common, especially in transport research such as automated vehicles, since this technology is not available yet and researchers are investigating this a priori (Kasper & Abdelrahman, 2020; Nordhoff et al., 2020). In an attempt to explore specific factors, we adjusted the model to fit this study’s aim.

This is the first study that adopted the UTAUT2 framework and tailored it accordingly to examine users’ intention to use new bicycle technologies that affect cycling safety and comfort. To develop the hypotheses, we have included psychological constructs from other domains, such as advanced driving assistance systems and automated vehicles.

Figure 1: Conceptual model (Extended UTAUT2)



Following the UTAUT2, it is hypothesised that users' intentions to adopt new technologies in e-bikes are related to performance and effort expectancy, social influences and hedonic motivation. Within this framework, an extended version of the framework is considered by including 'social norm' and 'perceived safety' as factors that could affect people's intention to accept new technologies. We excluded 'facilitating conditions', 'price value', and 'habit' of the UTAUT2 model since these technologies are not commercially available yet. Notably, according to the UTAUT2, the antecedents are independent and directly linked to the behavioural intention of technology adoption in e-bike. The elements in the conceptual framework are as follows:

Performance expectancy relates to individual beliefs concerning a system (Venkatesh et al., 2003). In the context of this study, performance expectancy is defined as the degree of usefulness an individual can get using new technologies on e-bikes. We assume that the performance expectancy construct will be a strong predictor.

Effort expectancy justifies the ease of use of a specific system (Venkatesh et al., 2003) and is also associated with the degree of consumers' ease of use (Venkatesh et al., 2012). In the context of this study, we believe that effort expectancy will positively influence behavioural intention.

Social influence is defined as an individual's perception of what others believe they should use a specific technology and to what extent others' opinion influences an individual to accept and use a specific technology (Venkatesh et al., 2003).

Hedonic motivation proves an individual's enjoyment using technology (Venkatesh et al., 2012). We believe hedonic motivation can be derived from the new technologies on e-bikes and fulfil individual satisfaction.

Social norm refers to individual's behaviour modification based on their belief of what others expect from them.

Perceived safety is frequently used in several studies predicting the influence of an individual to use technology due to their belief that it will improve their safety (Kapsler & Abdelrahman, 2020; Nordhoff et al., 2020). Hence, we construct the following hypothesis.

The following research hypotheses are tested from the conceptual model (Fig. 1):

H1: Performance expectancy positively influences behavioural intention to use new technologies on e-bikes.

H2: Effort expectancy positively influences behavioural intention to use new technologies on e-bikes.

H3: Social influence positively influences behavioural intention to use new technologies on e-bikes.

H4: Hedonic motivation positively influences behavioural intention to use new technologies on e-bikes.

H5: Social status positively influences behavioural intention to use new technologies on e-bikes.

H6: Perceived safety positively influences behavioural intention to use new technologies on e-bikes.

Survey

To investigate the aforementioned hypotheses, we conducted an online survey -translated into five languages (English, German, Greek, Dutch and French), which was distributed in five European countries, Austria, Belgium, Germany, Greece and the Netherlands, between November 2022 and January 2023. All constructs are measured by 3-5 standardised items assessed on 5-point Likert scales. The focus group of the survey comprised people who already use an e-bike or are willing to buy one. This group was chosen to collect more realistic results than asking people not interested in cycling. Countries were not selected randomly; on the contrary, they were chosen due to the varying quality of cycling infrastructure and cycling culture to understand users' perceived safety in different scenarios. In total, 1116 responses were collected.

Table 1: Sample

Variable	Austria	Belgium	Germany	Greece	Netherlands
Number of respondents	75	199	115	199	528
Gender					
Male	53	106	79	142	322
Female	18	90	35	56	195
Other	4	3	1	1	11
Age					
18-29	3	10	11	26	24
30-39	16	24	15	54	20
40-49	10	29	17	61	25
50-59	23	37	37	39	80
60-69	18	64	29	19	190
>70	5	35	6	0	189

We performed a Structural Equation Model (SEM) to analyse the behavioural framework. The SEM in this study contains three sets of equations: measurement equations, structural equations linking the latent constructs to observed characteristics of the participants, and structural equations relating the latent constructs to the dependent variables (user intention of new technologies on e-bikes). The model was estimated using the SPSS-AMOS.

3. RESULTS AND DISCUSSION

To investigate users' intention to use new technologies on e-bikes, a SEM model was analysed. Maximum likelihood method was used and the model was assessed through the five most commonly used goodness-of-fit indexes: Chi-square per degree of freedom (CMIN/DF: $1.0 < \chi^2 < 3.0$) (CMIN/DF : 2.80), Comparative Fit Index (CFI: >0.095) (CFI: 0.981), Tucker-Lewis index (TLI: >0.95) (TLI: 0.976), Root Mean Square Error of Approximation (RMSEA: < 0.07) (RMSEA : 0.04) and Standardised Root Mean Square Residual (SRMSR < 0.05) (SRMSR : 0.0266) (Hair et al., 2014; Schumacker & Lomax, 2010). In this study, standardised factor loadings are between 0.638 to 0.949, above the threshold of 0.5 (Hair et al., 2014, p. 618). The model was assessed for convergent and discriminant validity; Average Variance Extracted (AVE) was above the cut-off criterion of 0.50 (Fornell & Larcker, 1981; Hair et al., 2014), which illustrates the convergent validity. Composite reliability (CR) was also above the acceptance threshold of 0.7 (Hair et al., 2014), supporting internal consistency.

A significant positive relationship was found between performance expectancy and behavioural intention, hedonic motivation and behavioural intention, perceived safety and behavioural intention, social influence and behavioural intention, and effort expectancy and behavioural intention. Thus, this shows that performance expectancy, hedonic motivation and perceived safety are the stronger constructs and are important factors in user intention of the new technologies on e-bikes. Social influence is also an important and positive aspect in user intention, while effort expectancy has a mild positive significant role. In contrast, there is no significant relationship between social status and behavioural intention. The hypotheses and their structural results of this study are presented in Table 2.

Table 2: Results of structural relationships

Hypothesis	β	Significance	Results
H1	0.398	< 0.001	supported
H2	0.039	0.017	supported
H3	0.068	< 0.001	supported
H4	0.326	< 0.001	supported
H5	0.027	0.220	rejected
H6	0.134	< 0.001	supported

The variability of behavioural intention to use new technologies on e-bikes is explained by 84% of the proposed model. Investigating cross-country differences is evidence that performance expectancy has a strong and positive impact on user intention across all countries, while hedonic motivation has no significant impact in the Austrian sample. Additionally, perceived safety positively influences user intention in Belgium, Germany, and Netherlands. Social influence has a stronger relationship and is significant to the Dutch and Austrian responders. Finally, there is no significance on social status construct; this tendency is similar in all five countries. These results are presented in Table 3.

This model was also tested with a series of controls against user intention. In this attempt, we controlled our model to gender, age, lack of infrastructure, and perceived cycling safety (see Table 3). Note that age was tested as continuous while the rest variables were dummy coded. Results show that user intention significantly increases with increasing age in all counties except Belgium. Gender significantly impacts user intention in Belgium, while the rest variables are not statistically significant across all countries.

Table 3: Results of cross-country analysis

Variables	Austria	Belgium	Germany	Greece	Netherlands
<i>Dependent variable: Intention</i>					
Performance expectancy	0.455***	441***	0.461***	0.444***	0.369***
Effort expectancy	0.114	0.026	0.043	0.047	0.050*
Social influence	0.160**	0.004	-0.085	0.069	0.120***
Hedonic motivation	0.113	0.273**	0.284**	0.276***	0.359***
Social status	0.068	0.032	0.099	0.077	-0.016
Perceived safety	0.115	0.190**	0.186**	0.077	0.129***
Age	0.045*	0.027	0.059**	0.069**	0.066***
Gender (male)	0.044	0.044	0.035	- 0.002	- 0.003
Perceived cycling safety (high)	- 0.043	- 0.019	- 0.040	0.051	- 0.027
Lack of cycling infrastructure (yes)	0.006	- 0.012	- 0.021	- 0.026	0.007

***: p-value < 0.001, **: p-value < 0.05, *:p-value < 0.1

4. CONCLUSIONS

This study provides novel results for the user acceptance of new technologies on e-bikes as a potential solution to improve e-bike safety and comfort. We employed an extended framework of the UTAUT2, which applied to survey data from five European countries. We tested six constructs, while only five were supported (Performance expectancy, Social influence, Hedonic motivation, Perceived safety and Effort expectancy), with Performance expectancy, Hedonic motivation and Perceived safety having a strong relationship with users' intention to use new technologies on e-bikes in the aggregated sample.

Regarding the cross-country analysis, performance expectancy has a higher impact across all countries. The Netherlands shows a high impact on hedonic motivation and social influence, while there is a negative impact and no significance on social status. Perceived safety remains a strong impact in the Netherlands, Belgium and Greece. Additionally, we controlled the model with socio-demographic, infrastructure and safety variables. We found that user intention increases with increasing age in all countries but not in Belgium. However, no significant effects were found for the rest variables.

The findings from this study are an important added value to the literature since it lacks user acceptance. In addition, it offers new insights into deploying new technologies on e-bikes and can benefit different stakeholders, such as bicycle manufacturers and cities. While bicycle manufacturers and designers of such innovative systems are investigating these features to bring them into the market, they can integrate these insights to optimise and develop them better. Also, cities can develop and implement new policies for these emerging technologies for a smooth transition.

ACKNOWLEDGEMENTS

This work is supported by Accell Group.

REFERENCES

- Boronat, Pérez-Francisco, Calafate, & Cano. (2021). Towards a sustainable city for cyclists: Promoting safety through a mobile sensing application. *Sensors (Switzerland)*, 21(6), 1-18. doi:10.3390/s21062116
- ECF. (2023). Money for bikes: Tax incentives and purchase premiums for cycling in Europe. Retrieved from <https://ecf.com/resources/financial-incentives>
- Fishman, & Cherry. (2016). E-bikes in the Mainstream: Reviewing a Decade of Research. *Transport Reviews*, 36(1), 72-91. doi:10.1080/01441647.2015.1069907
- Fornell, & Larcker. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50. doi:10.2307/3151312
- Gadsby, & Watkins. (2020). Instrumented bikes and their use in studies on transportation behaviour, safety, and maintenance. *Transport Reviews*. doi:10.1080/01441647.2020.1769227
- Hair, Black, Babin, & Anderson. (2014). *Multivariate data analysis* (Fifth Ed. ed.). Prentice Hall.
- Kapousizis, Ulak, Geurs, & Havinga. (2022). A review of state-of-the-art bicycle technologies affecting cycling safety: level of smartness and technology readiness. *Transport Reviews*, 1-23. doi:10.1080/01441647.2022.2122625
- Kapsler, & Abdelrahman. (2020). Acceptance of autonomous delivery vehicles for last-mile delivery in Germany – Extending UTAUT2 with risk perceptions. *Transportation Research Part C: Emerging Technologies*, 111, 210-225. doi:10.1016/j.trc.2019.12.016
- Nikitas, Tsigdinos, Karolemeas, Kourmpa, & Bakogiannis. (2021). Cycling in the Era of COVID-19: Lessons Learnt and Best Practice Policy Recommendations for a More Bike-Centric Future. *Sustainability (Switzerland)*, 13(9). doi:10.3390/su13094620
- Nordhoff, Louw, Innamaa, Lehtonen, Beuster, Torrao, . . . Merat. (2020). Using the UTAUT2 model to explain public acceptance of conditionally automated (L3) cars: A questionnaire study among 9,118 car drivers from eight European countries. *Transportation Research Part F: Traffic Psychology and Behaviour*, 74, 280-297. doi:10.1016/j.trf.2020.07.015
- Oliveira, Nery, Costa, Silva, & Lima. (2021). A survey of technologies and recent developments for sustainable smart cycling. *Sustainability (Switzerland)*, 13(6). doi:10.3390/su13063422
- Panwinkler, & Holz-Rau. (2021). Causes of pedelec (pedal electric cycle) single accidents and their influence on injury severity. *Accident Analysis and Prevention*, 154. doi:10.1016/j.aap.2021.106082
- Plazier. (2022). E-bikes in rural areas: current and potential users in the Netherlands. *Transportation*. doi:10.1007/s11116-022-10283-y
- Schepers, Fishman, Den Hertog, Wolt, & Schwab. (2014). The safety of electrically assisted bicycles compared to classic bicycles. *Accident Analysis and Prevention*, 73, 174-180. doi:10.1016/j.aap.2014.09.010
- Schepers, Twisk, Fishman, Fyhri, & Jensen. (2017). The Dutch road to a high level of cycling safety. *Safety Science*, 92, 264-273. doi:10.1016/j.ssci.2015.06.005
- Schumacker, & Lomax. (2010). *A Beginner's Guide to Structural Equation Modeling: Third Edition (3rd ed.)*. New York: Taylor and Francis Group, LLC.
- Shimano. (2022). *State of the Nation 2022*. Retrieved from <https://lifestylebike.shimano.com/uk/stories/shimano-state-of-the-nation-2022>
- Statista. (2020). Number of e-bikes sold in Europe from 2009 to 2019. In (pp. CONEBI).
- Statistics Netherlands (CBS). (2021). 610 traffic deaths in 2020. Retrieved from <https://www.cbs.nl/en-gb/news/2021/15/610-traffic-deaths-in-2020>
- Sutton. (2022). European electric bike sales pass 5 million, all bikes 22 million. *cyclingindustry.news*. Retrieved from <https://cyclingindustry.news/european-electric->

[bike-sales-pass-5-million-all-bikes-22-million/#:~:text=New%20data%20published%20this%20week,%25%20year%2Don%2Dyear.](#)

Venkatesh, Morris G, Davis B, & Davis D. (2003). User Acceptance of Information Technology: Toward a Unified View. 27. *No.3*, pp. 425-478 (454 pages). doi:<https://doi.org/10.2307/30036540>

Venkatesh, Thong, & Xu. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *JSTOR*, 36, 157-178. doi:<https://doi.org/10.2307/41410412>

Design hourly volume estimation at freeway nodes using floating car data

Magdalena Schilling^{*1}, Marvin V. Baumann², Jörg Sonnleitner¹, Markus Friedrich¹, and Peter Vortisch²

¹Institute for Road and Transport Science, University of Stuttgart, Germany

²Institute for Transport Studies, Karlsruhe Institute of Technology, Germany

SHORT SUMMARY

The estimation of the Design Hourly Volume (DHV) is an essential step for a traffic assessment. At freeway nodes, not all ramps are detected by permanent traffic count (PTC) stations. Therefore the German HCM recommends additional short-term counts (STC) to determine the DHV. Since conducting STC is usually associated with high effort, the question arises whether the information obtained by STC can also be derived from Floating Car Data (FCD). We propose an approach for processing the FCD in order to apply it instead of the STC for the determination of the DHV at ramp junctions. The performance of the method is evaluated on five nodes, for which FCD from 2017 and a reference database covering all 8,760 hourly volumes of all ramps and main lanes of the road section are available. The result show, the usage of representative FCD days is possible.

Keywords: Design Hourly Volume Estimation, Floating Car Data, Highway Capacity Manual, Permanent Traffic Counts

1 INTRODUCTION

Design hourly volume (DHV) estimation is an essential step when it comes to estimating the Level-of-Service (LOS) of traffic facilities. In the Highway Capacity Manual (HCM, National Academies of Sciences (2016)) and the German HCM (FGSV (2015)), the DHV is determined based on a traffic volume estimation concept known as the n^{th} hour or respectively the hour of the year with the n^{th} highest traffic volume. To calculate this n^{th} hour precisely, a permanent traffic count (PTC) station is necessary at the corresponding traffic facility, since the traffic volume for all 8,760 hours of the year must be known. To determine the DHV at ramp junctions of nodes, supplementary short-term traffic counts (STC) are also necessary since, usually, not all ramps are recorded with PTC. The German HCM proposes a method for this process, which was validated and enhanced to a concept of the n^{th} highest saturated hour by Baumann et al. (2023).

The method proposed by the German HCM determines the representative turning flows at the node from STC on the ramps, which are then extrapolated to the DHV using the PTC available at the node. Conducting STC is generally associated with high effort, which is why the question arises if the required representative turning flows can be determined using Floating Car Data (FCD). Ceccato et al. (2022) demonstrate that the use of FCD is competitive compared to traditional data sources in terms of cost-effectiveness. Furthermore, Vogt et al. (2019) and Dabbas et al. (2020) show that the data fusion of FCD and PTC enables the estimation of origin-destination matrices for motorway networks. Travel times and route choice probabilities derived from FCD can moreover be used as input for Dynamic Traffic Assignment models to map OD matrices to link flows (Nigro et al. (2018); Tsanakas et al. (2022)). Nohekhan et al. (2021) use FCD, temporary volume counts (e.g., a week), and road characteristics to estimate hourly traffic volumes on off-ramps. FCD can also be used to determine travel time (Olszewski et al. (2018)), free flow speed (Diependaele et al. (2016)), or operating speed (Bruwer et al. (2021); Lobo et al. (2018)) on motorways.

The literature review demonstrates that traffic flow assessment using FCD is possible, but to the best of the authors' knowledge, there are no approaches in the literature that use FCD to determine the DHV in the context of the HCM, German HCM, or similar international guidelines. Therefore, this paper examines the potential of using FCD as a substitute for STC in estimating the DHV at nodes. We propose an approach for processing the FCD and evaluate the results afterward using FCD of route sections with a total length of 15 km and compared with STC results from Baumann et al. (2023).

2 METHICAL APPROACH

Concept and data availability scenarios for DHV estimation at freeway nodes

An example of how to combine PTC and STC using topological relationships of a freeway interchange is shown in the German HCM. It assumes that a cloverleaf interchange has eight PTC, one for each inflow or outflow, and at least two STC for each ramp junction. Each PTC defines a specific demand-situation, which needs to be analyzed. In the following, these demand-situations are referred to as PTC demand-situation. Each PTC demand-situation describes a temporary state with consistent traffic flows at the entire interchange, such that inflow equals outflow. In the German HCM these demand-situations are defined using the method ‘50th hour of the PTC’. In this example it leads to eight demand situations, which may occur on different weekdays and times of day. STC are usually conducted at a different date. Therefore, the German HCM uses the day hour of the PTC demand-situation to derive a second demand-situation based on the STC (STC demand-situation). In the next step a matrix estimation method is applied using the PTC demand-situation as boundary conditions and the STC demand-situation as initial matrix to derive DHV for each count station. This procedure is repeated for all eight PTC demand-situations of the cloverleaf interchange. After that, all eight demand-situations are evaluated. For each ramp junction, a separate saturation rate is estimated per demand-situation. The resulting saturation rate of a ramp junction is the worst-case saturation rate of all demand situations considered. For this concept for DHV estimation at freeway nodes we compare the usage of traffic flows derived from FCD instead of STC. Furthermore, the number of PTC can be varied. To understand the impact of these data sources, the following data availability scenarios are defined and will be analyzed:

1. ‘PTC: in-/outflow main lanes, STC: -’: This data availability scenario again uses no STC data, but the numbers of PTC stations are reduced to one count station for each inflow or out-flow on the main lanes of the node. This leads to eight PTC stations at a four-leg-interchange and to four PTC stations at a freeway exit.
2. ‘PTC: in-/outflow main lanes, STC: all’: This data availability scenario adds STC information for all counting stations (Figure 4, third row).
3. ‘PTC: in-/outflow main lanes, FCD: all’: This data availability scenario equals the data availability scenario above but uses FCD information instead of STC.
4. ‘PTC: in-/outflow main lanes, FCD: representative traffic days’: This data availability scenario uses representative traffic days obtained from FCD instead of single days.

Data basis

Hourly FCD-hits are available for 164 working days in 2017 on route sections with a total length of 15 km. ‘Hits’ refer to the number of vehicles recorded. The number of hits is a subset of the total traffic volume. The route sections are part of 63 ramp junctions for whom PTC are available. Missing hits on some count stations are derived based on adjacent hits. The ramp junctions are part of five nodes: two interchanges and three freeway exits.

Preprocessing to generate consistent FCD for all count stations

An initial plausibility check reveals some inconsistencies:

- There are some negative values as results of balancing checks considering the topology.
- Conflicts appear if several FCD route sections can be assigned to the same count station, leading to an over determination of some count stations by FCD route sections.

Thus, a matrix estimation procedure is executed to get consistent FCD using the VFlowFuzzy algorithm implemented in PTV Visum (PTV AG (2022)). This implementation allows to define tolerances for each count station in case the hits are too inconsistent. For each hour the related hourly hits are considered and tolerances are increased successively until a solution of consistent hits is found. If no solution is found, it is analyzed which count station has implausible hits. This allows us to ignore these values or increase the tolerance for these count stations.

Generate representative traffic days

In order to get more robust daily hits, we use the amount of FCD days to generate representative traffic days. A representative traffic day has a daily distribution, which occurs as often as possible in this or a similar way.

A method suitable for matching the properties of daily distributions is cluster analysis.

We define three types of traffic days:

- Monday
- Tuesday, Wednesday and Thursday
- Friday

For the cluster analyses of a traffic day type we use network load curves containing daily hits of all count stations for the days belonging to the traffic day type. Depending on the size of the distance measure and the clustering algorithm, there will be a different number of clusters and consequently a different distribution of days per cluster.

We use the average linkage cluster algorithm and the GEH value serves as distance measure for the comparison of hourly traffic volumes. The cluster containing most days is defined as main cluster. The smaller the maximal allowed distance, which is given as input by the user, the more clusters there are. However, the days in a cluster are more similar with a lower distance measure, the main cluster is more characteristic and it is less influenced by smoothing due to averaging of more divergent days.

We choose GEH 6, resulting in main cluster that represent about 50% of the days belonging to the traffic day type (table 1).

Table 1: result of cluster analysis using GEH-value ≤ 6 .

traffic day	days total	number of clusters	days in main cluster	share of days in main cluster
Mo	32	9	18	56%
TuWeThu	96	22	51	53%
Fri	34	9	16	47%

DHV estimation combining FCD and PTC on nodes

To adapt the DHV estimation at nodes as described above to representative FCD days, we calculate for each PTC demand-situation scenarios for all representative days. This leads to three scenarios per PTC. The day hour of the PTC demand-situation defines the hour of the representative day. Then the day hour of the representative day is used to get the traffic flow matrix. After that, the process using matrix estimation and determination of the worst-case saturation is the same.

3 EVALUATION

Combining PTC and FCD enables the DHV estimation for each ramp junction of a node. In this study, the DHV is defined as the 50th hour as it is the common standard in Germany. Since a consistent determination of the 50th hour based on the traffic volume is not possible at ramp junctions due to several traffic flows, the 50th highest saturated hour (calculated according to the methods of the German HCM) is used instead of the hour with the 50th highest traffic volume (Baumann et al. (2023)).

For all nodes a reference database is available that provides all 8,760 hourly volumes of 2017 for all ramps and main lanes (Baumann et al. (2023)). The reference database enables the calculation of the saturation of each ramp junction for all hours. So the actual 50th highest saturated hour can be determined for each ramp junction, which is referred to as the reference scenario in the following. In order to analyze the performance of the proposed method, the method is applied

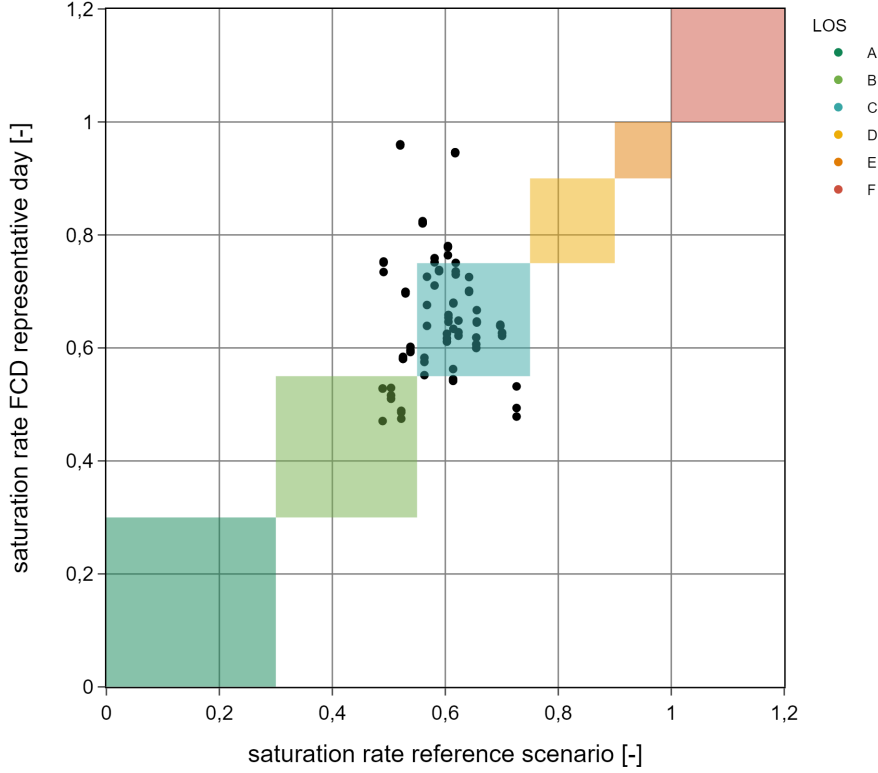


Figure 1: Exemplary results of a node for an estimation scenario using FCD representative days.

to all ramp junctions of the nodes considered, and for each ramp junction, the obtained results (referred to as estimation scenario in the following) are compared with the reference scenario.

Figure 1 shows the correlation between reference and estimation scenario exemplary for all ramp junctions of a node. Each point represents the result of one ramp junction. The colored squares illustrate the corresponding LOS according to the German HCM. If a point is located in one of the squares, the saturation rate of the estimation scenario results in the same LOS as that of the reference scenario. Otherwise, the estimation scenario differs from the target LOS of the reference scenario. For a further aggregation of the results, we introduce the metric of the ‘average LOS-accuracy’. Based on the results of the estimation scenario, this metric describes the relative share of the estimation scenario that achieves the target LOS of the reference scenario. Regarding the visualization in fig. 1, this corresponds to the proportion of points located within one of the colored LOS squares.

$$\text{average LOS-accuracy} = \frac{n(LOS_{es} = LOS_{rs})}{n} \quad (1)$$

with

$$\begin{aligned} n &= \text{number of estimation scenarios} \\ LOS_{es} &= \text{calculated LOS of estimation scenario} \\ LOS_{rs} &= \text{LOS of corresponding reference scenario} \\ n(LOS_{es} = LOS_{rs}) &= \text{number of estimation scenarios, which hit the LOS of their} \\ &\quad \text{reference scenario} \end{aligned}$$

4 RESULTS AND DISCUSSION

The aggregated results for 63 ramp junctions on the several data availability scenarios are shown in table 2, fig. 2 allows a less aggregated and more detailed view. Figure 2 shows the range of saturation for the STC and FCD scenarios for some ramp junctions. Additionally the results of the reference scenario and for the data availability scenario considering clusters as representative traffic days are included in the figure.

The results lead to the following conclusions:

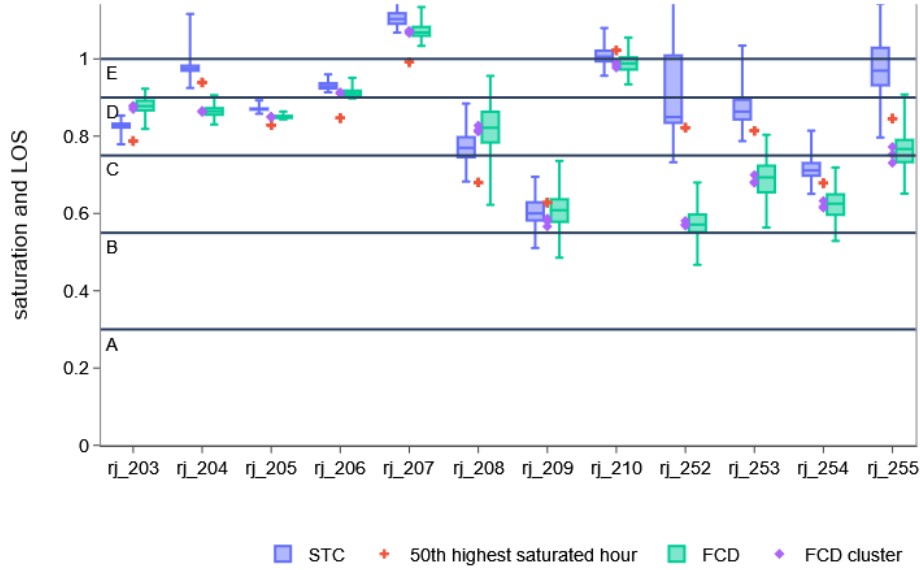


Figure 2: saturation using different data availability scenarios for some ramp junctions.

Table 2: LOS-accuracy for different data availability scenarios.

data availability scenario	LOS-accuracy	LOS-under-estimation	LOS-over-estimation	number of ramp junctions
'PTC: in-/outflow main lanes, STC: -'	33%	2%	65%	63
'PTC: in-/outflow main lanes, STC: all'	87%	3%	10%	63
'PTC: in-/outflow main lanes, FCD: all'	73%	10%	17%	63
'PTC: in-/outflow main lanes, FCD: representative traffic days'	78%	10%	12%	63

- STC or FCD on all ramps are crucial for an accurate estimation of the DHV.
- The method recommended by the German HCM - conducting STC at all ramps of a node with subsequent matrix correction or extrapolation at the nearest PTC stations - provides a data basis with which the hour with the 50th highest saturation is well met.
- Nevertheless, the saturation (and LOS) varies, depending on the day on which the STC is conducted.
- In this context, it must be taken into account that for the STC considered, it is assumed for reasons of convenience that the results of STC and PTC derive from the same year, while for practical reasons the STC is often conducted a year after the year of the PTC.
- The proposed FCD method underestimates the traffic flow on ramps in some cases.
- Using representative days slightly improves the LOS-accuracy.

For further research it would be interesting to expand the approach of using FCD as introduced in this paper to additional nodes or data sources. Further evaluation of the robustness of clustering to

get representative traffic days is necessary. Furthermore, it can be expected that the performance of the proposed method will increase in the future, as the availability of FCD will improve and will thus lead to more representative results.

ACKNOWLEDGEMENTS

This paper was based on research sponsored by the German Federal Ministry for Digital and Transport, represented by the Federal Highway Research Institute.

REFERENCES

- Baumann, M. V., Schilling, M., Friedrich, M., Reichert, S., Vortisch, P., & Wassmuth, V. (2023). Design hourly volume estimation at freeway nodes from short-term counts. (*unpublished manuscript, under review*).
- Bruwer, M. M., Andersen, S. J., & Merrick, W. (2021). Measuring the impact of roadworks on traffic progression using floating car data. Southern African Transport Conference 2021.
- Ceccato, R., Gecchele, G., Rossi, R., & Gastaldi, M. (2022). Cost-effectiveness analysis of Origin-Destination matrices estimation using Floating Car Data. Experimental results from two real cases. *Transportation Research Procedia*, 62, 541–548. (Publisher: Elsevier)
- Dabbas, H., Fourati, W., & Friedrich, B. (2020). Floating car data for traffic demand estimation-field and simulation studies. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1–8). IEEE.
- Diependaele, K., Riguelle, F., & Temmerman, P. (2016). Speed behavior indicators based on floating car data: results of a pilot study in Belgium. *Transportation research procedia*, 14, 2074–2082. (Publisher: Elsevier)
- FGSV. (2015). *Handbuch für die bemessung von straßenverkehrsanlagen: Hbs* (Vol. 299 B). FGSV-Verlag.
- Lobo, A., Amorim, M., Rodrigues, C., & Couto, A. (2018). Modelling the operating speed in segments of two-lane highways from probe vehicle data: a stochastic frontier approach. *Journal of Advanced Transportation*, 2018. (Publisher: Hindawi)
- National Academies of Sciences. (2016). *Highway capacity manual: Hcm* (6th ed.). Washington, D.C..
- Nigro, M., Cipriani, E., & del Giudice, A. (2018). Exploiting floating car data for time-dependent Origin–Destination matrices estimation. *Journal of Intelligent Transportation Systems*, 22(2), 159–174. (Publisher: Taylor & Francis)
- Nohekhan, A., Zahedian, S., & Haghani, A. (2021). A deep learning model for off-ramp hourly traffic volume estimation. *Transportation Research Record*, 2675(7), 350–362. (Publisher: SAGE Publications Sage CA: Los Angeles, CA)
- Olszewski, P., Dybicz, T., Jamroz, K., Kustra, W., & Romanowska, A. (2018). Assessing highway travel time reliability using Probe Vehicle Data. *Transportation Research Record*, 2672(15), 118–130. (Publisher: SAGE Publications Sage CA: Los Angeles, CA)
- PTV AG. (2022). *Ptv visum: Manual*. Karlsruhe.
- Tsanakas, N., Gundlegård, D., & Rydergren, C. (2022). O–D matrix estimation based on data-driven network assignment. *Transportmetrica B: Transport Dynamics*, 1–32. (Publisher: Taylor & Francis)
- Vogt, S., Fourati, W., Schendzielorz, T., & Friedrich, B. (2019). Estimation of origin-destination matrices by fusing detector data and Floating Car Data. *Transportation research procedia*, 37, 473–480. (Publisher: Elsevier)

Mobility Service Providers' Equilibrium Strategies in Multi-modal Networks

Claudia Bandiera*¹, Richard D. Connors², and Francesco Viti³

¹PhD student, Faculty of Science, Technology and Communication, University of Luxembourg, Luxembourg

²Senior Research Fellow, Faculty of Science, Technology and Communication, University of Luxembourg, Luxembourg and Institute for Transport Studies, University of Leeds, United Kingdom

³Associate Professor, Faculty of Science, Technology and Communication, University of Luxembourg, Luxembourg

SHORT SUMMARY

Due to the increasing introduction of new mobility solutions in the transport offers, the market equilibrium among Mobility Service Providers (MSPs) has become more complex. The focus of this paper is to develop a novel analytical approach to study competition and/or cooperation between multiple MSPs within a multi-modal network system. We formulate a novel Equilibrium Problem with Equilibrium Constraints (EPEC), where each MSP seeks to maximize their own profits at the upper level. At the lower level, users are divided into classes that capture their heterogeneity in terms of socioeconomic characteristics and activity-travel behaviour. We consider the multi-modal network link costs to be non-separable, therefore the lower-level equilibrium is formulated as a Variational Inequality (VI) problem. A solution approach is proposed and illustrated, based on a relaxation of the Diagonalization method. Finally, we apply the described methodology to a small example to illustrate some key properties of the proposed approach.

Keywords: EPEC, Multi-modal Network, Supernetwork, Variational Inequality

1 INTRODUCTION

In recent years, transportation systems have been offering travellers an increasing number of multi-modal options thanks to the introduction of new mobility solutions, such as ride hailing, shared and pooled mobility services, micromobility, on-demand services, etc. Consequently, the market equilibrium among Mobility Service Providers (MSPs) has become more complex, with different competitive or cooperative strategies being observed with the aim of attracting a sufficient share of customers and hence sustain a profitable business. In such scenario, users' modal choices are crucial in determining the durability of mobility services within the transportation system.

In Transport Network Design, the problem of studying the relationships between MSPs and users has been traditionally focused on uni-modal networks. Reflecting the complexity found in real transportation systems and individuals' mobility, more recently attention has been focused on multi-modal networks (Zhang et al., 2014). Although these works developed relatively complex models, there is limited research that includes multiple leaders, coexisting, competing or cooperating. From an economic and strategic point of view, it is essential to model the interactions between MSPs and travellers of the transport network to predict the response of these actors as a consequence of the variation in strategies of the entire system. In particular, scenarios offering a new transport service, introducing new regulations/incentives, or increasing users' heterogeneity, could substantially change the equilibrium of the whole network.

In the literature, few works have analyzed this type of problems. In the context of fast charging stations for electric vehicles, Guo et al. (2016) developed a Multi-agent Optimization Problem with Equilibrium Constraints (MOPEC)-based model to study interactions between multiple competitive investors and travellers assigned to a congested transport network. In cordon toll competition, Watling et al. (2015) formulated an Equilibrium Problem with Equilibrium Constraints (EPEC), through which they study the competing behaviour of two different public authorities from two cities aiming to maximize the social welfare of the corresponding residents. Yang et al. (2022), instead, defined a bi-level model to optimize pricing and relocation in a competitive one-way car-sharing market.

Albeit the above-cited works establish important developments in their respective area of application, all of them consider that users are homogeneous and assigned to networks in which the competition between suppliers is limited to uni-modal markets.

In this paper we developed a novel analytical approach to study competition and/or cooperation between multiple MSPs, while users with heterogeneous characteristics are assigned to a multi-modal network. We formulate a novel EPEC, where each MSP seeks to maximize their own profits at the upper level; the objective functions include MSP-specific costs and revenues. At the lower level, users are divided into classes that capture their heterogeneity in terms of socioeconomic characteristics and activity-travel behaviour, resulting in different daily trip chains. Due to the non-separability of the link cost functions, the lower-level equilibrium assignment is formulated as a Variational Inequality (VI) problem. The proposed methodology is applied to a small example to show key properties of the model, using an iterative solution approach, based on a relaxation of the Diagonalization method.

2 METHODOLOGY

This study aims to develop models for the economic assessment of different suppliers' strategies in a multi-modal network. Hence, the interactions between MSPs and users are modelled and illustrated using the concept of supernetworks (Sheffi, 1985). This representation has been adopted in the literature to tackle the complexity of multi-modal networks, being further expanded to include a time component connected to users' activity-based trip chains (Fu & Lam, 2014). Our methodology is applied to a static system in which time of departure/arrival from/to a location or duration of the activities performed at each destination are not considered in full detail, given the strategic, economical purpose of the developed model.

In order to build this supernetwork, we use endogenous information regarding users and MSPs. To represent a more realistic transportation system, travellers are divided into classes based on their socio-economic attributes and the trips' purpose; considering that based on the socio-economic characteristics users associate different costs to the travel time experienced in the network. The sequence of trips made in a day, their purpose and locations, instead, are included to explicitly represent the link between daily modal choices of users. It is here argued that each travel choice made by users is influenced by earlier decisions as well as by planned later trips during a day. We therefore define users of a certain class those performing a type of activity sequence (e.g. homework-leisure-home) in the same sequence of zone(s). Assuming that, during an ordinary weekday, users of the same class $k \in K$ perform the same sequence of trips. This sequence is then modelled as a directed graph, where a node $n \in N$ corresponds to a zone and a link $a \in A$ indicates a trip from one zone to the next (top Figure 1). The first and last location visited by a class of users represent that class's origin (O) and destination in the network (D). We assume that a single link can directly symbolize a trip without the need of defining all the different path alternatives present in the real network. Considering that, after computing a traffic assignment process, the network reaches equilibrium for which all used routes have the same generalised cost (Wardrop, 1952). We use then the information regarding MSPs to take into account the modes of transport available at each trip connection, and we expand the network into uni-modal layers (colored parallel networks in Figure 1). Each layer is owned by a specific MSP $j \in J$ that collects revenues based on how many travellers use their service (i.e. link flows), and accrue costs primarily depending on the size of their vehicle fleet v_j (capacity). We formulate the MSP profit objective function to be sufficiently general to describe different mobility services, such as car-sharing, bike-sharing, bus, train, e-scooter and taxi. The continuous upper level decision variables are controlled by the MSP fleet sizes $\mathbf{v} \in V \subseteq R^{|J|}$. MSPs decide how to strategically distribute these vehicles amongst the links of their network layer, with the purpose of maximizing their profit.

We consider each user class to be assigned to the multi-modal network following a fixed demand-based traffic equilibrium, using a path-based adaptation of the multi-class and multicriteria network equilibrium model (Nagurney, 2000). In particular, there will be cost (and revenue) components that vary with the usage of the service, and class-based cost components associated to the time spent using the service). Please see previous works (Bandiera et al., 2022a,b) for further details concerning the mathematical representation of MSPs and users, and previous results.

Given the non-separable nature of the network link cost functions, explicit path enumeration is used in this paper. While this approach is not yet readily applicable for large-scale networks, our current focus is to develop and understand the methodology presented here, for which small networks involving a limited number of multi-modal options for each class are sufficient. Undoubtedly, the

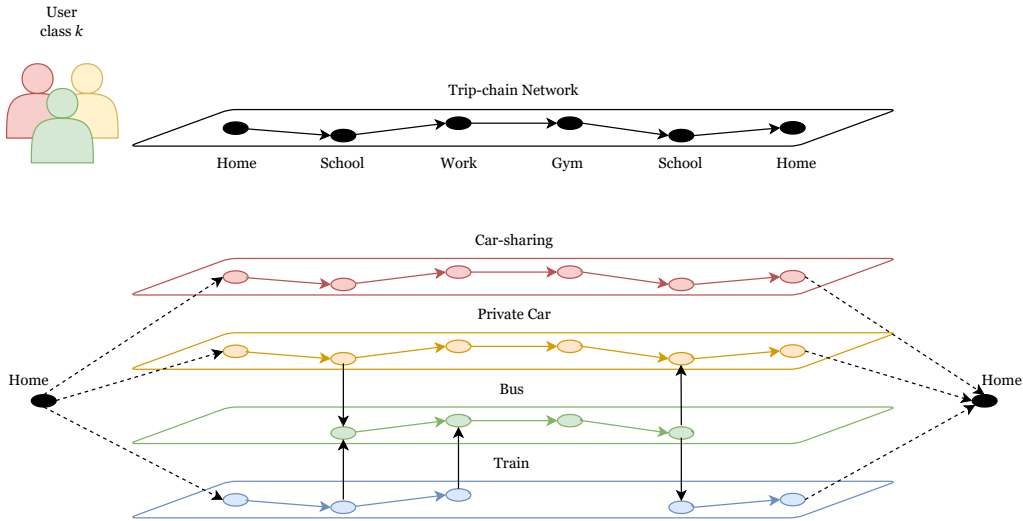


Figure 1: Multi-modal Trip-chain Supernetwork

choice of dividing users into classes increases the complexity of the model. However, the socio-economic characteristics are only affecting the cost perceived by users of a specific class and not the network expansion. On the other hand, defining users' classes based on the combination of daily trip chains and activity locations could be a non-trivial problem when considering large-scale networks. However, literature on the topic (Axhausen et al., 2002) shows that during typical weekdays the majority of travellers tend to perform home-work-home tours when using public transport or add an additional activity before/after work when travelling with private vehicles. Moreover, the combination of trip chains, activity sequences and locations are spatially limited and rather repetitive. Therefore, focusing on the most frequent tours, we cover most of travel the demand of an area.

The lower level equilibrium decision variables are the path flows represented by the vector $\mathbf{x} \in X \subseteq \mathbb{R}^z$, with z paths and X the set of demand-feasible flows. Travellers choose a path through the multi-modal network in order to perform their sequence of trips. A path can comprise three different types of links. In line with the economic assessment purpose of the model, links represent the main mode of transport connecting two zones. Access links (black dashed lines in Figure 1) allow users to access a mode of transport from their origin (Home), and egress from a mode of transport to reach their final destination (Home). These links play an important role inside the network, capturing costs related to a monthly subscription for a single service or a package containing a combination of them. Mode-specific links (horizontal links), instead, indicate trips made from one location to another using a specific mode of transport (designated by colour). These link costs include three stages of a trip: 1) accessing the selected mode of transport from the departure node; 2) travelling using the main mode of transport; 3) egress from the selected mode of transport and reaching the destination. Finally, interchange links (vertical black links) allow users to move from one mode of transport to another.

Figure 2 shows cost components for different links, which characterize each mode of transport. The lower level problem is complicated by the presence of multiple classes at the lower-level and the interdependency between flows on parallel links of the supernetwork. Concretely, some supernetwork links represent copies of the same real transport link of the underlying infrastructure network e.g. travel time on car-sharing links is influenced by travellers using private car and vice versa. Consequently, the corresponding link costs are non-separable. For this reason the users' equilibrium is formulated as a VI (Dafermos, 1980).

Let $C(\mathbf{x}, \mathbf{v})$ be the path cost function for the lower level problem, which depends on the capacities, \mathbf{v} , supplied by MSPs. Then a vector of path flows $\mathbf{x}^* \in X$ is a Wardrop equilibrium if and only if it satisfies the VI problem:

$$\langle C(\mathbf{x}^*, \mathbf{v}), \mathbf{x} - \mathbf{x}^* \rangle \geq 0 \quad \forall \mathbf{x} \in X \quad (1)$$

Given MSP fleet sizes \mathbf{v} , we denote the set of equilibrium solutions $X^*(\mathbf{v})$.

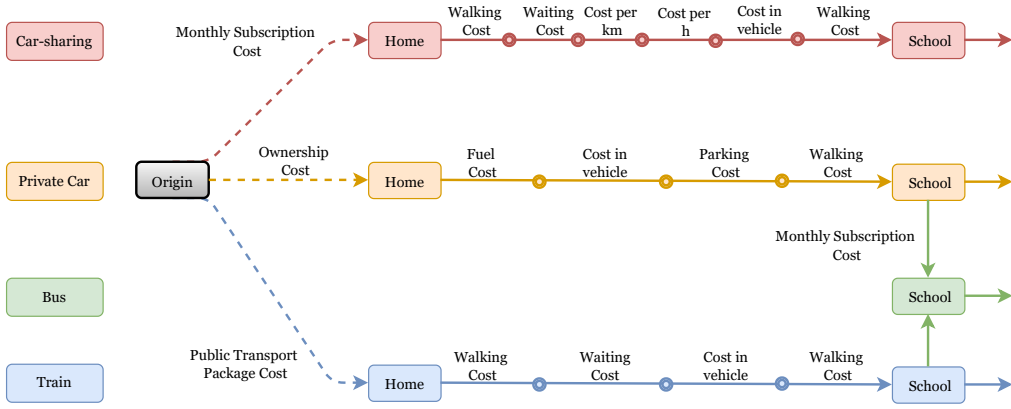


Figure 2: Example of detailed costs from Figure 1

Problem formulation and solution algorithm

We formulate the interaction between MSPs and users as an EPEC. An equilibrium at the upper level corresponds to no MSP wishing to unilaterally change their fleet size, given that the lower level satisfies the Wardrop equilibrium conditions.

Each MSP $j \in J$ seeks to maximize their profit, which is given by a continuous differentiable function, $p_j(v_j|\mathbf{x})$, depending on their fleet size, v_j , and on the path flows. The equilibrium path flows depend on the vector of MSP fleet sizes, so with the lower level constraint in place we have $p_j(v_j|\mathbf{x}^*(\mathbf{v}))$. Each MSP can change only their own fleet size: for a vector \mathbf{v} we denote a change in only the j -th component by $\mathbf{v}_{[j]}$. Collecting the profit functions into $\mathbf{p} = [p_j]$ (adopting the obvious vector notation), \mathbf{v}^* is an equilibrium solution if and only if

$$\mathbf{p}(\mathbf{v}^*|\mathbf{x}^*) - \mathbf{p}(\mathbf{v}_{[j]}^*|\mathbf{x}^*) \geq \mathbf{0} \quad \forall j \in J \quad (2)$$

$$\text{with } \mathbf{x}^* \in X^*(\mathbf{v}^*) \quad (3)$$

where the lower level equilibrium path flows $X^*(\mathbf{v})$ are defined above (see Equation 1).

EPEC problems are well known in literature for the difficulty of finding equilibrium solutions. In this paper, we solve the EPEC using the Diagonalization Method: applying iteratively at the upper level a minimization approach while at the lower level we calculate the equilibrium solution using the Extragradient Method (EM), often used in the context of traffic assignment (Nagurney, 2000). Due to challenge of achieving convergence of the general EPEC, relaxation approaches are usually applied to a standard Diagonalization method. Here we introduce a steplength variation scheme in the iterations of the upper-level problem, based on the Self-Regulated Averaging Method proposed by Liu et al. (2007).

3 EXAMPLE AND DISCUSSION

In this section, we showcase the implementation of the methodology to a small network. We consider two classes of users performing two different trip chains, described at the top of Figure 3. The example may represent a user class performing a home-work-home tour, and a second class performing a work activity in the same zone, but also chaining another activity in another zone before returning home. We consider that each trip connection is covered by four modal options: one-way car-sharing 1, one-way car-sharing 2, a transit service, and private car is available to all users. The multi-modal network resulting graphically representing the four services and the two user classes consists of 32 nodes and 40 links (Figure 3). For each OD, users can therefore choose from four paths, based on the modal options available. For simplicity, in this example we do not allow travellers to use more than one mode of transport. However, this simplification is not a restriction of the model nor of the solution algorithm.

Congestion effects inside the network are modelled using the conventional Bureau of Public Roads (BPR) function, considering that users choosing the car-sharing services, or the private car, experience a travel time that is influenced by the presence of the other modes of transport on the same infrastructure. The transit service, instead, is considered to have a dedicated lane throughout the network, hence this service is affected only by the number of public transport users.

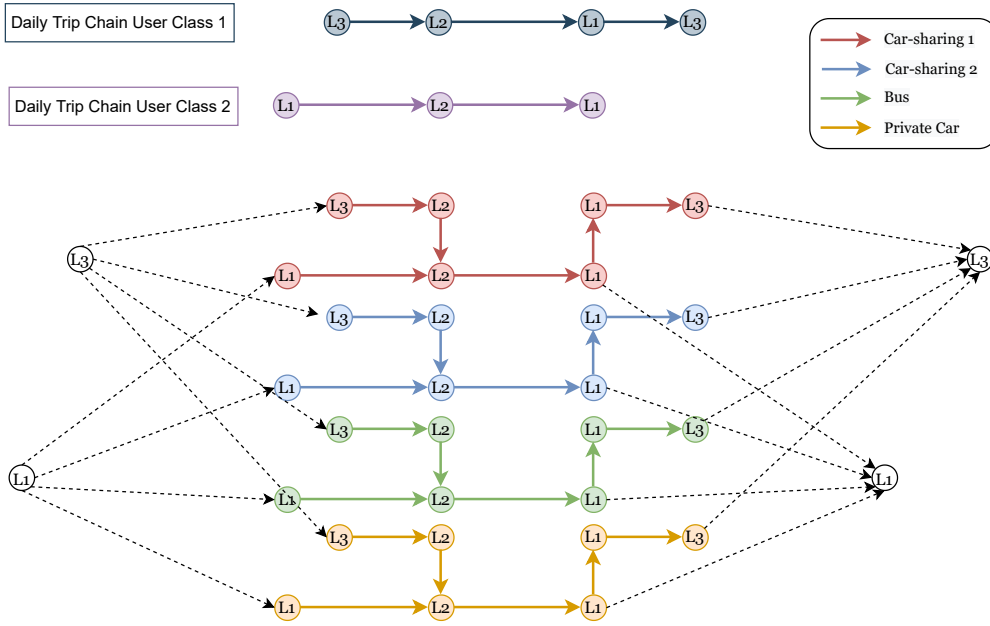


Figure 3: Example competition between MSPs

In this example, we try to understand the behaviour of the two car-sharing services competing inside the multi-modal network, applying the solution algorithm described in Section 2. We consider the two MSPs to adopt different price strategies: Car-sharing supplier 1 sells a cheaper monthly package compared to Car-sharing supplier 2, whereas the latter charges users a cheaper rate per hour and per kilometer.

Here we focus on the properties of the EPEC and the algorithm. To do so, we computed the full objective function (profit) surface for each MSP over a range of fleet sizes. This is computationally expensive and impractical in most cases, but allows us to verify solutions proposed by the algorithm. Figure 4 shows the objective functions surfaces of the two car-sharing suppliers. We examine whether the algorithm converges, and where to, starting from different initial conditions for the fleet sizes. In each case, the algorithm converges to the same point, indicated in magenta. Note that at the solution, Figure 4 left hand plot shows MSP1 profit is maximised (with v_2 fixed, varying v_1) and similarly, MSP2 profit is maximised in the right hand plot (fixing v_1). In this scenario, it seems that the car-sharing 2, offering a cheaper fixed price for the package, manages to attract more users, with a bigger fleet size ($v_1 = 74; v_2 = 82$) and more profitable service ($p_1 = 421.5; p_2 = 538$). In Figure 4 we also indicate on the axis the best strategy for each MSP, which occurs in both cases when their competitor is not operating in those zones.

It is also interesting to observe the variation of the total travel cost for the lower level (Figure 5). Obviously, an increase in the fleet size of both suppliers is translated in a reduced travel cost for users, but this is not economically feasible for the MSPs. It is particularly interesting to see by looking at the total cost surface that when the fleet size is sufficiently large, only one of the two suppliers could survive in the market.

The proposed methodology shows promising results on the proposed network. The aim of future developments is to expand its application to bigger networks considering the competition and cooperation between multiple suppliers at the upper level with multiple user classes at the lower level. Through this approach it will be possible to study different dynamics that occur in the transportation network due to the presence of heterogeneous actors with diverse purposes.

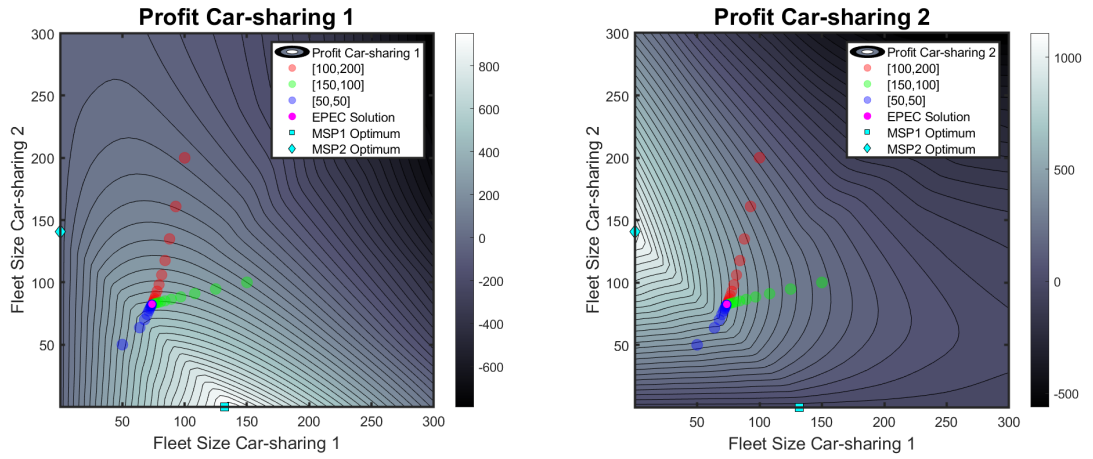


Figure 4: Profits variation with fleet sizes and equilibrium solution

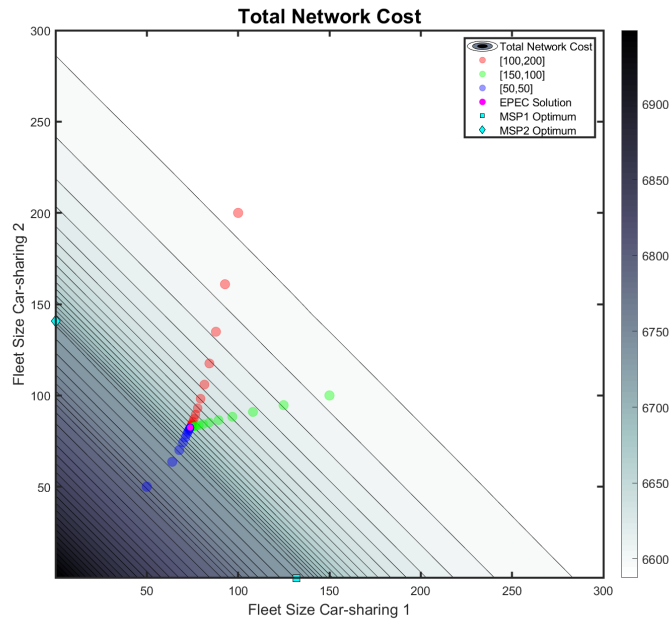


Figure 5: Total travel cost variation with fleet sizes

ACKNOWLEDGEMENT

This research is part of the project MaaS4All funded by Fonds National de la Recherche Luxembourg (GRANT NUMBER: 13769009).

REFERENCES

- Axhausen, K. W., Zimmermann, A., Schönfelder, S., Rindsfuser, G., & Haupt, T. (2002). Observing the rhythms of daily life: A six-week travel diary. *Transportation*, 29(2), 95–124.
- Bandiera, C., Connors, R. D., & Viti, F. (2022a). *Competition and Cooperation between Suppliers in Multimodal Network Design Problems*. https://tristan2022.org/Papers/TRISTAN_2022_paper_2182.pdf. ([Online])

- Bandiera, C., Connors, R. D., & Viti, F. (2022b). *Evaluating Mobility Service Providers' Strategies in an Activity-Based Supernetwork*. <https://transp-or.epfl.ch/heart/2022/abstracts/269.pdf>. ([Online])
- Dafermos, S. (1980). Traffic equilibrium and variational inequalities. *Transportation Science*, *14*, 42-54.
- Fu, X., & Lam, W. H. K. (2014). A network equilibrium approach for modelling activity-travel pattern scheduling problems in multi-modal transit networks with uncertainty. *Transportation*, *41*, 37-55.
- Guo, Z., Deride, J., & Fan, Y. (2016). Infrastructure planning for fast charging stations in a competitive market. *Transportation Research Part C-emerging Technologies*, *68*, 215-227.
- Liu, H. X., He, X., & He, B. (2007). Method of successive weighted averages (mswa) and self-regulated averaging schemes for solving stochastic user equilibrium problem. *Networks and Spatial Economics*, *9*, 485-503.
- Nagurney, A. (2000). A multiclass, multicriteria traffic network equilibrium model. *Mathematical and Computer Modelling*, *32*, 393-411.
- Sheffi, Y. (1985). *Urban transportation networks* (Vol. 6). Prentice-Hall, Englewood Cliffs, NJ.
- Wardrop, J. G. (1952). Road paper. some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, *1*(3), 325-362.
- Watling, D. P., Shepherd, S. P., & Koh, A. (2015). Cordon toll competition in a network of two cities: formulation and sensitivity to traveller route and demand responses. *Transportation Research Part B-methodological*, *76*, 93-116.
- Yang, S., Wu, J., Sun, H., Qu, Y., & Wang, D. Z. W. (2022). Integrated optimization of pricing and relocation in the competitive carsharing market: A multi-leader-follower game model. *Transportation Research Part C: Emerging Technologies*.
- Zhang, L., Yang, H., Wu, D., & Wang, D. (2014). Solving a discrete multimodal transportation network design problem. *Transportation Research Part C: Emerging Technologies*, *49*, 73-86.

Investigating the preferences for the use of urban ridepooling

Thomas Schatzmann^{*1}, Felix Zwick², and Kay W. Axhausen¹

¹Institute for Transport Planning and Systems, ETH Zürich, Switzerland

²MOIA GmbH, Germany

SHORT SUMMARY

This study investigates the preferences for the use of the urban ridepooling service MOIA in Hamburg, Germany. A survey with over 4,000 (non-)users was conducted and a discrete choice model was estimated to understand users' preferences to use the service. The study provides insights into the sociodemographic characteristics of ridepooling users, their preferences towards the service and first findings on the preferences towards an intermodal combination with public transportation. The results show that factors such as travel cost, time, trip distance and purpose are significant in influencing the use of ridepooling services. According to the choice experiment, intermodal travel is a viable choice for trip distances above 10 km, primarily for public transport subscription holders. The findings of this study can inform the design and marketing of future ridepooling services, and contribute to the broader debate on the potential benefits and challenges of shared mobility services in improving urban mobility and reducing the negative impacts of transportation on the environment.

Keywords: ridesharing, shared mobility, public transport, discrete choice modeling, intermodality, multimodality

1 INTRODUCTION

Urban ridepooling has emerged as a popular mobility solution, particularly in urban areas, as it provides a more efficient and affordable way for individuals to travel short to medium distances. Ridepooling is a service that allows multiple passengers who are traveling in the same direction to share a vehicle, thus reducing the number of vehicles on the road and decreasing congestion, noise and greenhouse gas emissions (Shaheen & Cohen, 2018; Zwick et al., 2021). In this paper, we investigate the preferences for the use of urban ridepooling on the example of the MOIA service.

MOIA is a ridepooling service launched in 2017 by Volkswagen Group, which uses electric vehicles that can carry up to six passengers. It operates the largest European ridepooling fleet with over 250 vehicles in Hamburg, Germany. The recent service expansion with a larger service area (270 km² instead of 200 km²), integration of wheelchair accessible vehicles and tariff integration into the public transport (PT) system as part of a funding project served as an occasion for a scientific long-term monitoring of these measures, in the context of which this research work took place. We surveyed 4,167 MOIA users and non-users in October and November 2022 to understand their sociodemographics and general mobility behavior, and estimated a discrete mode choice model to analyze their mode preferences. Specifically, we examine how factors such as age, gender, income, and travel distance influence individuals' choices between ridepooling, private cars, PT, and slow modes of transportation. The survey was conducted *before* the introduction of the described measures in January 2023 and will be repeated in fall 2023.

The study builds upon previous scientific investigations of the use of MOIA as part of MOIA's accompanying study by Karlsruhe Institute of Technology and TU Munich from 2019 to 2021 (Kagerbauer et al., 2021). Kistorz et al. (2021) reported the findings of the survey of over 12,000 MOIA (non-)users in 2019. They found that MOIA is used across all age groups and genders, and enriches multimodal travel behavior. A detailed investigation of users with mobility impairments and work-related trips was conducted. In contrast to the previous study, we estimate a discrete choice model and specifically investigate the intermodal use of ridepooling. The intermodal use of ridepooling was also investigated by Diebold et al. (2021) on the example of ioki in Hamburg. The

service is designed differently to MOIA and delivers customers for a small fee of 1€ in addition to a PT ticket to the next railway station in a 15 km² service area. Thus, it is not surprising that the share of intermodal trips is very high at 72% and that the average age is with 34 rather young.

We contribute to the literature on ridepooling by examining the preferences for the use of urban ridepooling, and identifying the factors that influence individuals' willingness to use these services. The estimated discrete choice models provide novel insights into important level-of-service attributes and values of travel time (VTT) that inform policy-makers and urban transportation planners in their efforts to promote sustainable and efficient transportation systems through new digital and smart mobility services.

2 METHODOLOGY

This study employed an online two-stage approach using a combination of revealed and stated preference surveys (RP & SP), which were administered to participants. RP data provide valuable information about mode choices in real markets, but often lack in variability of the underlying variables to construct appropriate models and forecasts (Ortúzar & Willumsen, 2011). To better comprehend the trade-off confronted by individuals in choosing between multiple modes, SP methods such as stated choice experiments (SCE) have been used as these are often richer in trade-off information by design (Louviere et al., 2003; Train, 2009). Hence, the mode choice experiment presented in the SP survey made use of individualized reference trips from data gathered in the RP survey. A pooled RP-SP Multinomial Logit (MNL) model was then applied to examine the influence of level-of-service (LOS) and sociodemographic attributes on the choice between the alternatives presented. In addition, population weighted willingness-to-pay indicators were derived to complement the analysis of mode choices in a multimodal setting.

Recruitment and survey design

Given that MOIA operates predominantly in the city of Hamburg, we focused on participants residing in the city and its surrounding areas. The recruitment process included two distribution channels, with current MOIA members being internally recruited by MOIA's marketing department and other respondents being externally sourced via two regional panel providers. Ultimately, the sample size for the RP survey amounted to 4,167 individuals. Comprehensive data cleaning and participants not filling out the SCE reduced the sample size to 3,823 individuals for the SP survey. Both surveys were conducted using the survey software Qualtrics.

The RP survey entailed questions about the participants' sociodemographic profile on personal and household level, their mobility tools and behavior as well as a ridepooling assessment with respect to MOIA. The sociodemographic questions were standardized in accordance with the German transport census 2017 (MiD - Mobilität in Deutschland; Nobis & Kuhnimhof (2019)) to allow for re-weighting to population level after model estimation. For the purpose of this study, reference values for home-based work and leisure trips were generated via the Google API for each participant. The modes considered in the choice experiment and model were walking (W), cycling (B), car (C), public transport (PT), taxi (T), MOIA (M) and an intermodal alternative (MPT) consisting of MOIA and PT, where MOIA was assumed to be feeder for PT. A typical ride with MOIA can be characterized as a sequence of three stages: Walking to the pick-up location, traveling in a MOIA vehicle and walking from the drop-off location to the final destination. To simulate this process, MOIA's internal virtual stop network with over 12,000 stops was utilized to effectively route these trips. The RP data contained MOIA as an option only if all of the specified locations were located within the MOIA service area as of 2023. In addition, level-of-service attributes (e.g. an arrival window due to ridepooling or wait time for the vehicle to arrive) were included in the choice experiment.

The overall mode choice experiment entailed six experimental designs which include seven alternatives displayed to a respondent. Dependent on the trip distance, purpose, driver's license ownership and car availability, each respondent was assigned to a block of eight choice situations of one of those designs. In the end, six D-efficient pivot designs were implemented using NGene (Rose & Bliemer, 2009; ChoiceMetrics, 2021).

Modeling approach

The data gathered allowed for the estimation of pooled RP-SP MNL choice models. The utility functions are depicted in Equations 1 to 3. For ease of readability, the subscripts for choice task t are omitted.

$$V_{i,n,s} = \alpha_{i,s} + \delta_{shift,i,s} \cdot z_{shift,n,s} + \beta_{LOS,i} \cdot x_{LOS,n} + \gamma_{socio,i} \cdot x_{socio,n} + f_c(x_{dist,n,s}, z_{leis,n,s}) \cdot x_{cost,i,n,s} + f_{tt,i}(x_{dist,n,s}, z_{leis,n,s}) \cdot x_{tt,i,n,s} \quad (1)$$

$$f_c(x_{dist,n,s}, z_{leis,n,s}) = (\beta_{cost,com} + \beta_{cost,leis} \cdot z_{leis,n,s}) \cdot \left(\frac{x_{dist,n,s}}{8\text{km}} \right)^{\lambda_{cost,dist}} \quad (2)$$

$$f_{tt,i}(x_{dist,n,s}, z_{leis,n,s}) = (\beta_{tt,com,i} + \beta_{tt,leis,i} \cdot z_{leis,n,s}) \cdot \left(\frac{x_{dist,n,s}}{8\text{km}} \right)^{\lambda_{tt,dist,i}}, \quad (3)$$

where alternative $i \in J = \{W, B, C, PT, T, M, MPT\}$ and data source $s \in \{RP, SP\}$. $\alpha_{i,s}$ represents the alternative-specific constants (ASC) and $\delta_{shift,i,s}$ denotes shifts on the ASCs (trip purpose and MOIA membership) for both the RP & SP model component. $\beta_{LOS,i}$ denotes the influence of alternative-specific LOS attributes like access/egress time ($\beta_{i,aet}$) and arrival time window for M ($\beta_{M,latewin}$). Wait time while transferring ($\beta_{waittime}$), wait time for the vehicle ($\beta_{waitveh}$) and number of transfers (β_{trans}) were estimated jointly where applicable for alternatives PT, T, M and MPT. $\gamma_{socio,i}$ captures alternative-specific sociodemographic characteristics for age, gender and household income. The travel time and cost coefficient ($\beta_{tt,com,i}$, $\beta_{cost,com}$) is modeled as a non-linear function of beeline distance, which also includes a shift for leisure trips ($\beta_{tt,leis,i}$, $\beta_{cost,leis}$) to account differences in sensitivities across these trip purposes.

Since an intermodal alternative was introduced in the choice experiment, where MOIA is assumed to be an access mode to PT (even though that this might already be a use case for MOIA in reality, the RP survey did not capture it), two model specifications were tested. While MNL 1 treated the in-vehicle travel times for MOIA as separate effects (main mode: $\beta_{M,tt,com}$ and access mode: $\beta_{MPT,M,tt}$), MNL 2 estimated these two effects jointly in $\beta_{M,tt,com}$ (see Table 1). As a consequence, this affects the VTT for MOIA, which is discussed in the next section. The proposed modeling approach accounts for the impact of trip purpose and distance, various LOS and sociodemographic attributes on mode choice in the presence of ridepooling. The models were estimated in preference space, using R and the Apollo package (R Core Team, 2020; Hess & Palma, 2019).

3 RESULTS AND DISCUSSION

An analysis of the sample in comparison with the German MiD, which was restricted to Hamburg, uncovered two noteworthy observations: Firstly, the sample was representative in terms of gender and holders of a driver's license. Secondly, it demonstrated a slight inclination towards younger participants, as well as an over-representation of individuals with a high degree of education (university or diploma) and high-income households (more than 6,500 Euro per month). Consequently, post-estimation reweighting was necessary for the measures of interest using sample enumeration.

In the sample, 90% of participants (3,715) responded affirmatively to having booked a ride with MOIA within the past year, which might be an artefact of the recruitment process, but also demonstrates the popularity of MOIA's ridepooling service in Hamburg. The sample's sociodemographic profile in general is very similar compared to the work done in 2019 by Kagerbauer et al. (2021). MOIA users are on average 44 years old, well educated and mostly live in one-person, high-income households. With regard to their last trip booked, two insights were relevant for the choice experiment. It was found that 60% of the participants reported their last trip as being for leisure, 12.9% for travel to a train station or the airport, and 10.6% for work. Additionally, 14% of the trips were combined with other modes of transportation, with 70% of those including PT. This indicated the necessity of incorporating and testing an intermodal alternative in the DCE, also because the survey did not differentiate between a combination with short-distance and long-distance PT modes.

Model results

Figure–1 shows the choice frequencies (market shares) divided by data source. Examining the RP data, car and PT were the most commonly chosen alternatives (33.8% and 33% respectively), followed by bike and walk (18.2% & 13.4%). MOIA exhibits a market share of 1.6%, which is higher than MOIA’s actual share of roughly 0.1%. This might be related to the large proportion of MOIA users in the sample. The SP data illustrates a decrease in car, PT and walk market shares, in favor of the intermodal alternative and MOIA. Taxi was the least chosen mode with 0.1% and 0.7% market shares in the RP and SP data, respectively. This is unsurprising as taxis are not often the first choice for commutes or frequent leisure trips. It is important to note that the SP modal splits presented here are an artefact of the experimental design and the associated distance classes, and should not be compared to real-world modal splits (Glerum et al., 2013). A closer investigation of non-trading and lexicographic choice behavior revealed that in 42% of all choice tasks the least expensive option was chosen, which either indicates a highly price-sensitive sample or an experimental design that transparently exposed all fares. However, this can be attributed to the assumption of zero costs for PT season ticket subscribers. On an individual level, 18% of all participants always (i.e. 8 times) chose the least expensive mode. As a result, in terms of cost, both PT and the intermodal alternative were viable options, which could explain the observed popularity of the latter.

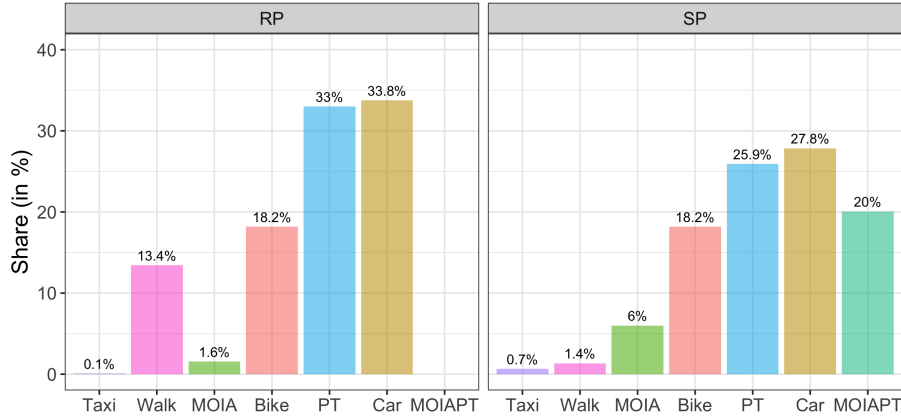


Figure 1: Choice frequencies

The model estimates are presented in Table 1. The units of the temporal variables are minutes, while those of the costs are Euros. Building upon the recommendations of Wasserstein et al. (2019), the table does not present associated p-values. Thus, researchers should recognize the presence of uncertainty and should not assume that effects exist simply due to the statistical significance or lack thereof. The model fit, $Adj.p^2$, of both models is almost equal, although slightly higher for MNL 1, which likely is a consequence of not pooling the travel time attribute for MOIA of the intermodal alternative ($\beta_{MPT,tt,com}$). However, the VTT differ between the two models, as shown in Figures 2 & 3 and discussed after examining the estimates.

The most substantial shifts on the ASC’s were given by $\delta_{M,s,nonuser}$ and thus controlling for MOIA membership, demonstrating that non-users were less prone to choose MOIA and MOIAPT. Furthermore, it was important to also account for trip purpose related shifts on the ASC as they influence the subsequent calculation of the VTT. Noteworthy findings concerning the sociodemographical variables considered are the following: With respect to gender, a substantial effect was only observed for the choice of bicycles. Men were more likely to do so than women. An effect for high income households was only apparent for private cars. As such, this does not seem to be the case for choosing a ridepooling service like MOIA. The strongest impact among all variables was observed for education. While a high education degree has a positive impact on the probability of choosing a bicycle, it has negative one for all other modes. This finding might be related to the fact that cycling is a more sustainable way of traveling compared to motorized modes and thus appeals more to well educated people.

Table 1: Estimation results

Reference: Walking Parameter	MNL 1		MNL 2	
	Estimate	Rob. t-ratio	Estimate	Rob. t-ratio
$\alpha_{W,RP}$	0.000	(NA)	0.000	(NA)
$\alpha_{B,RP}$	-1.712	(-2.038)	-1.727	(-2.053)
$\alpha_{C,RP}$	-2.442	(-2.961)	-2.538	(-3.069)
$\alpha_{PT,RP}$	-1.455	(-1.765)	-1.509	(-1.831)
$\alpha_{T,RP}$	-5.499	(-7.951)	-5.616	(-8.203)
$\alpha_{M,RP}$	-3.253	(-3.557)	-3.237	(-3.526)
$\delta_{B,RP,leis}$	-0.904	(-1.609)	-0.922	(-1.644)
$\delta_{C,RP,leis}$	-0.121	(-0.212)	-0.079	(-0.139)
$\delta_{PT,RP,leis}$	-0.597	(-1.053)	-0.552	(-0.976)
$\delta_{M,RP,leis}$	-0.044	(-0.070)	-0.579	(-0.931)
$\delta_{M,RP,nonuser}$	-1.495	(-1.456)	-1.487	(-1.443)
$\alpha_{W,SP}$	0.000	(NA)	0.000	(NA)
$\alpha_{B,SP}$	-1.268	(-1.476)	-1.266	(-1.469)
$\alpha_{C,SP}$	-2.729	(-3.236)	-2.814	(-3.320)
$\alpha_{PT,SP}$	-1.796	(-2.134)	-1.845	(-2.189)
$\alpha_{T,SP}$	-3.154	(-5.944)	-3.250	(-6.210)
$\alpha_{M,SP}$	-0.846	(-0.921)	-0.832	(-0.902)
$\alpha_{MPT,SP}$	-3.275	(-3.760)	-2.672	(-3.080)
$\delta_{B,SP,leis}$	0.296	(0.491)	0.267	(0.443)
$\delta_{C,SP,leis}$	0.364	(0.601)	0.437	(0.725)
$\delta_{PT,SP,leis}$	0.259	(0.434)	0.314	(0.524)
$\delta_{M,SP,leis}$	0.182	(0.286)	-0.329	(-0.526)
$\delta_{M,SP,nonuser}$	-1.571	(-4.100)	-1.592	(-4.165)
$\delta_{MPT,SP,leis}$	1.119	(1.741)	0.737	(1.150)
$\delta_{MPT,SP,nonuser}$	-0.781	(-5.017)	-0.777	(-4.915)
$\gamma_{B,male}$	0.157	(1.012)	0.157	(1.006)
$\gamma_{B,age31-65}$	0.072	(0.360)	0.075	(0.375)
$\gamma_{B,age66-86}$	-0.412	(-1.082)	-0.422	(-1.098)
$\gamma_{B,educhigh}$	0.189	(1.178)	0.195	(1.207)
$\gamma_{B,inc1.7-5.5k}$	0.059	(0.134)	0.060	(0.135)
$\gamma_{B,inc6.5k+}$	0.163	(0.350)	0.163	(0.346)
$\gamma_{C,male}$	-0.037	(-0.254)	-0.037	(-0.252)
$\gamma_{C,age31-65}$	0.232	(1.243)	0.236	(1.260)
$\gamma_{C,age66-86}$	0.187	(0.522)	0.179	(0.499)
$\gamma_{C,educhigh}$	-0.514	(-3.438)	-0.523	(-3.487)
$\gamma_{C,inc1.7-5.5k}$	0.566	(1.344)	0.561	(1.321)
$\gamma_{C,inc6.5k+}$	0.650	(1.455)	0.645	(1.434)
$\gamma_{PT,male}$	-0.047	(-0.319)	-0.051	(-0.341)
$\gamma_{PT,age31-65}$	-0.026	(-0.139)	-0.027	(-0.145)
$\gamma_{PT,age66-86}$	-0.191	(-0.536)	-0.213	(-0.598)
$\gamma_{PT,educhigh}$	-0.288	(-1.910)	-0.294	(-1.942)
$\gamma_{PT,inc1.7-5.5k}$	-0.112	(-0.272)	-0.122	(-0.295)
$\gamma_{PT,inc6.5k+}$	-0.430	(-0.978)	-0.443	(-1.002)
$\gamma_{M,male}$	-0.044	(-0.251)	-0.048	(-0.270)
$\gamma_{M,age31-65}$	0.139	(0.655)	0.149	(0.692)
$\gamma_{M,age66-86}$	0.917	(1.431)	0.964	(1.469)
$\gamma_{M,educhigh}$	-0.638	(-3.627)	-0.639	(-3.588)
$\gamma_{M,inc1.7-5.5k}$	0.047	(0.084)	0.053	(0.093)
$\gamma_{M,inc6.5k+}$	0.007	(0.012)	0.003	(0.005)
$\gamma_{MPT,male}$	-0.153	(-0.984)	-0.157	(-0.998)
$\gamma_{MPT,age31-65}$	0.290	(1.458)	0.295	(1.473)
$\gamma_{MPT,age66-86}$	0.490	(1.303)	0.490	(1.297)
$\gamma_{MPT,educhigh}$	-0.497	(-3.127)	-0.508	(-3.179)
$\gamma_{MPT,inc1.7-5.5k}$	0.105	(0.256)	0.093	(0.226)
$\gamma_{MPT,inc6.5k+}$	0.029	(0.066)	0.015	(0.035)
$\beta_{cost,com}$	-0.292	(-14.970)	-0.301	(-15.438)
$\beta_{cost,leis}$	0.127	(7.165)	0.139	(7.809)
$\beta_{W,tt,com}$	-0.130	(-7.736)	-0.133	(-7.842)
$\beta_{B,tt,com}$	-0.144	(-12.704)	-0.148	(-13.089)
$\beta_{C,tt,com}$	-0.037	(-3.537)	-0.035	(-3.327)
$\beta_{C,aet}$	-0.086	(-12.696)	-0.087	(-12.828)

Reference: Walking Parameter	MNL 1		MNL 2	
	Estimate	Rob. t-ratio	Estimate	Rob. t-ratio
$\beta_{PT,tt,com}$	-0.061	(-10.903)	-0.062	(-11.107)
$\beta_{PT,aet}$	-0.052	(-9.911)	-0.053	(-10.085)
$\beta_{T,tt,com}$	-0.015	(-0.686)	-0.012	(-0.571)
$\beta_{M,tt,com}$	-0.087	(-8.255)	-0.092	(-8.893)
$\beta_{M,aet}$	-0.022	(-2.309)	-0.018	(-1.819)
$\beta_{M,latewin}$	-0.039	(-7.818)	-0.039	(-7.810)
$\beta_{MPT,M,tt}$	-0.018	(-2.823)	(NA)	(NA)
$\beta_{waittime}$	-0.020	(-2.532)	-0.019	(-2.362)
β_{trans}	-0.269	(-12.445)	-0.268	(-12.350)
$\beta_{waitveh}$	-0.021	(-4.934)	-0.022	(-5.000)
$\beta_{W,tt,leis}$	0.009	(0.476)	0.010	(0.567)
$\beta_{B,tt,leis}$	0.016	(1.382)	0.022	(1.875)
$\beta_{C,tt,leis}$	-0.014	(-1.420)	-0.014	(-1.384)
$\beta_{PT,tt,leis}$	0.019	(2.608)	0.022	(3.116)
$\beta_{T,tt,leis}$	-0.016	(-0.559)	-0.018	(-0.604)
$\beta_{M,tt,leis}$	0.025	(1.791)	0.057	(5.079)
$\lambda_{cost,dist}$	-0.531	(-14.592)	-0.502	(-14.057)
$\lambda_{W,tt,dist}$	-0.212	(-2.780)	-0.206	(-2.684)
$\lambda_{B,tt,dist}$	-0.276	(-6.043)	-0.270	(-5.891)
$\lambda_{C,tt,dist}$	-0.254	(-3.311)	-0.216	(-1.987)
$\lambda_{PT,tt,dist}$	-0.077	(-0.777)	-0.033	(-0.325)
$\lambda_{T,tt,dist}$	0.357	(1.146)	0.373	(1.097)
$\lambda_{M,tt,dist}$	0.090	(1.074)	0.017	(0.220)
μ_{RP}	1.000	(NA)	1.000	(NA)
μ_{SP}	1.053	(0.982)	1.034	(0.658)
LL(0,RP)		-8147.186		-8147.186
LL(0,SP)		-34631.630		-34631.630
LL(final,RP)		-4728.957		-4724.122
LL(final,SP)		-23942.468		-23967.600
LL(final,model)		-28671.426		-28691.722
Adj. ρ^2 (model)		0.330		0.329
Number of respondents		3823		3823
Number of observations		28907		28907
Number of parameters		86		85

Figures-2 and 2 present the VTTs for all main modes of both models. As can be inferred from the corresponding legend, the VTT for MOIAPT is not displayed since the main mode is PT, whose parameter was jointly modeled with the PT alternative. Moreover, the VTTs are only modeled for the range of distance the modes were available in the data. The Delta method was used to estimate the VTTs and their 95%-confidence interval (Daly et al., 2012). Despite a very similar pattern of VTT for commute trips, contrastingly, the two models yield notable differences in the VTT for leisure trips. The models showed that, in general, travel cost tends to have a more substantial influence on mode choice than travel time. This also holds for the corresponding shift of the parameter for leisure trips and distance elasticity. A non-linear interaction of cost with household income was also tested, but turned out to have no effect. However, the difference in VTTs among the modes is primarily the result of the different travel time parameters (and their respective shifts), as the cost coefficient is generic (i.e. the same for all modes). Interestingly, for bicycle, public transport, and MOIA, leisure travel time sensitivities are less negative compared to their commutes, suggesting that participants find the time spent using these modes more enjoyable. This manifests even more in MNL 2, where the VTT for MOIA is much lower due to the joint estimation of MOIA as a main and access mode. In addition, the distance elasticity of travel time for MOIA is positive (but not significant), indicating that travel time sensitivity is barely influenced by distance. As such, for MNL 1, this suggests that the VTT for MOIA is comparatively more driven by travel cost rather than time in relation to the other modes, and hence for shorter distances more similar to car and PT. For MNL 2, the leisure effect induced by pooling the travel time parameters seemed to affect the VTT more strongly than cost, resulting in a substantially lower value. This might also be a result of the rather simply constructed intermodal alternative, and needs further investigation in the future. Another interesting outcome from the models regarding MOIA is that a larger

window of arrival time (e.g. late arrival due to pooling and hence possible rerouting) is perceived as almost twice as worse than waiting for the vehicle or accessing/egressing it.

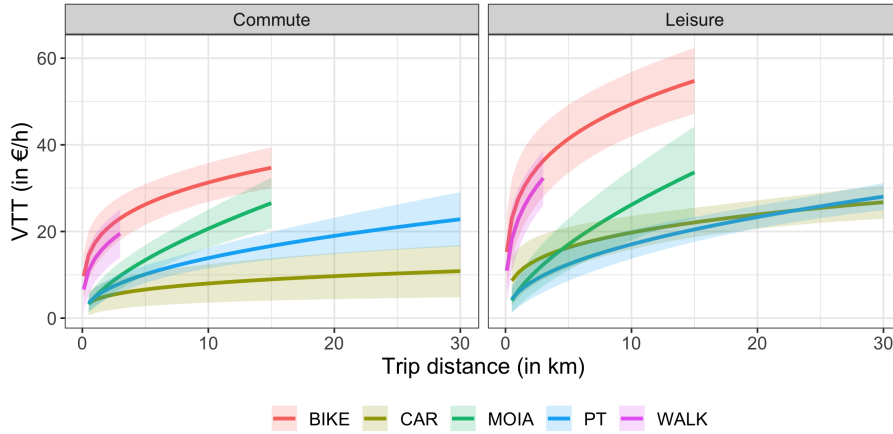


Figure 2: Values of travel time by Beeline-distance (MNL 1)

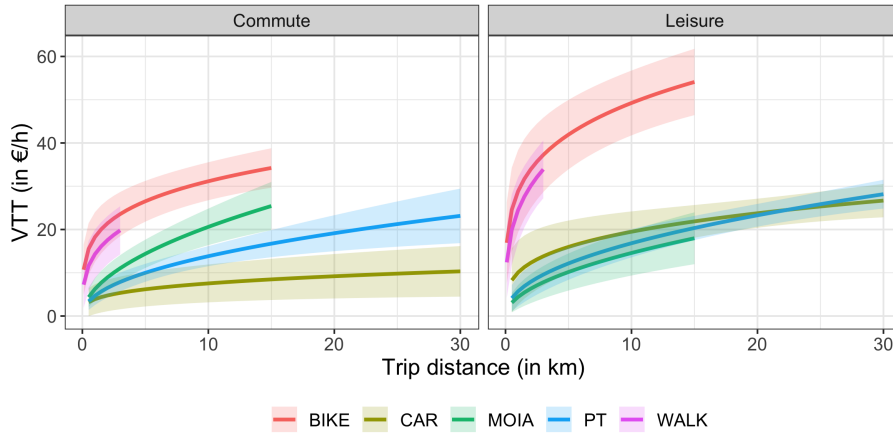


Figure 3: Values of travel time by Beeline-distance (MNL 2)

4 CONCLUSIONS

This study presents a comprehensive investigation of mode choice preferences in the presence of ridepooling for the city of Hamburg. A combination of RP and SP surveys was employed to examine the effect of sociodemographic characteristics and level-of-service attributes on the choice between walking, cycling, cars, PT, taxi, MOIA and an intermodal alternative for commute and leisure trips. Insights from the RP survey were used and tested in the subsequent mode choice experiment and model. Key drivers in choosing a ridepooling service like MOIA and its associated VTT are primarily level-of-service attributes such travel cost and time as well as trip purpose. There are notable differences in the VTT for leisure, however, if the travel time parameter is separately estimated for MOIA as a main and access mode, or jointly. Even if the hypothetically introduced intermodal alternative provides novel insights, the rather simple construction of it can be considered as a limitation of the study and needs further research. In addition, more complex models such as for example Mixed Multinomial or Nested Logit models could be estimated to examine unobserved taste heterogeneity and nesting structures between the modes.

ACKNOWLEDGEMENTS

This research is part of the publicly funded project *Auf dem Weg zum Hamburg-Takt*, Förderkennzeichen ÖPNV-2021-001/203#004 – 2021.

REFERENCES

- ChoiceMetrics. (2021). Ngene 1.3 user manual and reference guide: The cutting edge in experimental design [Computer software manual]. Retrieved from <http://www.choice-metrics.com/downloads>
- Daly, A. J., Hess, S., & de Jong, G. (2012). Calculating errors for measures derived from choice modelling estimates. *Transportation Research Part B*, *46*(2), 333–341.
- Diebold, T., Czarnetzki, F., & Gertz, C. (2021). On-Demand-Angebote als Bestandteil des ÖPNV. Nutzungsmuster und Auswirkungen auf die Verkehrsmittelentscheidung in einem Hamburger Stadtrandgebiet. *Internationales Verkehrswesen*, *73*(3), 88–94. Retrieved from <https://trid.trb.org/view/1884855>
- Glerum, A., Stankovikj, L., Thémans, M., & Bierlaire, M. (2013). Forecasting the demand for electric vehicles: Accounting for attitudes and perceptions. *Transportation Science*, *48*(4), 483–499.
- Hess, S., & Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, *32*, 100170.
- Kagerbauer, M., Kostorz, N., Wilkes, G., Dandl, F., Engelhardt, R., Glöckl, U., ... Zwick, F. (2021). *Ridepooling in der Modellierung des Gesamtverkehrs - Methodenbericht zur MOIA Begleitforschung* (Technical Report). Karlsruhe: Karlsruher Institut für Technologie (KIT).
- Kostorz, N., Fraedrich, E., & Kagerbauer, M. (2021, jan). Usage and User Characteristics—Insights from MOIA, Europe’s Largest Ridepooling Service. *Sustainability*, *13*(2), 958. Retrieved from <https://www.mdpi.com/2071-1050/13/2/958> doi: 10.3390/su13020958
- Louviere, J. J., Hensher, D. A., & Swait, J. D. (2003). *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge.
- Nobis, C., & Kuhnimhof, T. (2019). *Mobilität in Deutschland - MiD Ergebnisbericht* (Technical Report). Bonn, Berlin: Studie von infas, DLR, IVT und infas 360 im Auftrag des Bundesministers für Verkehr und digitale Infrastruktur.
- Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling Transport*. John Wiley & Sons, Chichester.
- R Core Team. (2020). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Rose, J. M., & Bliemer, M. C. J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, *29*(5), 587–617.
- Shaheen, S., & Cohen, A. (2018). Shared ride services in North America: definitions, impacts, and the future of pooling. *Transport Reviews*, *0*(0), 1–16. Retrieved from <https://doi.org/10.1080/01441647.2018.1497728> doi: 10.1080/01441647.2018.1497728
- Train, K. (2009). *Discrete choice methods with simulation*. Cambridge University Press, Cambridge.
- Wasserstein, R. J., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond "p<0.05". *The American Statistician*, *73*(1), 1–19.
- Zwick, F., Kuehnel, N., Moeckel, R., & Axhausen, K. W. (2021, jan). Agent-based simulation of city-wide autonomous ride-pooling and the impact on traffic noise. *Transportation Research Part D: Transport and Environment*, *90*, 102673. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S1361920920308580> doi: 10.1016/j.trd.2020.102673

Investigating preferences for powertrains when buying a car in Germany

Ilka Dubernet*¹ and Dennis Seibert²

¹Ph.D., Research Associate, Institute of Transport Research, German Aerospace Center (DLR), Germany

²M.Sc. Economics, Research Associate, Institute of Transport Research, German Aerospace Center (DLR), Germany

SHORT SUMMARY

Germany’s Federal Climate Change Act requires the transport sector to reduce its greenhouse gas emissions by almost half until 2030. The government has set an ambitious goal of increasing the number of fully electric vehicles to 15 million within the same period. Data on the factors influencing the purchase of low- or zero-emission vehicles is still scarce. Based on stated preference data collected in spring 2022, we found that battery electric vehicles are the preferred choice of respondents in the hypothetical situation of buying a new car, but not in the used car market. Our results support the high demand for battery electric vehicles, suggesting that stagnating shares in new registrations could be due to supply shortages. Purchase price and energy costs are the most important factors leading to the choice of an electric powertrain, indicating that purchase premiums currently put into place in Germany are highly effective.

Keywords: transformation of transport, electrification and decarbonization of transport, powertrain purchase decision, low- or zero-emission vehicles, market share, discrete choice modelling, stated preference survey

1 INTRODUCTION

Despite long-lasting political negotiations on global level for reducing greenhouse gas emissions (GHG emissions) the transportation sector in most countries is still not showing a downward trend (Lamb et al., 2021; Intergovernmental Panel on Climate Change, 2022). This is especially true for road transport, which is still highly dependent on fossil oil. Hence, cutting down greenhouse gas emissions in transportation is still a challenging task (Creutzig et al., 2015) even though policy instruments, such as energy or carbon tax, exist (e.g., Kok et al. (2011); Stepp et al. (2009); Haasz et al. (2018); Whitehead et al. (2021); Dahl & Sterner (1991)).

Until 2030, Germany’s Federal Climate Change Act (KSG) (Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, 2021) requires the transport sector to reduce its GHG emissions by around 48% compared to 1990. The sector’s emissions must then not exceed 85 Mt CO_2 equivalents. A halving of those emissions in less than ten years represents a challenge based on the stagnant development of GHG emissions over the past 30 years as described above. Additionally, the sector missed its KSG sector target in 2021, which puts pressure on the government to set up an immediate climate action program.

To achieve the climate protection targets for transport, the German government is taking measures aimed in particular at a transition of powertrains, i.e., increasing the number of vehicle kilometres travelled electrically via the gradual electrification of the passenger car fleet. These measures include, for example, purchase premiums, the suspension of the motor vehicle tax or the lower taxation of company cars for zero- or low-emission vehicles. The German government has set itself the goal of increasing the number of fully electric passenger cars to 15 million by 2030 (Sozialdemokratischen Partei Deutschlands, BÜNDNIS 90 / DIE GRÜNEN, Freie Demokraten, 2021).

Vehicle purchase decisions by households determine not only the short-term mode choice decision, but also the size and composition of the national vehicle fleet for longer periods. Modelling this behaviour accurately is therefore a key factor to secure sustainable transport planning and policy setting (Kickhöfer et al., 2019).

Car-specific variables such as fuel efficiency (e.g. Alberini et al. (2022)) and a suite of sociodemographic variables, such as age, gender, presence of young children or employment status have been

identified for significantly influencing travel behaviour (e.g. Bhat et al. (2009); Frondel & Vance (2018); Alberini et al. (2022)).

At this point, it is not clear under which circumstances companies and households choose low- or zero-emission vehicles and which factors play a role in their decision. Data on powertrain selections when purchasing a vehicle is still scarce.

To help gather data and gain a more profound understanding of the influencing factors on the decision process when purchasing a car with a specific powertrain, we conducted an online survey among the German population in spring 2022. Based on the reviewed literature, we aim to fill the gaps linking consumer purchase decisions of new and used vehicles with different powertrains. The main objective of the survey is to determine trade-offs and influencing factors on the decision of buying a car with a classic or alternative powertrain technology. To address these objectives and derive parameters for car stock models, we estimated an error component logit mixture model and derived demand elasticities and cross- elasticities respective to changes in the most important continuous variables.

2 METHODOLOGY

Survey

To examine car powertrain preferences, we conducted a survey in which respondents were asked to make trade-offs between different car powertrain alternatives in hypothetical car purchase situations. The survey consists of five blocks of questions: screening questions, car-specific questions, a stated preference part, attitudinal questions, and socio-demographic questions.

In eight labelled discrete choice games, respondents had to choose between the four powertrain options, petrol, diesel, full plug-in hybrid (PHEV) and battery electric power (BEV) for their car purchase. In addition, control attributes that can have an effect on powertrain choice were added, such as fuel price and specific attributes for alternative powertrains such as charging time, range driving with full electrical power and the availability of charging stations.

Figure 1 shows an example of a choice situation presented to the respondents.

Which vehicle would you most likely choose?
(Situation 6 of 8)

	Petrol	Diesel	Plug-in Hybrid	Electric
Power train				
Purchase price	17.000 €	17.000 €	34.100 €	28.400 €
Fuel price/energy cost for 100km	12 €	10 €	5 €	6 €
Charging time at Fast Charging Station				2 hours and longer
Range purely electric			50 km	200 km
Availability public charging station			High	Low
	Your Choice	Your Choice	Your Choice	Your Choice

Figure 1: Example of a Choice Situation Presented to a Respondent (Purchase of a New Car)

The survey contained a maximum of 62 questions. Recruited respondents were participants of a German automotive panel by an external panel provider. After a soft launch with 209 respondents, data was collected in March 2022 via a web-link to the survey. After data cleaning, the final sample contains 16,032 choices of 2,004 respondents. The average response time to answer all questions was 14.47 minutes.

Method

The choice of the powertrain technology is modelled using an error component model, which allows analysing the relative importance of the factors listed above by exploiting the Mixed MNL structure to allow for inter-alternative correlation and heteroscedasticity (Train, 2009). The error components are added in a “pseudo panel” setting: for each alternative, a different random variable, normally distributed with variance σ_{EC} , is added to the utility. The value of this variable is considered fixed across the choices of each individual, representing personal a priori preferences.

Different nested logit specifications, e.g. nests for alternative and classic powertrain, were tested as well, but were not found to lead to statistically significant improvements in model fit.

The model only represents the choice of powertrain technology in a new or used car scenario. In particular, the decision to purchase a new vehicle or to replace an old one, as well as the decision to search on the used or on the new vehicle market are taken as fixed, and are the responsibility of previous steps in car fleet models for which the estimated parameters shall be used in the future. Selection of the variables to include in the model was supported through the methodology described by Hillel et al. (2019). In particular, this method showed that the answers to the attitudinal questions were a strong predictor of the powertrain technology choice.

Test of various model specifications suggested that the model works better with most continuous variables transformed by a logarithm, showing a decreasing sensitivity to changes of the variable with increasing values of the variable.

3 SELECTED RESULTS AND DISCUSSION

The model results are presented in Tables 1 and 2. Estimation statistics are as follows: the final Likelihood of our mode is -13301.96, AIC 26681.92 and BIC 26981.53. The adjusted Rho-square is 0.58. Most parameters are significantly different from 0 at the 95% level (two stars) or 90% level (one star).

Age was found to be an important determinant of the choice. Indeed, respondents between 35 and 45 years old are much more prone to choose an electric or hybrid vehicle than the other age classes. Surprisingly, while the availability of charging stations is important to the respondents, they do not seem to react differently to “medium” and “high” availability. Whether it is because the “medium” level is sufficient for most decision makers, or because the respondents found it difficult to differentiate the two levels from their descriptions, unfortunately cannot be determined without additional data.

The availability of a private garage or parking spot does not have a significant effect, even though including this parameter leads to a significant increase in model fit. A difference is visible, including the interaction: depending on whether the decision maker is homeowner (who likely can decide to install a charging station, parameter $\beta_{garage,owner}$) or renter (who needs to accept the decisions of their landlord, parameter $\beta_{garage,renter}$).

The variance parameter for the error components σ_{EC} is much higher than the scale parameter for the logit error terms (fixed to 1). This indicates a strong influence of personal a priori preference in the choice of a powertrain, with little variations between the situations for a person that cannot be explained by the parameters that are part of the model.

In our survey, respondents with a positive environmental attitude (affect and cognition) and environmental behaviour prefer an alternative powertrain over a classical one. Respondents with a high score on environmental affect are less likely to purchase a diesel car. The effect of environmental cognition on alternative powertrain preferences is the strongest.

The main purpose of this model is to study the determinants of demand for electric vehicles. To this end, Table 3 shows the elasticities and cross-elasticities of demand respective to changes in the most important continuous variables, that is, the percentage change in the demand for each of the alternatives, given a 1% change in the continuous variable.

Elasticities are computed by predicting the demand for the exact situations in the sample, for a small change in each of the variables, and comparing to the predicted demands without the change in the variables.

In general, the demand is relatively inelastic, in the sense that a given percentage change in the studied variables always leads to a smaller change in the demand.

BEV Price is the variable that leads to the largest relative changes in BEV demand, followed by BEV energy costs, BEV range and then PHEV price. PHEV range only has a small effect on demand, which is consistent with the selling argument of PHEV that limitations in electric range can be compensated using an internal combustion engine.

While smaller than purchase price, the elasticities relative to energy costs are still in a similar range, meaning that an alternative to policies having an effect on purchase price would be policies that make energy costs for BEVs lower compared to other alternatives, either through subsidies or taxation.

Interestingly, charging time has a very limited effect on demand: decision makers very likely want to be able to perform their full trips in one drive, and charge the vehicle while they are performing their activities, making range a much more important factor than charging time.

Table 1: Model Estimates (dummies)

Parameter	Estimate	Std. Dev.	T-Test (0)	
$\alpha_{petrol,new}$	0			
$\alpha_{diesel,new}$	-3.1	0.37	-8.4	**
$\alpha_{phev,new}$	-1.7	0.5	-3.4	**
$\alpha_{bev,new}$	-1.2	0.45	-2.5	**
$\alpha_{petrol,used}$	0			
$\alpha_{diesel,used}$	-1.7	0.19	-9.1	**
$\alpha_{phev,used}$	-1.3	0.37	-3.6	**
$\alpha_{bev,used}$	-1.2	0.3	-4.1	**
$\beta_{classic,[18,25)}$	0			
$\beta_{classic,[25,35)}$	0			
$\beta_{classic,[35,45)}$	0			
$\beta_{classic,[45,55)}$	0			
$\beta_{classic,55+}$	0			
$\beta_{alternative,[18,25)}$	0			
$\beta_{alternative,[25,35)}$	0.26	0.3	0.85	
$\beta_{alternative,[35,45)}$	0.65	0.31	2.1	**
$\beta_{alternative,[45,55)}$	0.3	0.29	1	
$\beta_{alternative,55+}$	0.3	0.28	1.1	
$\beta_{petrol,children}$	0			
$\beta_{diesel,children}$	1	0.27	3.7	**
$\beta_{phev,children}$	0.35	0.29	1.2	
$\beta_{bev,children}$	0.4	0.28	1.5	
$\beta_{garage,owner}$	0.18	0.2	0.91	
$\beta_{garage,renter}$	-0.2	0.21	-0.98	
$\beta_{av,low}$	0			
$\beta_{av,medium}$	0.51	0.049	10	**
$\beta_{av,high}$	0.52	0.047	11	**

To test the effect of variables taking only discrete values, market shares under alternative “scenarios” were computed. Figure 2 presents the results for the “base” scenario (running predictions on the sample without changes), “high availability” scenario (availability of charging stations for BEV set to “high” in all choice situations) and “all garage” scenario (all respondents are assumed to have access to a private garage or parking spot). Two price variations are also included: -15% for all BEV alternatives, and a simulation of the “Umweltbonus” currently in effect in Germany (-9000 Euros for new BEVs with a price under 40000 Euros, -7500 Euros for new BEVs with a price over 40000 Euros). Predictions are simulated for random values of the parameters sampled from the asymptotic distribution of the parameters, to get a feeling for results variance.

Availability of charging stations does have a noticeable effect on BEV market shares, both in the new and used market, but the difference remains small. Access to a private garage does not have any noticeable effect. In comparison, the effect of the two price policies is much larger.

Overall, the results seem to point to the fact that the technology is mature enough, such that the most important factor leading to a purchase decision for a BEV is the price.

Table 2: Model Estimates (Continuous Variables)

Parameter	Estimate	Std. Dev.	T-Test (0)	
σ_{EC}	2.7	0.056	49	**
λ_{inc}	-0.051	0.042	-1.2	
$\lambda_{inc,fuel}$	0.049	0.062	0.79	
β_{price}	-2	0.055	-36	**
$\beta_{fuel\ price}$	-1.2	0.09	-13	**
$\beta_{fuel\ price,educated}$	-1.5	0.064	-23	**
$\beta_{loadtime}$	-0.12	0.037	-3.2	**
β_{km}	-0.82	0.092	-8.9	**
$\beta_{phev,range}$	0.15	0.073	2.1	**
$\beta_{bev,range}$	0.72	0.071	10	**
$\beta_{petrol,dist}$	0			
$\beta_{diesel,dist}$	0.72	0.12	5.9	**
$\beta_{phev,dist}$	0.26	0.094	2.7	**
$\beta_{bev,dist}$	0.27	0.12	2.4	**
$\beta_{petrol,COG}$	0			
$\beta_{diesel,COG}$	-0.096	0.12	-0.83	
$\beta_{phev,COG}$	0.39	0.11	3.4	**
$\beta_{bev,COG}$	0.64	0.12	5.3	**
$\beta_{petrol,AFF}$	0			
$\beta_{diesel,AFF}$	-0.23	0.11	-2.1	**
$\beta_{phev,AFF}$	0.24	0.11	2.2	**
$\beta_{bev,AFF}$	0.58	0.11	5.4	**
$\beta_{petrol,BEH}$	0			
$\beta_{diesel,BEH}$	0.24	0.13	1.8	*
$\beta_{phev,BEH}$	0.36	0.12	2.9	**
$\beta_{bev,BEH}$	0.4	0.12	3.3	**

4 CONCLUSIONS

This paper presents the findings of a stated preference survey on powertrain preference when purchasing a new or used vehicle. We report on the survey design, experiences made during data collection, preparation for model estimations and the results of an error component logit mixture model. It was shown that the collected data set holds rich information with a promising number of cases suitable for modelling the preferences.

Overall, the results seem to point to the fact that the technology is mature enough, such that the most important factor leading to a purchase decision for a BEV is the price. These results may help policymakers for continued or future measures aimed in particular at the transition of powertrains. In particular, the results show that while the current incentive put into place in Germany (“Umweltbonus”, a reduction in purchase price) is very effective, a modification of the difference in price of energy per kilometre between technologies would also have a potentially important impact on the preference for electric powertrains. Under the hypothesis that the consumers see these changes as a long-term trend that will impact them for the whole lifetime of their vehicle, rather than a potentially short-lived policy.

Our survey also reflects a snapshot of how high the registration numbers of electric vehicles could be if supply shortages implying long delivery times and high purchase prices were overcome quickly.

Table 3: Demand Elasticities

Parameter	Elasticities							
	Petrol		Diesel		PHEV		BEV	
	Used	New	Used	New	Used	New	Used	New
BEV Price	0.2	0.27	0.2	0.29	0.21	0.28	-0.6	-0.47
PHEV Price	0.2	0.24	0.22	0.26	-0.66	-0.6	0.2	0.21
Petrol Price	-0.57	-0.66	0.28	0.24	0.26	0.19	0.25	0.16
Diesel Price	0.18	0.15	-0.7	-0.8	0.18	0.12	0.15	0.1
BEV Charge Time	0.012	0.016	0.012	0.017	0.013	0.017	-0.036	-0.029
BEV Range	-0.071	-0.099	-0.071	-0.1	-0.076	-0.1	0.21	0.17
PHEV Range	-0.015	-0.019	-0.017	-0.021	0.051	0.047	-0.015	-0.017
BEV Energy Costs	0.14	0.2	0.14	0.21	0.15	0.21	-0.42	-0.35
PHEV Energy Costs	0.14	0.18	0.15	0.19	-0.47	-0.44	0.14	0.15
Petrol Energy Costs	-0.4	-0.39	0.2	0.14	0.19	0.11	0.17	0.093
Diesel Energy Costs	0.12	0.11	-0.49	-0.58	0.13	0.092	0.11	0.075

Against the background of halving emissions in less than ten years, it is crucial how quickly battery electric vehicle’s new registrations (and thus indirectly its stock) increase.

Finally, in our survey only the demand is modelled, but not the supply. In the current (pandemic and the Russia-Ukraine war) situation, manufacturers cannot produce as many electric vehicles as they are in demand. This is still foreseeable in the coming years and could therefore be a major problem for the ramp-up in passenger car traffic.

REFERENCES

- Alberini, A., Horvath, M., & Vance, C. (2022). Drive less, drive better, or both? behavioral adjustments to fuel price changes in germany. *Resource and Energy Economics*, 101292.
- Bhat, C. R., Sen, S., & Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B: Methodological*, 43(1), 1–18.
- Creutzig, F., Jochem, P., Edelenbosch, O. Y., Mattauch, L., Vuuren, D. P. v., McCollum, D., & Minx, J. (2015). Transport: A roadblock to climate change mitigation? *Science*, 350(6263), 911–912.
- Dahl, C., & Sterner, T. (1991). Analysing gasoline demand elasticities: a survey. *Energy economics*, 13(3), 203–210.
- Federal Ministry for the Environment, Nature Conservation and Nuclear Safety. (2021). *Federal climate change act*. (http://www.gesetze-im-internet.de/englisch_ksg/englisch_ksg.pdf)
- Fronzel, M., & Vance, C. (2018). Drivers’ response to fuel taxes and efficiency standards: evidence from germany. *Transportation*, 45(3), 989–1001.
- Haasz, T., Vilchez, J. J. G., Kunze, R., Deane, P., Fraboulet, D., Fahl, U., & Mulholland, E. (2018). Perspectives on decarbonizing the transport sector in the eu-28. *Energy strategy reviews*, 20, 124–132.
- Hillel, T., Bierlaire, M., Elshafie, M., & Jin, Y. (2019). Weak teachers: Assisted specification of discrete choice models using ensemble learning. In *8th Symposium of the European Association for Research in Transportation (hEART 2019)*. Retrieved from https://transp-or.epfl.ch/heart/2019/abstracts/hEART_2019_paper_117.pdf

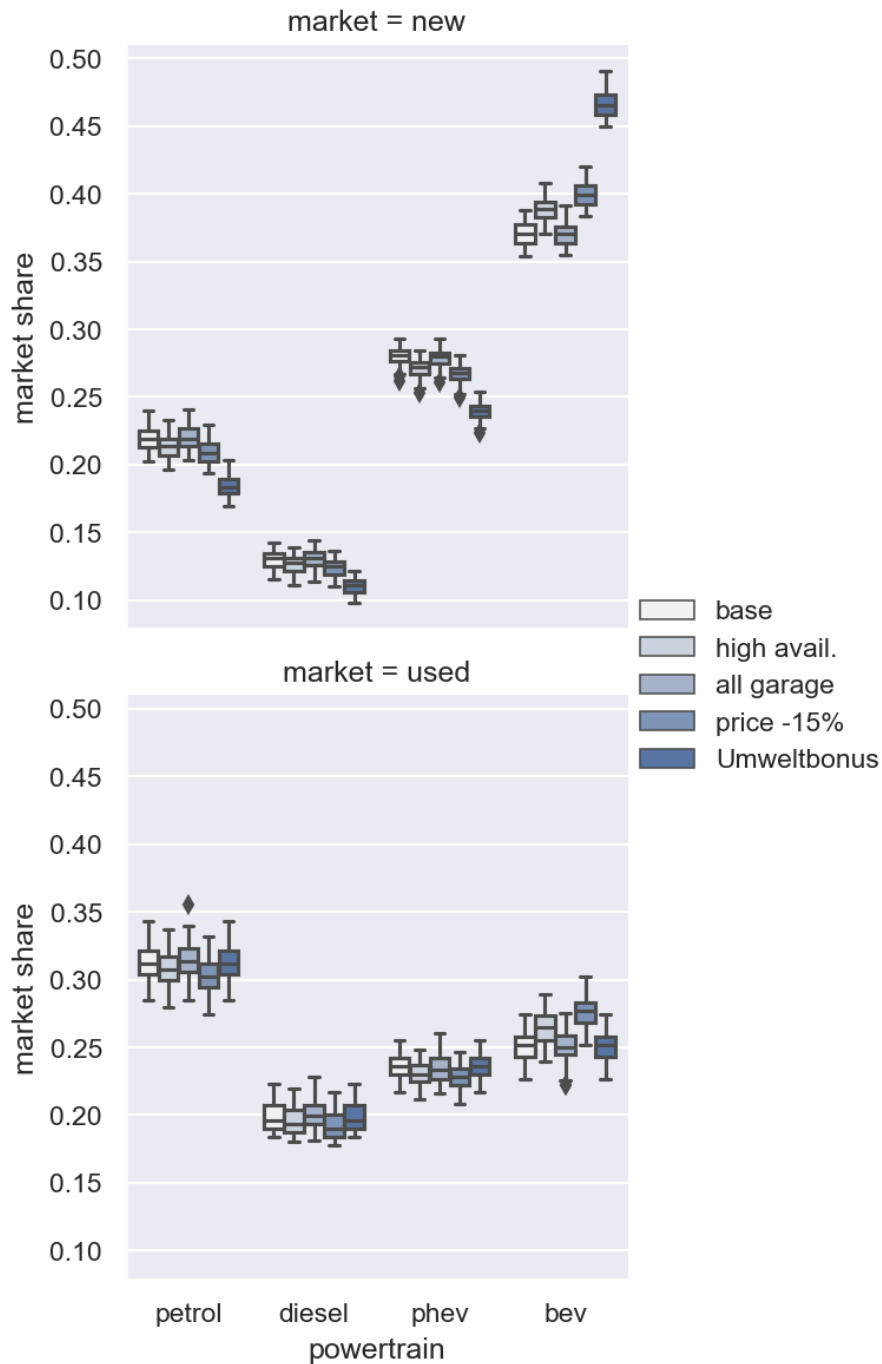


Figure 2: Market Shares Under Different Scenarios

Intergovernmental Panel on Climate Change, W. G. t. (2022). *Climate change 2022, mitigation of climate change*. IPCC, Switzerland.

Kickhöfer, B., Bahamonde-Birke, F. J., & Nordenholz, F. (2019). Dynamic modeling of vehicle purchases and vehicle type choices from national household travel survey data. *Transportation Research Procedia*, 41.

Kok, R., Annema, J. A., & van Wee, B. (2011). Cost-effectiveness of greenhouse gas mitigation in transport: A review of methodological approaches and their impact. *Energy Policy*, 39(12), 7776–7793.

Lamb, W. F., Wiedmann, T., Pongratz, J., Andrew, R., Crippa, M., Olivier, J. G., ... others (2021). A review of trends and drivers of greenhouse gas emissions by sector from 1990 to 2018. *Environmental research letters*.

- Sozialdemokratischen Partei Deutschlands, BÜNDNIS 90 / DIE GRÜNEN, Freie Demokraten. (2021). *Koalitionsvertrag 2021 – 2025 zwischen der Sozialdemokratischen Partei Deutschlands (SPD), BÜNDNIS 90 / DIE GRÜNEN und den Freien Demokraten (FDP)*. (<https://cms.gruene.de/uploads/documents/Koalitionsvertrag-SPD-GRUENE-FDP-2021-2025.pdf>)
- Stepp, M. D., Winebrake, J. J., Hawker, J. S., & Skerlos, S. J. (2009). Greenhouse gas mitigation policies and the transportation sector: The role of feedback effects on policy effectiveness. *Energy Policy*, *37*(7), 2774–2787.
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
- Whitehead, J., Plötz, P., Jochem, P., Sprei, F., & Dütschke, E. (2021). Policy instruments for plug-in electric vehicles: an overview and discussion. *International Encyclopedia of Transportation*, *1*, 496–502.

Modelling Travel Time Anticipation Under Rational Inattention and Endogenous Information Constraints

Dimitrios Pappelis*¹, Emmanouil Chaniotakis², Tim Hillel³, Maria Kamargianni⁴

¹ Research Assistant, BSEER Energy Institute, University College London, UK, WC1H0NN

² Lecturer, BSEER Energy Institute, University College London, UK, WC1H0NN

³ Lecturer, Civil Engineering, University College London, UK, WC1H0NN

⁴ Professor, BSEER Energy Institute, University College London, UK, WC1H0NN

SHORT SUMMARY

Transportation research has been traditionally grounded on the economic theory of Rational Expectations, assuming that individuals are fully informed, optimizing, and self-interested decision makers. However, this assumption fails to sufficiently explain the inertia that characterizes travellers' behaviour in face of uncertainty. In recent years, there has been a rising interest in the theory of Rational Inattention, arguing that individuals choose to make seemingly suboptimal choices due to the cost of acquiring and processing available information. In this paper, we present a continuous quadratic Rational Inattention model of travel time anticipation. We showcase that its properties satisfy behavioural hypotheses derived from data collected through a case study in the city of Turin on within-day travel re-evaluation. We conduct simulation experiments and propose an alternative 2-stage framework for enhancing existing neoclassical travel behaviour models, indicating potential biases and discrepancies in the forecasted market shares, specifically with regards to rare travel time occurrences.

Keywords: Choice modelling, Dynamic travel behaviour, Inertia, Rational inattention, Traffic information

1. INTRODUCTION

Transportation planning and policy making rely on models to predict and explain the behavior of travellers. Traditionally, research on this front has been based on the economic theory of Rational Expectations, assuming that individuals are fully informed, optimizing, and self-interested decision makers. However, this assumption fails to sufficiently explain the resistance to change that characterizes travellers' behaviour in face of uncertainty. In recent years, there has been a rising interest in the Rational Inattention (RI) theory, originally developed by Christopher Sims (2003). The argument is that individuals consciously choose to make seemingly suboptimal choices due to the cost of acquiring and processing available information. In recent years, Matejka and McKay (2015) expanded the theory for discrete choice under imperfect information and cognitive capacity constraints. As such, RI has emerged as a compelling and neat framework for further understanding the behavior of decision makers in complex and dynamic environments.

In the context of transport modelling, Rational Inattention is still relatively unexplored. Fosgerau et al. (2019) and Jiang et al. (2020) defined the problems of route and departure time choice under

RI and provided simulation findings. Fosgerau et al. (2020) established the general equivalence between discrete choice and RI models, providing an alternate point of view in the interpretation of typical RUM models. From an application perspective, Habib (2022) investigated empirical use-cases and focused on estimable specifications of discrete choice RI models.

In this paper, we present a continuous-quadratic RI model of travel time anticipation. We showcase that its properties satisfy our behavioural hypotheses derived from data collected from a case study in the city of Turin on within-day travel demand shift choices. We proceed to assess the model capabilities through numerical experiments and then propose a 2-stage framework for enhancing existing neoclassical models of travel behaviour, given the open challenges associated with data collection for RI phenomena. We indicate how ignorance of the priors and information capacity constraints could lead to potential biases and discrepancies in the forecasted market shares, especially with regards to rare travel time occurrences.

2. METHODOLOGY

Data Collection

The motivation of this paper originates in the investigation of within-day re-evaluation and day-to-day learning as described by Pappelis et al. (2022). In that study, a joint Revealed Preference and Stated Preference (RP-SP) experiment was applied to collect "pseudo" panel data on within-day demand shift choices. The primary objective was to investigate individuals' adaptation strategies when faced with travel time fluctuations on their habitual schedule, and how the accumulated experience affects their future actions. Participants, whose travel patterns were initially recorded using a smartphone tracking application, were provided with travel information for an upcoming habitual trip, either during an activity or en-route to their destination. Given this information, they were asked to record their response in the form of an adaptation strategy. The strategy could involve modifying trip characteristics such as departure time, mode, or route, or changing the target activity through replacement or cancellation. At the end of each day, participants updated their anticipation of travel time for the following day based on accumulated experience and reported whether they would consider long-term adjustments to their habitual schedule.

The described experiment allowed for the exploration of individuals' responses to travel time fluctuations and the implications on their future travel behavior. It was applied in the metropolitan area of Turin (IT) between February and April 2022, as part of a wider travel demand survey. Recruited individuals formed a stratified sample of the travel survey participants, which is representative of the population in the Turin region (a survey company was hired for recruitment). The RP data collection was performed using a smartphone-based travel survey tool, the MobyApp. The habitual activity and travel patterns were tracked from the application in the form of travel diaries over the course of 7 days. In total, 365 individuals accessed the experiment and 351 of them completed it, resulting in 702 tracked trips and 4212 observations.

The dataset revealed some interesting behavioural findings with regards to inertia effects of travel behaviour and the concept of false certainty adoption. For instance, Figure 1 displays the number of trips categorized by re-evaluation strategy, based on the daily fluctuations in travel time. The level of fluctuation is determined by the travel factor parameter, which is multiplied by the habitual travel time for a specific trip of the participant in each scenario. The analysis shows that for medium levels of travel time fluctuation, the dominant re-evaluation strategy is 'No change,' suggesting that many individuals may prefer to stick with their habitual option rather than make changes, even if from a utility maximization perspective this can be seen as "irrational". This

finding aligns with the concept of resistance to change, a heterogeneous factor across the population. As travel time increases, schedule constraints and conflicts may increase stress, leading individuals to consider changing their travel plans (such as adjusting departure time, mode, or route). For extreme levels of travel time fluctuation, we observe the highest likelihood of cancellation or replacement of the activity.

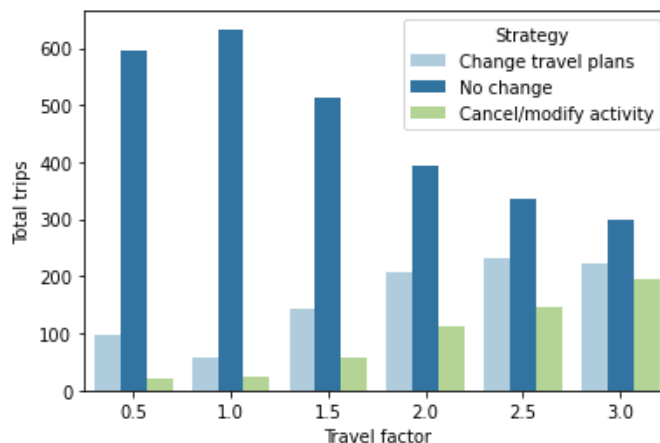


Figure 1 Resistance to change for different levels of travel time fluctuation

It is also important to study how the prior expectation of travel time evolves with accumulated experience. Figure 2 depicts the participants’ scaled anticipated travel time after each day, against the 2-day and 3-day moving average of different travel time orders used throughout the experiment. We observe significant sluggishness and inertia in the travel time anticipation of the participants, being influenced from their prior beliefs and experience. While Rational Expectations theory would imply that external stimuli would cause stronger and fast responses, we observe much milder adaptations and a “magnet effect” towards the reference level of travel time.

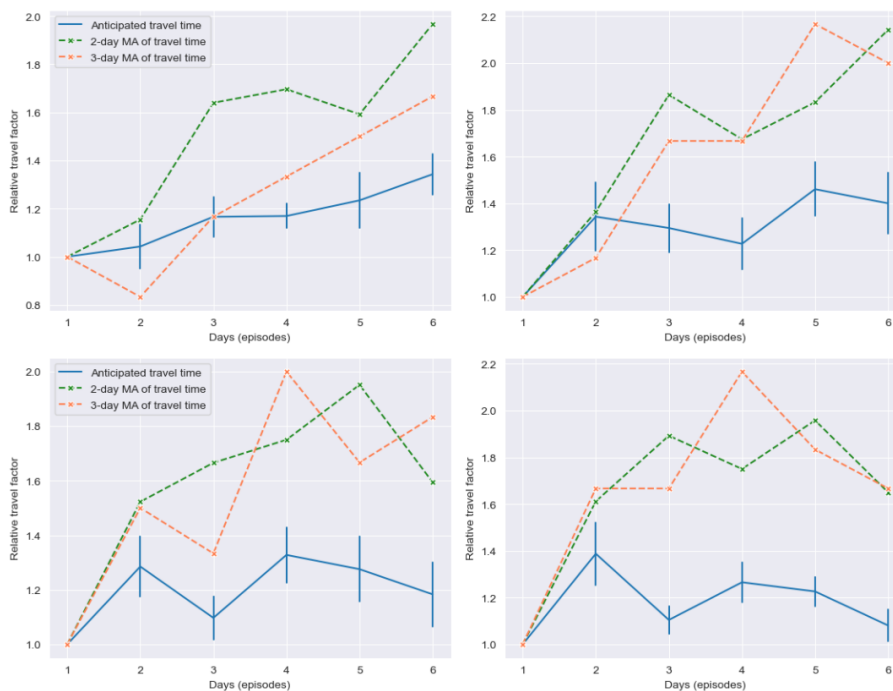


Figure 2 Inertia and sluggishness in travel time anticipation

Modelling Framework

Based on these behavioural observations, we proceed to define the travel anticipation problem as a static model of choice under Rational Inattention (Mackowiak et al., 2021). Consider an agent who plans to perform a daily trip and receives an information signal s , in order to set her travel time anticipation a , subject to unknown network conditions t . Let the utility have the following log-quadratic form,

$$U(a, t) = -b(a - t)^2$$

The agent is tracking the unknown random state of the network, which under perfect information would be equal to her anticipation. Naturally, this would allow the agent to construct her subsequent travel plans most accurately (e.g., departure time, mode, route). However, as the true travel time is infeasible to observe constantly and travel information comes at a perceptual cost, the agent chooses to receive noisy information that determines the posterior beliefs that she may hold. The utility parameter b is a scaler, which can account for agent's heterogeneity with regards to traffic information seeking. Under the general quadratic form, we assume that over or under-estimation of travel time incurs equal losses. In many cases, delayed arrivals might incur costlier losses, so it is worth studying different variants of the utility function going forward. The objective of the agent is to maximize the expectation of her utility less the cost of information $C(f)$, which is a function of the information strategy,

$$\max_f \int U(a, t) f(a, t) dt da - C(f) \quad (1)$$

The joint probability $f(a, t)$ is sufficient to describe the choice of information and action, as they are derived such that no two signals lead to the same action. Otherwise, the agent would be wasting attentional resources by distinguishing between signals that do not directly affect their actions. As a result, it is possible to make a one-to-one association between the signal and action and analyse the relationship between attention, allocation, information acquisition, and decision-making in a unified framework. The objective function (1) is maximized subject to the following constraints,

$$\int f(a, t) da = g(t), \forall t \quad (2)$$

The prior belief of the agent is described by the pdf $g(t)$. Constraint (2) ensures the consistency of the prior and posterior beliefs of the agent under Bayesian rationality.

$$C(f) = \lambda \cdot I(a; t) = \lambda \cdot [H[g(t)] - E[H[t|a]]] \quad (3)$$

The cost function (3) is defined in terms of the mutual information between the agent's anticipation and the actual travel time. It is based on the difference between the entropy of the prior distribution of travel times and the conditional entropy of the distribution of travel times given the agent's prediction. The parameter λ typically referred to as the "attention cost" or "information cost" reflects the required effort of acquiring and processing the information.

$$H[g(t)] = - \int g(t) \log g(t) dt \quad (4)$$

Entropy (4) is quantified using Shannon's definition, which measures the amount of information present in the probability distribution of travel time. The cost function penalizes travel time predictions that require more attention to achieve a specific level of accuracy. By minimizing the difference between the prior and conditional entropy based on the prediction, the cost function encourages accurate predictions that require less attention. The solution to the agent's problem for

an unknown network state t has a probabilistic logit form. The solution of the agent’s problem for an unknown state of the network t is has the following probabilistic logit form.

$$f(a|t) = \frac{p(a)e^{U(a,t)/\lambda}}{\int_z p(z)e^{U(z,t)/\lambda} dz}$$

In most cases, RI problems do require numerical solution methods. A well-studied exception is the case of quadratic utility, Gaussian prior uncertainty, and an unbounded action space, where Gaussian signals are optimal. Interestingly, for a bounded or truncated action space, the solution of the continuous problem is discrete, indicating that the agent contemplates only specific levels for a given choice, a phenomenon commonly observed in the stickiness of product prices. In the context of travel time, this would imply that travellers choose from a finite set of levels when updating their anticipation and might, for instance, set a regular departure time and standard “safety” departure when expecting a range of potential delays.

3. RESULTS AND DISCUSSION

The collection of data for the practical estimation of RI models is challenging, mainly because the concept of cognitive capacity constraints is abstract and difficult to measure. In the context of travel time anticipation and travel behavior, an ideal dataset would need to capture multiple factors simultaneously, including the agent's beliefs (i.e., their prior perception of the probability distribution of travel times), the world (i.e., network conditions such as travel time), attention allocation (i.e., the choice of signal or level of information), and action (i.e., the agent's choice). The design of such sophisticated experiments is an ongoing task in economics research. In absence of this complete dataset, we proceed to perform numerical experiments on the travel time anticipation model and then propose a 2-stage approach to enhance traditional neoclassical models of travel behaviour.

Numerical Experiment

To assess and showcase the properties of the modelling framework, we perform numerical experiments that justify our behavioural hypotheses derived from the data analysis. A triangular prior distribution is assumed for the agents’ belief, a common approach in related studies (Figure 1).

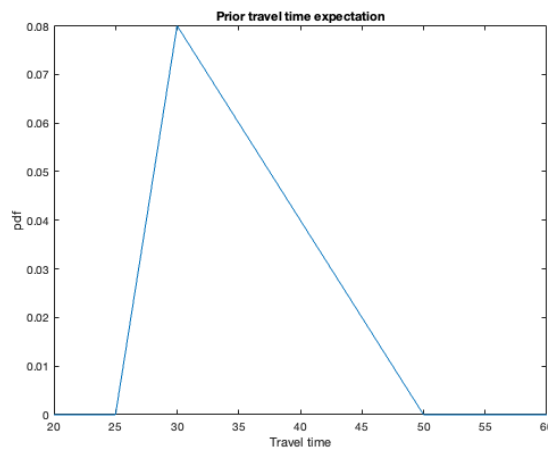


Figure 3 Triangular prior anticipation of trip travel time distribution

We then proceed to solve the RI problem (Eq.1-4) for two different levels of the marginal cost of information λ . The optimization problem was solved using the GAP-SQP geometric algorithm proposed by Armenter et al. (2021). Figure 2 presents the joint probability of anticipated travel times, as well as the conditional probability of the non-zero solutions (discrete choice set). It is apparent that the responsiveness of an action to a given state can be increased by altering the stakes or reducing the cost of information. When the stakes are high or the cost of information is low, individuals are more motivated to make accurate predictions of the travel time and allocate their attention accordingly, thus the plurality in possible actions. This increased attention leads to greater responsiveness of the action to the state, as individuals are more likely to adjust based on the information available to them. On the contrary, for lower stakes or high values of the information constraints, the agent might only consider few alternatives and apply them over a range of states of the network.

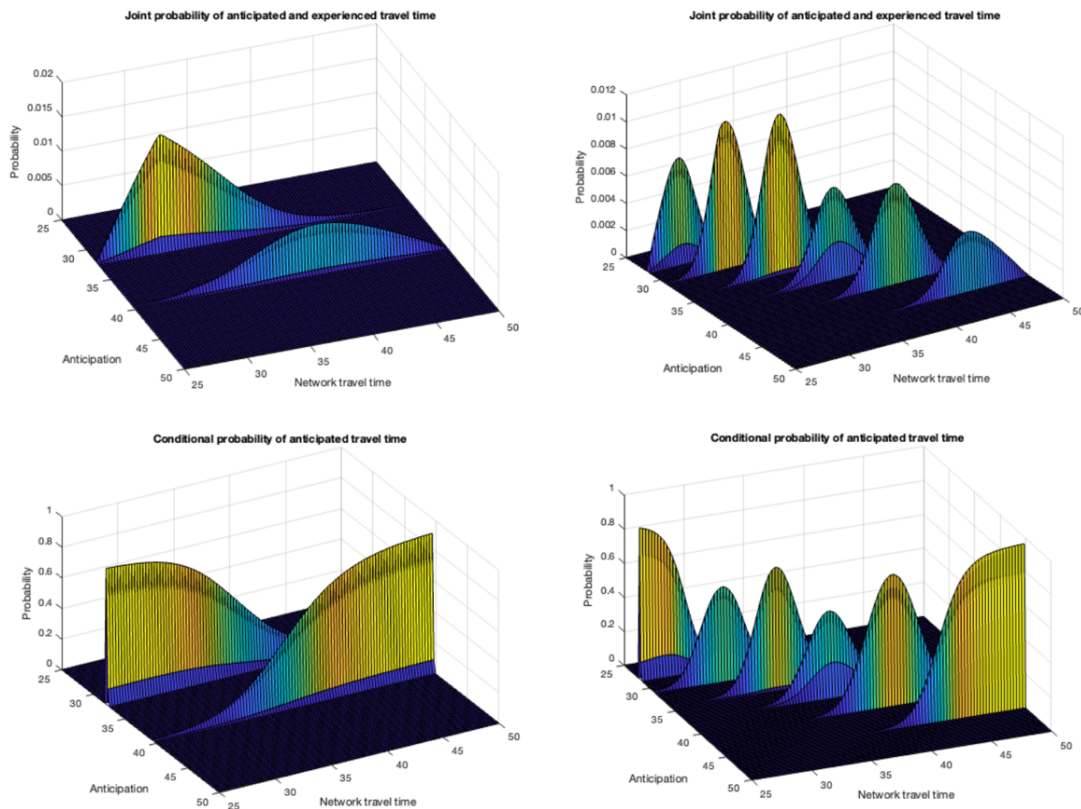


Figure 4 Simulated joint and conditional probabilities for higher ($\lambda=0.03$, left) and lower ($\lambda=0.005$, right) values of marginal information cost

Empirical Findings

The theory of Rational Inattention and the endogenous processing of information raise important questions about what traditional empirical methods, such as controlled experiments, capture in a transportation setting. This is particularly relevant for travel re-evaluation behavior, where it is most often assumed that individuals are fully aware and process all available advanced information. Furthermore, in a revealed preference setting, such effects might already be captured in the data, thus there is a need to not only disentangle preferences, but also consider their equilibrium relationships with the supply side.

Given these open research challenges, we extend the travel re-evaluation framework developed by Pappelis et al. (2022). At this point, it is important to clarify that -in this context of RI- we are not referring to the cognitive constraints of the participant with regards to the experiment setting and attributes, which is also important to be controlled, but with the inattention to information (e.g., journey planners, radio) that would be observed in the transition to a real-world setting. Figure 5 illustrates a two-stage sequential framework for incorporating RI effects in the demand shift models. In the first stage, we utilize the continuous RI model to solve for travel time anticipation. In the second stage, we use the output of the RI model as a more realistic depiction of travel time when simulating dynamic demand shift decisions.

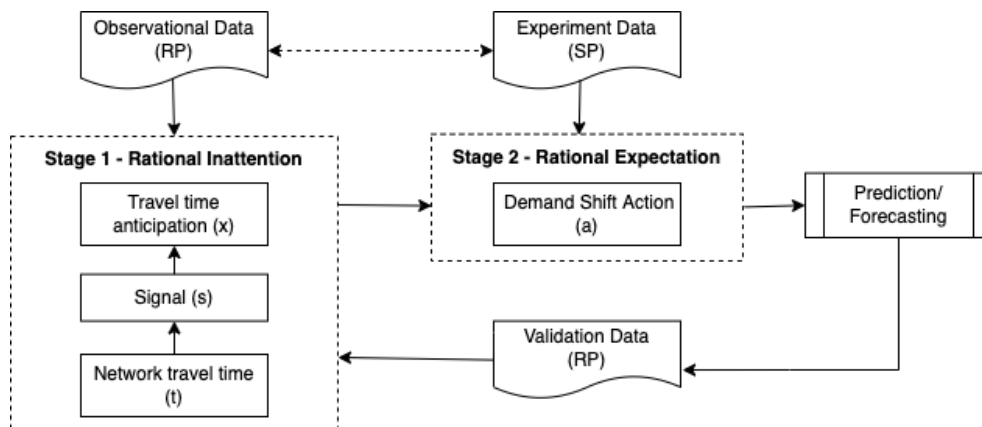


Figure 5 Sequential approach for incorporating travel time anticipation under RI

The model selected for the evaluation of the framework is the static Mixed Nested Logit, which was designed to generate the probability of specific adaptation strategies being selected, when faced with travel time fluctuation during a habitually performed trip. The nesting structure and the alternatives of the travel re-evaluation model are depicted in Figure 6 (see full paper for complete specification).

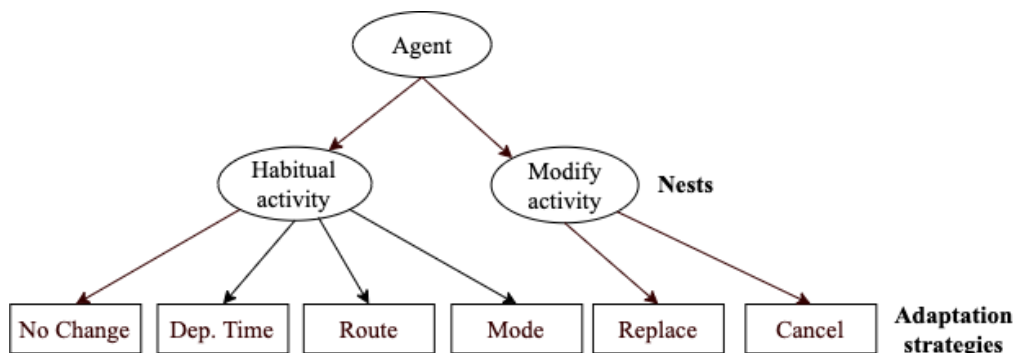


Figure 6 Demand shift model choice alternatives and nesting structure (Pappelis et al., 2022)

Applying the 2-stage framework, we proceed to perform sensitivity analysis on the information cost parameter λ of the travel time variable, maintaining the assumption of the triangular distribution, and then comparing the simulated market shares for different ranges of the travel time distribution. We observe that for severe delays (travel factor >2.5), the Rational Expectations model might overestimate the aggregate response of the travellers, especially when it comes to cancellation of a given trip. Comparing it to the extreme case of a marginal information cost above the threshold of any signal acquisition, a significant discrepancy of over 20% can be observed in

the market share of the “Habit” alternative. On the contrary, for lower levels of travel time fluctuation (travel factor <1.5), the Rational Inattentive agent might falsely overreact due to false signals, when she would be better off following her habitual schedule. Such discrepancies indicate the importance of measuring and accounting for the prior beliefs, the information processing constraints and marginal cost of information λ in travel behaviour modelling and forecasting.

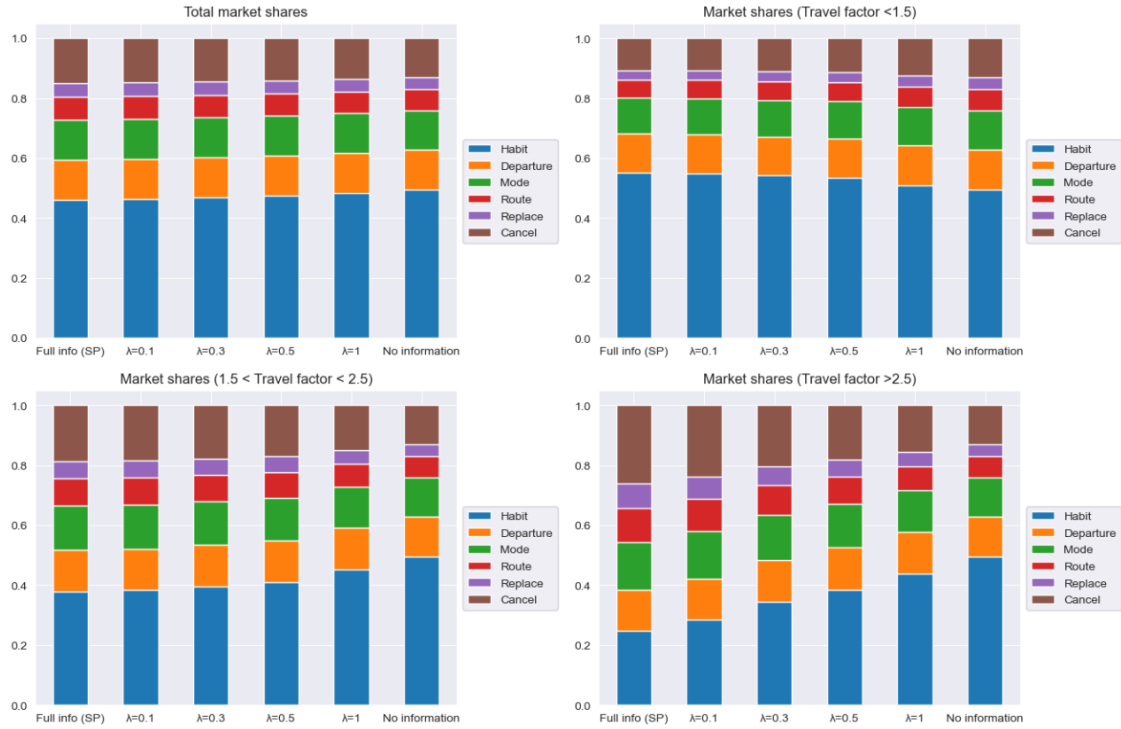


Figure 7 Forecasted shares for different levels of marginal information cost and travel time fluctuation

4. CONCLUSION

In conclusion, our paper highlights the potential benefits of incorporating Rational Inattention theory into transportation modelling and travel time anticipation in particular. Future steps include the extension of the framework to a dynamic setting, allowing for individuals to acquire informative signals which can also be used as predictors of future actions. Finally, the relevance and applicability of the RI theory in transportation needs to be further examined through the design of sophisticated data collection experiments.

5. ACKNOWLEDGEMENTS

The research reported in this paper is supported by European Union’s Horizon 2020 research and innovation programme under Grant Agreement No. 815269, project HARMONY.

6. REFERENCES

Armenter, R., Muller-Itten, M. and Stangebye, Z., 2021. Geometric methods for finite rational inattention.

Fosgerau, M. and Jiang, G., 2019. Travel time variability and rational inattention. *Transportation Research Part B: Methodological*, 120, pp.1-14.

Fosgerau, M., Melo, E., De Palma, A. and Shum, M., 2020. Discrete choice and rational inattention: A general equivalence result. *International economic review*, 61(4), pp.1569-1589.

Habib, N. and Mohammed, K., Rational Inattention in Discrete Choice Models: Estimable Specifications with Empirical Applications of Ri-Multinomial Logit (Ri-Mnl) and Ri-Nested Logit (Ri-Nl) Models. 2022. Available at SSRN 4182187.

Jiang, G., Fosgerau, M. and Lo, H.K., 2020. Route choice, travel time variability, and rational inattention. *Transportation Research Part B: Methodological*, 132, pp.188-207.

Maćkowiak, B., Matějka, F. and Wiederholt, M., 2021. Rational inattention: A review.

Matějka, F. and McKay, A., 2015. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1), pp.272-298.

Pappelis, D., Kamargianni, M., and Chaniotakis, E., 2022. On Modelling the Dynamics of Within-day Travel Behaviour. *Proceedings of the 102nd Annual Meeting of the Transportation Research Board*, January 2023, Washington DC, United States.

Sims, C.A., 2003. Implications of rational inattention. *Journal of monetary Economics*, 50(3), pp.665-690.

The spatial variation of travel time valuations: A general equilibrium model and application in project appraisal

Daniel Hörcher*¹ and Daniel J. Graham²

¹Research Associate, Transport Strategy Centre, Imperial College London, U.K.

²Professor and Head of the Centre for Transport Studies, Imperial College London, U.K.

SHORT SUMMARY

Current practices in transport policy appraisal are mostly restricted to partial equilibrium modelling, creating a natural need to explore new ways to understand the spatial general equilibrium impacts of transport interventions. The emerging literature of quantitative spatial models (QSM) offers new opportunities. However, the direct application of the QSM methodology in transport is hindered by the assumption of unidimensional ‘iceberg’ travel costs. Due to the presence of both temporal and pecuniary travel costs, the theoretical characterisation and empirical measurement of the monetary value of travel time savings has been a central theme of transport research for decades. We bridge a gap between spatial and transport economics by developing a quantitative spatial model with endogenous travel time valuations, revealing its previously neglected spatial heterogeneity. The model yields OD-specific values of time in spatial general equilibrium. Numerical implementation of the model highlights the relevance of our contribution in practical transport appraisal.

Keywords: transport appraisal; value of time; spatial general equilibrium.

1 INTRODUCTION

Transport appraisal models help divert heated debates on large-scale infrastructure projects to a somewhat more objective quantitative basis. Experience suggests that travel time savings is one of the biggest sources of benefits when a transport intervention reduces the distance and/or journey time between geographic locations. For this reason, the theoretical underpinning and empirical estimation of the monetary value of travel time received increased attention in transport research (Small, 2012). The standard transport appraisal methodology is often criticised, however, for its partial equilibrium approach, i.e., its inability to predict and quantify the impact of the spatial reorganisation of economic activity after the implementation of transformative transport investments (Mackie et al., 2011). Previous attempts in spatial general equilibrium transport appraisal, such as the so called ‘land use-transport interaction’ models, have not reached a consensual acceptance among economists, mainly due to the absence of microfoundations behind various assumptions on model specification and the arbitrary (theoretically inconsistent and/or statistically potentially biased) identification of model parameters.

This research reflects on recent developments at the crossroad between urban economics and economic geography: a new class of models often referred to as *quantitative spatial economics* (Redding & Rossi-Hansberg, 2017) seems to be more widely acknowledged as a tool for empirically relevant economic analysis in spatial general equilibrium. Quantitative spatial models (QSMs) are based on discrete-continuous demand models in which the choices of residential and workplace location are governed by Fréchet distributed multiplicative idiosyncratic shocks and, given the location choice, consumers optimise their consumption of housing, goods and services on a continuous scale. QSMs achieved a breakthrough in spatial economics due to very advantageous analytical properties: the existence and uniqueness of spatial equilibrium can be derived analytically and the model can be calibrated in a series of theoretically consistent econometrics exercises. Pioneering QSM papers, such as Allen & Arkolakis (2014), Ahlfeldt et al. (2015), Donaldson (2018), Monte et al. (2018), Heblich et al. (2020) and Allen & Arkolakis (2022), have been publishing in the leading journals of economics over recent years.

We believe that in the long run, QSMs may be suitable to replace the existing partial equilibrium-based transport appraisal methodologies in practical policy development as well. However, a prerequisite for that is to bring this new approach closer to transport research and align it more closely with advances in contemporary transport modelling.

In this research we focus on a common assumption of QSMs which limits their ability to replicate important characteristics of transport provision. QSMs capture the inconvenience of travel through the well-known *ad-valorem* or ‘iceberg’ formulation. This unidimensional measure of cost is not suitable to distinguish the disutility of travel time loss from pecuniary expenditure. The practical consequence is that a very fast but relatively expensive transport service may appear identical to a slow but relatively cheap alternative. Arguably, these two types of services (e.g., an expensive high-speed rail link and an affordable commuter service) may have fundamentally different structural impacts on the spatial economy. Several QSMs are based on even simpler assumptions, completely neglecting the monetary cost of travel. Such models are even less suitable to appraise pricing policies, a set of measures frequently advocated by transport economists.

This paper documents an initial attempt to capture temporal and monetary costs through separate time and money constraints facing households in the QSM framework. We derive an analytical expression of endogenous travel time valuations for each residence–workplace combination of the spatial model. This generates spatial heterogeneity in the valuation of travel time savings which is mostly neglected in the current transport appraisal practice. In a numerical implementation of the model, we provide an illustration of how this novel approach can be applied in transport policy evaluation and project ranking.

2 METHODOLOGY

The paper’s key contributions lie in the we model household preferences. Therefore, in the present short paper, we detail the demand side of the methodology, suppressing other components of the spatial model into a brief description.

Household preferences

Let us define the utility of a representative worker who resides in location i and commutes to location j as

$$U_{ij} = \left(\frac{C_{ij}}{\beta} \right)^\beta \left(\frac{L_{ij}}{\gamma} \right)^\gamma \cdot z_{ij}. \quad (1)$$

In this specification, C_{ij} denotes consumption of a variety of goods, L_{ij} is a measure of leisure time, β and γ are structural parameters, and z_{ij} is an idiosyncratic taste shock associated with the combination of locations i and j . Households are ex-post heterogeneous in their location preferences. In our notation we suppress the unique identifier of households; note, however, that z_{ij} takes a different value for each household.

Commuters face two constraints through which individual labour supply x_{ij} will affect utility. This way we adopt the modelling approach of Arnott (2007) and Hörcher et al. (2020) in a spatial setting. First, wage w_j at workplace j and the monetary price of commuting τ_{ij} determine the budget available for consumption, where P_i is the price index of the consumption variety.

$$x_{ij} (w_j - \tau_{ij}) = P_i \cdot C_{ij} \quad [\kappa] \quad (2)$$

Second, leisure time L_{ij} , time spent at work (T , exogenous), and commuting time t_{ij} cannot exceed \bar{L} , the daily time endowment of households.

$$\bar{L} = L_{ij} + x_{ij} (T + t_{ij}) \quad [\mu] \quad (3)$$

With Lagrange multipliers κ and μ , first-order condition of the optimal choice of individual labour supply implies

$$\kappa(\tau_{ij} - w_j) + \mu(T + t_{ij}) = 0, \quad (4)$$

which equates the monetary benefit of the marginal trip to work with its monetary as well as time cost. Rearrangement leads to an expression of the ratio of the marginal utilities of time and money:

$$\frac{\mu}{\kappa} = \frac{w_j - \tau_{ij}}{T + t_{ij}} = v_{ij}. \quad (5)$$

We interpret this ratio as a monetary valuation of the incremental relaxation of the worker's time endowment. We call it the (marginal) *value of time* and denote by v_{ij} . This quantity has been one of the key variables of the literature of transport economics since its emergence (DeSerpa, 1971; Small, 2012; Jara-Díaz, 2020). The value of time provides a suitable exchange rate between travel time savings and monetary expenditures, thus allowing the analyst to quantify in monetary terms the benefit of journey time reduction after transport improvements. As one would expect, the worker's wage is among the determinants of the value of time, as foregone time could always be used to earn income through work. Equation (5) also reveals that the value of time depends on the monetary and time cost of commuting as well. The core consequence from a spatial economic point of view is that the value of time will likely differ between commuters by the place of residence and work, which is often neglected in mainstream transport policy appraisal. To emphasise this feature of the model, we keep the subscripts of v_{ij} throughout the forthcoming analysis.

After simple algebraic manipulations, first-order conditions lead to the following expressions for the optimal consumption, labour supply and leisure quantities.

$$C_{ij} = \xi \frac{v_{ij}}{P_i}, \quad (6)$$

where $\xi = \left(1 + \frac{\gamma}{\beta}\right)^{-1}$. Naturally, consumption decreases with the price index at the residential location (P_i) and increases with the β parameter of our direct utility function. More surprisingly, v_{ij} enters this formula directly. That is, someone with a high value of time is expected to consume more. It may be more appropriate in the present context to interpret v_{ij} as a *net hourly wage* instead of a travel time valuation, where both the money cost and the temporal duration of commuting is part of the net wage. With this interpretation, it is more convincing that consumption increases with the net hourly wage, indeed.

Combining the first-order conditions with respect to C_{ij} and L_{ij} with the monetary budget constraint in (2) and (5), we find

$$x_{ij} = \xi(T + t_{ij})^{-1}. \quad (7)$$

This rule suggests that individual labour supply increases with the utility of consumption through β and decreases with the utility of leisure time through γ . Commuting time has a negative impact on x_{ij} . Interestingly, the gross wage cancels out in this formula, so it has no direct impact on individual labour supply under the present assumptions. The consumer problem's solution with respect to leisure time is

$$L_{ij} = \left(1 + \frac{\beta}{\gamma}\right)^{-1}. \quad (8)$$

In this simple formula, β and γ have the expected impact on leisure time, and all the remaining endogenous variables cancel.

The last four results yield the following indirect utility function for a given combination of residential and working locations.

$$u_{ij} = \left(\frac{v_{ij}/P_i}{\gamma + \beta}\right)^\beta \left(\frac{1}{\gamma + \beta}\right)^\gamma \quad (9)$$

In the specific case $\beta + \gamma = 1$, indirect utility simplifies to

$$u_{ij} = \left(\frac{v_{ij}}{P_i}\right)^\beta z_{ij} = \left[\frac{w_j - \tau_{ij}}{P_i(T + t_{ij})}\right]^\beta z_{ij}. \quad (10)$$

That is, the net hourly wage (or in a different interpretation, the marginal value of time), the local price index, the β parameter, and idiosyncratic taste are the only determinants of a residence–workplace combination's attractiveness to households.

Spatial equilibrium

In the rest of this modelling exercise, we follow standard practices in the quantitative spatial economics literature with some adjustments necessitated by the methodology introduced in the previous subsection. In particular, household heterogeneity is represented *à la* Eaton & Kortum (2002), the production side of the model follows Monte et al. (2018) while the present version of the model neglects competition in the housing market, just as in Hayakawa et al. (2021). The

main differences are routed in (i) the endogeneity of individual labour supply and (ii) the unique specification of indirect utility, as derived above.

The idiosyncratic utility shock is specified as a draw from a Fréchet distribution:

$$F_{ij}(z) = \exp(-A_i B_j z^{-\epsilon}), \quad (11)$$

where the average amenity (i.e. the scale parameter) is defined as the product of residence and workplace dependent local fundamentals A_i and B_j , and ϵ governs the spread of individual preferences. These assumptions lead to location choice probabilities that take the form of a commuting gravity equation.

$$\lambda_{ij} = \frac{A_i B_j \left[\frac{w_j - \tau_{ij}}{P_i(T + t_{ij})} \right]^{\beta\epsilon}}{\sum_r \sum_s A_r B_s \left[\frac{w_s - \tau_{rs}}{P_r(T + t_{rs})} \right]^{\beta\epsilon}} \quad (12)$$

To model the production side of the economy we follow a conventional approach in new economic geography and Monte et al. (2018) more closely. Varieties of the consumption good are produced under monopolistic competition, using labour as the sole input. A fixed factor of production and the constant marginal cost of producing a unit of one of the symmetric varieties in location j imply increasing returns to scale. Productivity is an exogenous characteristic of each location. Following the usual derivations, profit maximisation and the zero profit assumption yield an equilibrium unit price for a variety produced in j and sold in i under *ad valorem* trade cost. This setup provides a separate gravity equation of trade flows, measuring the fraction of spending in i on goods produced in j . We use the trade gravity equation to compute the vector of equilibrium wages. Finally, CES preferences and monopolistic competition yield a price index for each location that we use in the commuting gravity equation above. Unfortunately, the length limit of this short paper does not allow for a more detailed elaboration of the model.

In the numerical implementation of the model, we solve for spatial equilibrium by iteratively re-evaluating the equilibrium conditions of the model and updating the wage and price index vectors.

3 RESULTS AND DISCUSSION

Figure 1 shows the layout of our simulation framework. This toy network includes 18 locations arranged in a grid. We mimic a system of two cities connected by a transport link. Each city has a central node and eight spokes around it. The local fundamentals of the locations are set to the same values except for the dark shaded ones: productivity in nodes A5 and B5 are set to $E_{A5} = 2.0$ and $E_{B5} = 1.5$, while in the remaining locations $E_j = 1$. This implies that city A is somewhat more efficient in production. Nodes A2 and B2 feature higher amenity levels than other places; $A_{B2} = B_{A2} = 2.0$ while all the remaining amenity variables are normalised to one, thus allowing us to observe the impact of amenities by comparing A2 to A8 and B2 to B8. The attribute levels of transport links between these nodes are depicted in Figure 1. Commuting times and costs in city B are 10 percent higher than in city A, but there is no difference in intracity goods transport. Finally, commuting through the intercity link is significantly costlier than within the two cities, which is expected to limit the attractiveness of intercity residence–workplace combinations while the trade impedance is milder. We set the structural parameters of the model following typical values in the quantitative spatial economics literature.¹

Figure 2 depicts the core results of this exercise: the pattern of heterogeneity in travel time valuations. Recall from (5) that v_{ij} may potentially differ between each residence–workplace combination (or origin–destination pair, OD-pair, in transport terminology). In our toy network, each link can be used by commuters of multiple OD-pairs. To derive the mean value of time of each link, we take the flow-weighted average of the relevant i – j combinations.

¹In particular, $\sigma = 4$, $\epsilon = 6.83$, $\beta = \gamma = 0.5$, $\bar{L} = 1$, $T = 8/24$, $F = 50$, and population is $M = 10,000$.

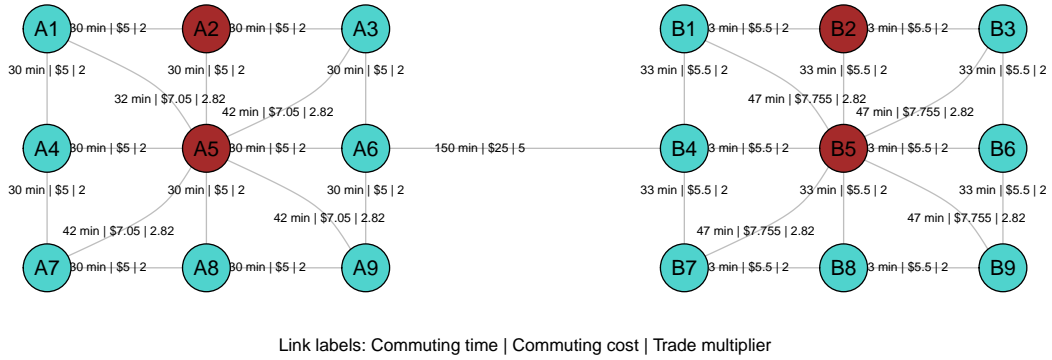


Figure 1: Network layout of the simulation framework.

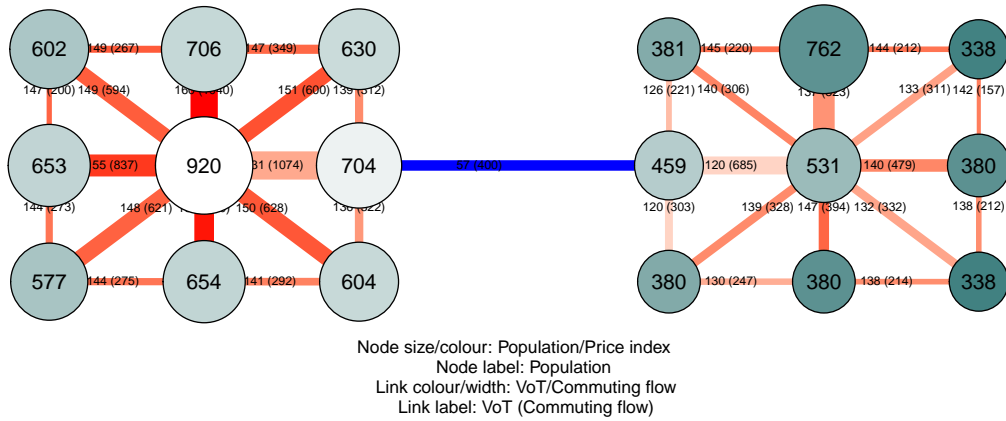


Figure 2: Travel time valuations, commuting flow, and the distribution of residential population and price indices in spatial equilibrium.

The result in Figure 2 reveals considerable heterogeneity in the average values of time by link. Time is valued generally higher in the more productive and better connected city A where the mean value of time is 157.5 as opposed to 145.4 in city B. The difference in commuting flows is even greater between the two cities. Note that the correlation between commuting flows and travel time valuation is positive but weak: 0.173 under the current set of parameters. That is, as opposed to flows, the forces behind the value of time differ from gravity between residential and workplace locations. The pattern of flows follows the regular characteristics of monocentric cities: traffic between peripheral locations is moderate due to the lower level of employment in these areas. Commuting movements are sparse between the two cities due to the relatively high travel time and cost in this market, even though the only intercity link is shared by more OD-pairs than regular intracity links. The value of time is also significantly lower for long-distance commuters who have a higher share on links A5–A6 and B4–B5.

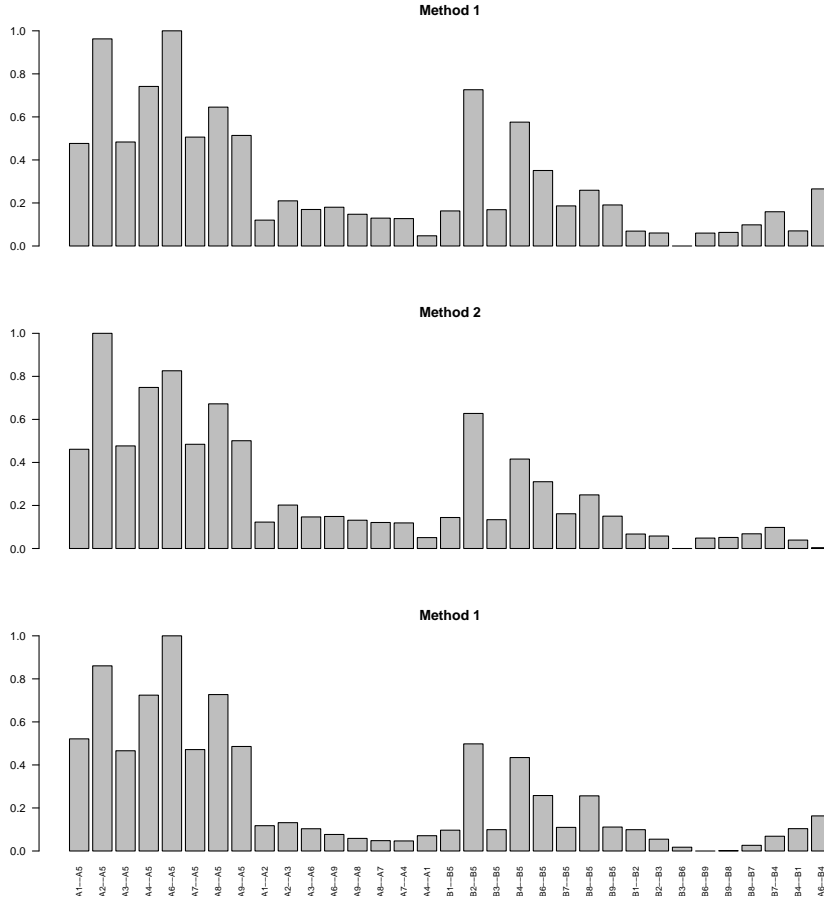


Figure 3: Relative efficiency of link-level travel time reduction according to three appraisal methods.

Let us now explore whether the spatial differentiation of the value of time may affect policy decisions in prioritising investments in a transport network. In a series of numerical simulations, let us reduce travel time on each link of our toy network one by one, by 10 minutes, and rank these potential improvements according to the economic benefit they generate. We approximate the welfare effect by three distinct methods:

Method 1 – Multiply traffic flow on each link by the 10-minute savings and the mean value of time considering every trip in the network.

Method 2 – Repeat Method 1 with the link-specific values of time derived in the previous subsection.

Method 3 – Compute the aggregate welfare effect of a link-level improvement in general equilibrium by computing the difference of aggregate indirect utilities across all OD-pairs, before and after the improvement.

Table 1: Ranking of link-level travel time reductions according to three appraisal methods

Link endpoints	Method 1	Method 2	Method 3
A6-A5	1	2	1
A2-A5	2	1	2
A4-A5	3	3	4
B2-B5	4	5	6
A8-A5	5	4	3
B4-B5	6	10	10
A9-A5	7	6	7
A7-A5	8	7	8
A3-A5	9	8	9
A1-A5	10	9	5

In Figure 3 we observe similarities in the three patterns of the relative performance of the link-level investment projects we simulated. Table 1 compares the ranking of ten alternative projects according to the three appraisal methods. Note that the divergence in policy recommendations is more substantial in terms of rankings. The general equilibrium approach agrees with the most naive method in that Links A6–A5 and A2–A5 should be prioritised first, but most of the remaining rankings differ between these methods. The project ranking of Methods 2 and 3 do not differ in more than 1 unit which reveals the value of the link-specific differentiation of the value of time (that we apply in Method 2). The are only two infrastructure segments from city B in this list, and Method 1 consistently overestimates their ranking relative to Methods 2 and 3. For example, link B4–B5 ranks 6th in the first column, but when the relatively low value of time of travellers is taken into account (Method 2) or general equilibrium impacts are considered (Method 3), this project falls back to the 10th place.

4 CONCLUSIONS

This paper builds on the emerging literature of quantitative spatial models with the aim of making this spatial general equilibrium approach more suitable to assess the economic impact of large-scale transport policies. We relax the assumption of ‘iceberg’ commuting costs that expresses the disutility of travel as a multiplier of consumer utility. We replace this assumption by integrating a leisure–labour trade-off and distinct time and money constraints into the consumer problem of the standard QSM approach. This implies that the monetary value of travel time becomes an endogenous, heterogeneous, and location-dependent outcome of the spatial equilibrium. Numerical results showcase the degree of heterogeneity in travel time valuations and the possible bias that the use of homogeneous value of time estimates imply in practice.

This manuscript documents the first stage of a research project. Subsequent stages include (i) further theoretical work, with particular attention being paid to an adequate representation of the housing market, (ii) the adaptation of our theoretical framework to granular spatial data and the analysis of real geographies, (iii) the development of empirical identification strategies to estimate structural parameters for the new model, and (iv) an analytical exploration of the properties of equilibrium/equilibria in this framework. Naturally, the endogeneity of travel time valuation is not the only precondition of the widespread adaptation of quantitative spatial models in transport policy appraisal. Ideally, QSMs should be adopted to this purpose such that they do not lose the advantageous properties they have in terms of theoretical and empirical consistency, analytical tractability and modularity. Our hope is that this research will serve as a starting point for a more widespread awareness and use of QSMs in the transport field as well.

REFERENCES

- Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6), 2127–2189.
- Allen, T., & Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3), 1085–1140.
- Allen, T., & Arkolakis, C. (2022). The welfare effects of transportation infrastructure improvements. *The Review of Economic Studies*, 89(6), 2911–2957.
- Arnott, R. (2007). Congestion tolling with agglomeration externalities. *Journal of Urban Economics*, 2(62), 187–203.
- DeSerpa, A. C. (1971). A theory of the economics of time. *The Economic Journal*, 81(324), 828–846.
- Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5), 899–934.
- Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5), 1741–1779.

- Hayakawa, K., Koster, H. R., Tabuchi, T., & Thisse, J. (2021). *High-speed rail and the spatial distribution of economic activity: Evidence from Japan's Shinkansen* (CEPR Discussion Papers No. 15771).
- Heblich, S., Redding, S. J., & Sturm, D. M. (2020). The making of the modern metropolis: evidence from london. *The Quarterly Journal of Economics*, 135(4), 2059–2133.
- Hörcher, D., De Borger, B., Seifu, W., & Graham, D. J. (2020). Public transport provision under agglomeration economies. *Regional Science and Urban Economics*, 81, 103503.
- Jara-Díaz, S. (2020). Transport and time use: The values of leisure, work and travel. *Transport Policy*, 86, A7–A13.
- Mackie, P., Graham, D., & Laird, J. (2011). The direct and wider impacts of transport projects: a review. In de Palma et al. (Ed.), *A Handbook of Transport Economics*. Edward Elgar Publishing.
- Monte, F., Redding, S. J., & Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12), 3855–90.
- Redding, S. J., & Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9, 21–58.
- Small, K. A. (2012). Valuation of travel time. *Economics of Transportation*, 1(1-2), 2–14.

Investigating Link- and Network-level Bicycle Traffic Flow Characteristics using a Microsimulation Approach

Ying-Chuan Ni^{*1}, Michail Makridis¹, and Anastasios Kouvelas¹

¹Institute for Transport Planning and Systems, Swiss Federal Institute of Technology (ETH) Zurich, Switzerland

SHORT SUMMARY

Implementing cycling infrastructure for road users has become a popular transport policy for cities to create a sustainable urban environment nowadays. A thorough understanding of bicycle traffic is required to evaluate new infrastructure designs. To fill the remaining knowledge gap in this aspect, this study aims to investigate bicycle traffic flow characteristics on dedicated bike lanes. A microsimulation tool is used to simulate various scenarios and compute bicycle traffic states. From the simulation results, bicycle flow characteristics presented at both link and network levels are identified and discussed. The findings are expected to be applied to future research regarding large-scale bicycle traffic flow modeling.

Keywords: Bicycle flow; Dedicated bike lane; Fundamental diagram; Macroscopic fundamental diagram; Microscopic traffic simulation; None-lane-based traffic

1 INTRODUCTION

Cities have been planning to expand dedicated cycling infrastructure over urban areas to enhance the usage of this active transport mode and foster a more sustainable transport environment (Pucher & Buehler, 2017). In E-Bike City project, we aim to allocate around 50% of the existing urban road space to cyclists and other slow modes (D-BAUG ETH Zurich, 2022). It is envisioned that cycling will become a primary transport mode in the city. Therefore, urban transport systems need to be re-designed to meet the growing cycling demand. However, traffic and transport modeling nowadays often regards cycling as a auxiliary mode and ignores its congestion dynamics. An in-depth and universal understanding of bicycle traffic flow characteristics is still lacking, which hinders the planning for quality cycling infrastructure.

In fact, there were plenty of studies which focused on bicycle traffic flow in the past decades. At microscopic level, bicycle flow was often simulated by cellular automata (CA), which is a simple discrete time and space model (Gould & Karner, 2009; Jiang et al., 2004). Although there has been much research endeavor attempting to overcome its discrete space limitation by introducing various extensions, CA still fails to account for the large behavioral heterogeneity nature of bicycle flow. There were also studies developing social force models which can better consider the two-dimensionality of cycling motion (Liang et al., 2018; Zhao & Zhang, 2017). However, they were known to be too computationally expensive for large-scale modeling purposes. Twaddle et al. (2014) provided a thorough review and comparison of different bicycle modeling approaches. On the other hand, microsimulation tools seem to be a good option which lies between these two types of approaches. Grigoropoulos et al. (2021) used SUMO to simulate the traffic performance of a bicycle route in various scenarios. Nevertheless, SUMO simulates bicycles through its built-in sublane function, which divides a bike lane into multiple sublanes for overtaking (Lopez et al., 2018). The suitability of this setup in simulating bicycle traffic is questioned.

There were studies which investigated bicycle traffic flow with empirical data. By performing a series of experiments, Wierbos (2021) analyzed the macroscopic bicycle flow properties, including capacity, capacity drop, jam density, and queue discharge rate in different scenarios, such as narrowing bottleneck, merging, and queuing at a stop line. However, the flow performance at congested states was not reported. Through field observation, Li et al. (2015) constructed FDs for bicycle-only paths with the presence of bottlenecks. It was found that bicycle traffic can maintain a relatively high flow rate even when the density is larger than the critical density, which was

different from the simulation results of the CA studies. However, the conclusion was based on an arbitrary curve-fitting. There was no detailed description regarding the empirical observation after the hypothetical critical density. Guo et al. (2021) also plotted FDs for the wide ring-shaped track bicycle flow experiments they conducted. A similar trend that a constant flow rate remains across a certain region of densities was discovered. It was inferred that this phenomenon was resulted from the staggered formation of bicycles at high density situations which allowed the lateral space to be utilized more efficiently. Therefore, the flow rate would not start decreasing right from the onset of congestion. Still, we remain skeptical about the outcome of such effect on FDs since (1) only few data points were generated in the study and (2) the FDs only specifically described the traffic states of small wide ring-shaped track experiments, which might be very different from the real-world cycling environment. More detailed explanations and reflection about the observed phenomena are required. Hence, the FDs of dedicated bike lanes are yet to be explored.

At the network-level, the macroscopic fundamental diagram (MFD) of a bike lane network is of interest. Little research effort has shed light on MFD for bicycle traffic in particular. Huang et al. (2021) used empirical data to investigate the impact of bicycle flow and infrastructure design on the shape of car MFD. Loder et al. (2021) first intended to capture the tri-modal interactions by using multi-modal MFDs for each mode. The bicycle MFD was generated and fitted with empirical data following the same method proposed for car MFD without careful consideration for the unique bicycle flow properties. Later on, Huang et al. (2022) applied the concept of 3D-MFD on car-bicycle traffic. Both empirical data and Vissim simulation were adopted to generate MFDs. However, there was no specific discussion on the shape and properties of the resulting bicycle MFD. To a certain extent, these studies still focused on the mixed car-bicycle traffic flow on urban roads. Besides, they were based on scarce and heterogeneous bicycle flow data collected in the field. A basic understanding of MFD for a dedicated bike lane network is still lacking.

To allocate more dedicated road space to bicycles considering its congestion dynamics within an urban network, the aggregated bicycle traffic needs to be precisely described. FD and MFD are two proper ways to model the network traffic performance. Hence, this study seeks to investigate the characteristics of bicycle traffic flow by generating its FDs and MFDs.

2 METHODOLOGY

This section first describes the selection of microsimulation tool and the calibration of bicycle-related simulation parameters. In the second part, the simulation environments and output analysis methods are explained.

As mentioned in section 1, empirical bicycle flow data containing complete traffic states are still lacking. Therefore, instead of relying on empirical data, a microsimulation approach is adopted in this study to investigate bicycle traffic characteristics at the aggregated level. Among all the traffic simulation tool, PTV Vissim is believed to possess the most sophisticated bicycle simulation function, which is not a simple projection of car traffic (PTV Group, 2023).

Compared to car driving behavior, tactical-level decisions play an even bigger role in cycling motion. In addition, bicycle traffic is not regulated by lane markings, which makes it more complex. Vissim outperforms other simulation tools by including a built-in lateral model, which enables it to simulate the special features of none-lane-based traffic, such as the overtaking behavior of two-wheelers. Moreover, the diamond queue function represents the cyclists by a diamond shape and makes the queue configuration more realistic, as pointed out in Gavriilidou et al. (2019). This also influences the resulting standstill (jam) density and queue discharge rate. Vissim also allows users to set the look ahead/back distance and number of interacting objects of each agent. This function is helpful for simulating road users, like cyclists, which have better anticipation ability.

Although there may still be several detailed cycling behavior features which are not presented in the bicycle simulation function in Vissim, it is hypothesized in this study that the function is sufficient to reproduce the bicycle traffic dynamics at aggregated levels. All the bicycle modeling parameters which need to be calibrated, including desired acceleration distribution, desired speed distribution, car-following model parameters, and lateral model parameters, follow the setup suggested in Kathis et al. (2021).

To plot bicycle flow FDs, eight 300-m-long dedicated bike lanes with three different lane widths, 1.5 m, 2 m, and 2.5 m, are built in Vissim. Bottlenecks are placed in the middle of lanes 4 ~ 8 to generate congested situations. Two types of bottleneck are considered. The first type of bottleneck on lanes 4 ~ 6 reduces the path width by 0.5 m, while the second is 1-m-wide on lanes 7 ~ 8. Table 1 lists all the created bike lanes. The one-minute-aggregated density, speed, and flow data are obtained from a 10-m-long segment before the bottlenecks on these bicycle paths. A five-hour scenario with varying demand profile is implemented.

Table 1: Dedicated bike lanes for FD generation

Dedicated bike lane no.	Lane width (m)	Bottleneck width (m)
1	1.5	none
2	2.0	none
3	2.5	none
4	1.5	1.0
5	2.0	1.5
6	2.5	2.0
7	2.0	1.0
8	2.5	1.0

In addition to FDs, this study also aims to derive a bicycle flow MFD. An arterial with 2-m-wide bike paths in two directions on both sides of the road is built in Vissim. Six intersections divide the arterial into seven road sections. The green time at every intersection is 40 s, while the cycle time equals to 70 s. Each road link is 150 m long. Accordingly, the signal offset is set to 35 s. The average flow and density data are collected every five minutes from the ten road links in the middle, excluding the inflow and outflow links. A fifteen-hour scenario is designed to mimic the demand profile of a typical weekday, while the first half-an-hour is considered a warm-up period for the network to be filled up.

3 RESULTS AND DISCUSSION

This section describes the analyzed simulation results and research findings based on the results.

Bicycle flow fundamental diagrams

Figure 1 shows the density-flow FDs of bike paths with different widths. In order to compare the FDs of different lane widths in a convenient way, the density value is transformed into two-dimensional. The y-axis also becomes flow rate per 1-m-width so that it can align with the transformed density. By doing so, we keep the speed information (slope) meaningful.

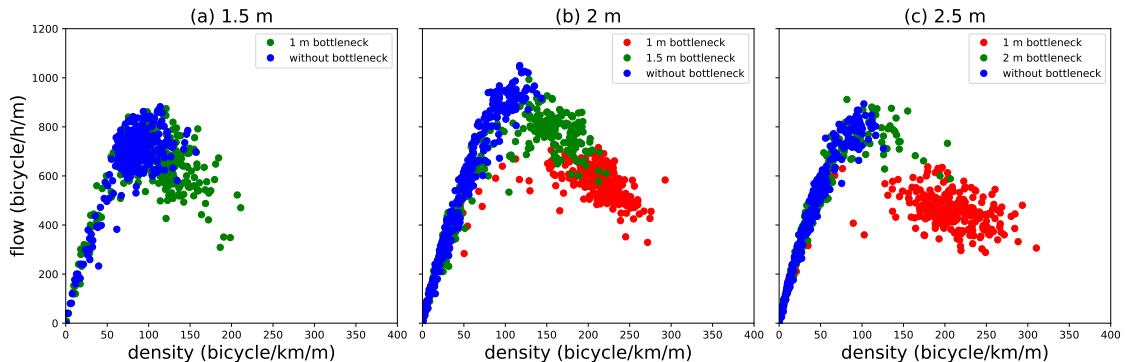


Figure 1: Flow-density FDs of bicycle flow on bike lanes with different lane widths

The effect of different lane widths can then be discussed. As can be seen in Figure 1a, few data points in the decreasing branch can be observed in the 1 m bottleneck case. Only little congestion effect is shown, which indicates that bicycle traffic flow on 1.5 m width does not have significant difference with the flow on 1 m width. Still, it implies that bicycles in free flow condition need more than 1-m-wide lane space to execute overtaking maneuvers. The narrowing bottleneck does impact the traffic volume. In addition, a large scatter of blue dots can also be observed near the critical density. This may be resulted from the stochastic stop-and-go disturbance caused by the desired speed heterogeneity.

Figure 1b plots the FD of the bicycle flow on 2-m-wide lane. Compared to the 1.5-m-wide lane, the capacity is increased, indicating that there are more overtaking behaviors. By looking at the blue dots in the free flow branch, one can see that it does not have the same large scatter near the critical density as shown in the 1.5-m-width case, indicating that a wider lane width can handle disturbance better. Furthermore, the green data points become more obvious in the decreasing branch. Even though the lane width reduction is the same, the congestion effect is more significant when the path width is 2 m before the bottleneck. This indicates that a 2-m-wide bicycle lane allows cyclists to utilize the existing lane space more efficiently. By looking at the red dots, one can also see that the capacity drop becomes more significant when the bottleneck is only 1 m wide, which is caused by the larger lane width narrowing.

The FD of the bicycle flow on 2.5-m-wide lane, as shown in Figure 1c, has a smaller capacity per unit lane width than the lane with 2 m width does, which seems to be non-intuitive. This means that widening the lane width does not necessarily increase the degree of lane space utilization. On the other hand, no obvious decreasing branch can be found in the case of the 2-m-wide bottleneck (green dots). These both indicate that the capacity of the bike lane with 2 m width is closer to the performance limit of dedicated bicycle infrastructure compared to the other two lane widths.

Table 2 summarizes the FD attributes of each lane width. Note that the free flow speed is not included here and will be discussed later. The capacity is determined by the point with the largest flow rate, while the critical density is the density of the capacity point. The jam density values are obtained by running simulations and placing a stop line on each lane. As can be seen in the table, the jam density values are different in cases of different lane widths, which also demonstrates the different degrees of lane space utilization.

Table 2: Shape attributes of FDs of bike lanes with different widths

	1.5 m width	2 m width	2.5 m width
Capacity (bicycle/h/m)	882.24	1049.91	893.53
Critical density (bicycle/km/m)	113.92	117.32	102.23
Jam density (bicycle/km/m)	408.00	460.50	410.8

To sum up, bicycle flow on 2-m-wide bike lane can utilize the lane space most efficiently. There is no improvement when widening the lane width to 2.5 m. This can also be understood by considering that the lateral space required by a cyclist is between 1 m and 1.3 m in the adopted Vissim setup.

Other than the difference between various lane widths, there are a few special characteristics of non-lane-based bicycle traffic flow regarding the free flow speed. As can be seen in Figure 2a, which is the speed-density FD of the bike lane with 2 m width in the case without bottleneck, the speed decreases rapidly as the density increases. This is the result of large desired speed heterogeneity. Bicycle flow is significantly slowed down by slow cyclists. Table 3 computes the average speed values in six divided density ranges. It is believed that this phenomenon is more significant in bicycle traffic than in car traffic. In addition to the speed heterogeneity, the effect of overtaking is another special characteristics of bicycle flow which was often overlooked. Figure 2b shows the results of a single-file bicycle traffic flow with no overtaking allowed. It can be observed that the speed decreases more rapidly as the density increases than in the case with overtaking. In the single-file case, cyclists loss the possibility to overtake slow cyclists, which further degrades the bicycle flow performance. Therefore, it is important to consider these two aspects when modeling bicycle traffic flow.

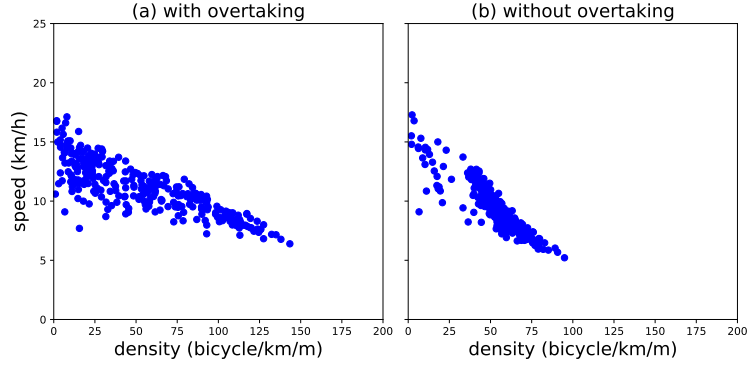


Figure 2: speed-density FDs of bicycle flow on a 2-m-wide bike lane (a) with overtaking (b) without overtaking

Table 3: Free flow speeds of bike lanes in six density ranges

Density (bicycle/km/m)	Average free flow speed (km/h)		
	1.5 m width	2 m width	2.5 m width
0 ~ 25	13.85	13.19	13.50
25 ~ 50	10.66	11.75	11.98
50 ~ 75	10.09	10.83	10.67
75 ~ 100	8.38	9.82	8.81
100 ~ 125	6.78	8.32	7.44
125 ~ 150	5.28	7.19	—

On the other hand, the existence of a constant flow branch around critical density pointed out by the previous studies cannot be observed from the FDs in Figure 1. It is suspected that the constant flow may actually stem from the better anticipation behavior of cyclists on the short ring-shaped track. The phenomenon may not appear in FDs of bike lanes in the real-world.

Bicycle flow macroscopic fundamental diagram

Figure 3 presents the MFD of the built bicycle lane arterial. The density variation is shown by the color of each data point. From the figure, one can observe an MFD curve with low scatter, showing the relationship between accumulation and network traffic performance can also be applied to bicycle flow. There are only few deviated data points which are caused by hysteresis, as shown by the color representing the normalized standard deviation of density. Since the simulation was carried out on a homogeneous arterial with equal link lengths and signal timing plans, no obvious decreasing branch can be observed.

The MFD is further analyzed by using the method of cuts (MoC) proposed by Daganzo and Geroliminis (2008). Each practical cut can be generated according to a moving observer speed and the number of blocks the observer can pass through γ . In the previous subsection, it was mentioned that the free flow speed varies a lot across different density values for bicycle flow. Therefore, this study modifies the MoC for bicycle traffic. The free flow speeds calculated in Table 3 are used to generate cuts. For each γ value, there is one corresponding speed. The saturation flow rate is determined by several Vissim simulation runs. Note that unlike typical car traffic, the maximum flow rates at a free flow lane segment and a stop line can be slightly different due to the staggered queue formation. No cuts in the decreasing branch are generated since there is no congested traffic state in the designed homogeneous arterial scenario. Figure 4 shows the cuts derived from the modified method. Compared to the examples which use only a single average speed value in Figure 5, the modified method produces more accurate cuts.

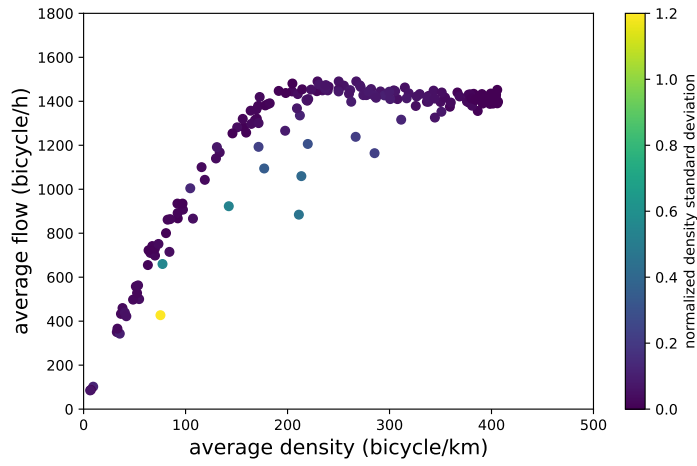


Figure 3: MFD of the bike lane arterial

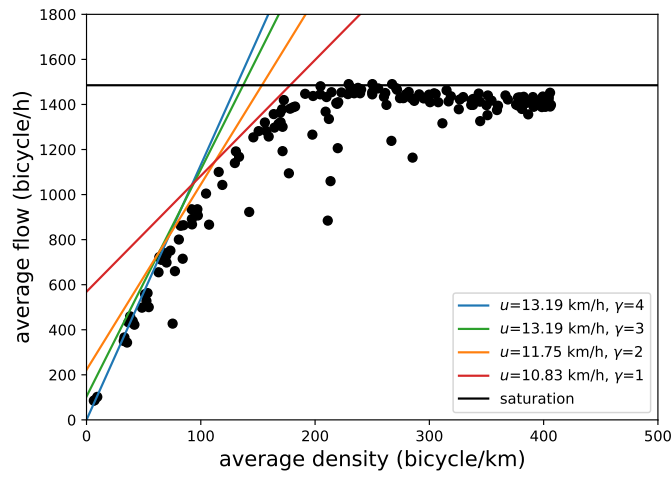


Figure 4: MFD of the bike lane arterial with cuts generated by the modified MoC

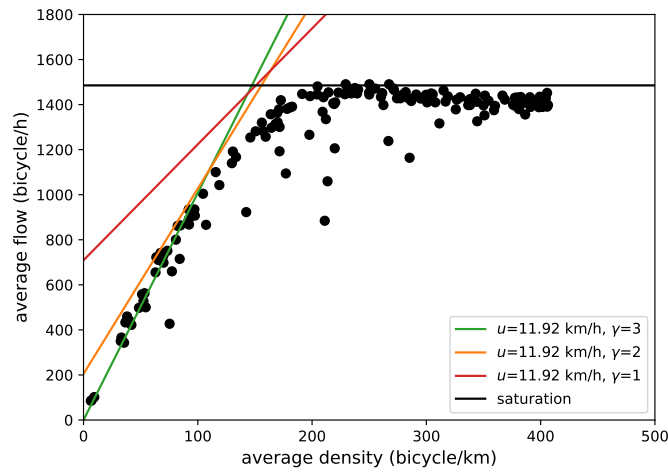


Figure 5: MFD of the bike lane arterial with cuts generated by a single speed value

In addition, a little decreasing trend can be observed after the critical density in the MFD although no spillback situation occurs in the simulated scenario. The decreased average flow rate may be the effect of anticipative cycling behavior. The speed of bicycles discharging from intersections decreases when there are too many queuing cyclists at the downstream links. This is a special phenomenon resulted from the look ahead behavior designed in Vissim. In this setup, cyclists look further downstream to determine their following speed.

4 CONCLUSIONS

This study investigates the characteristics of bicycle traffic flow on dedicated infrastructure by looking at FDs of different lane widths and an MFD of a bike lane arterial which are generated from microsimulation. The identified features are discussed and compared with the findings in the previous studies pertaining to bicycle flow modeling.

At link-level, it is found that lane widths greatly influences the FD attributes. A proper width can be found based on the lateral space required by a cyclist so that they can utilize the lane space in the most efficient manner. At free flow situations, the speed decreases as the density increases due to the desired speed heterogeneity. However, the overtaking behavior mitigates such effect by enabling bicycles to utilize the lateral space. On the other hand, the existence of a constant flow branch, which was pointed out in the previous ring-shaped track experiment, on real-world bike lanes is questioned.

This study also applies the MoC to derive the upper-bound MFD for bicycle traffic on dedicated infrastructure. To the best of our knowledge, this is the very first attempt to discuss bicycle MFD by using carefully-examined bicycle flow attributes. It is found that the modified method can better capture the shape of an MFD for none-lane-based bicycle traffic where the behavioral heterogeneity is large.

The results in this study largely depend on the setup in the microsimulation tool. Although the representativeness of the calibrated built-in bicycle simulation model may be argued, the overall macroscopic characteristics of bicycle flow observed from the FDs and MFD are deemed to be valid. The findings can be used for more complex network-wide bicycle traffic modeling, such as cell transmission model or MFD-based traffic assignment, and therefore assist dedicated cycling infrastructure planning by transport authorities which seek to increase bicycle usage.

ACKNOWLEDGEMENTS

This work was supported by E-Bike City project of the Department of Civil, Environmental and Geomatic Engineering at ETH Zurich.

REFERENCES

- Daganzo, C. F., & Geroliminis, N. (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, 42(9), 771–781. <https://doi.org/10.1016/j.trb.2008.06.008>
- D-BAUG ETH Zurich. (2022). E-Bike City: Designing sustainable streets. <https://ebikecity.baug.ethz.ch/en/>
- Gavriilidou, A., Daamen, W., Yuan, Y., & Hoogendoorn, S. (2019). Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour. *Transportation Research Part C: Emerging Technologies*, 105, 468–484. <https://doi.org/10.1016/j.trc.2019.06.012>
- Gould, G., & Karner, A. (2009). Modeling Bicycle Facility Operation. *Transportation Research Record: Journal of the Transportation Research Board*, 2140(1), 157–164. <https://doi.org/10.3141/2140-17>
- Grigoropoulos, G., Hosseini, S. A., Keler, A., Kathis, H., Spangler, M., Busch, F., & Bogenberger, K. (2021). Traffic Simulation Analysis of Bicycle Highways in Urban Areas. *Sustainability*, 13(3), 1016. <https://doi.org/10.3390/su13031016>

- Guo, N., Jiang, R., Wong, S., Hao, Q.-Y., Xue, S.-Q., & Hu, M.-B. (2021). Bicycle flow dynamics on wide roads: Experiments and simulation. *Transportation Research Part C: Emerging Technologies*, 125, 103012. <https://doi.org/10.1016/j.trc.2021.103012>
- Huang, Y., Sun, D., Li, A., & Axhausen, K. W. (2021). Impact of bicycle traffic on the macroscopic fundamental diagram: some empirical findings in Shanghai. *Transportmetrica A: Transport Science*, 17(4), 1122–1149. <https://doi.org/10.1080/23249935.2020.1832157>
- Huang, Y., Sun, D., & Zhang, S. (2022). Three-dimensional macroscopic fundamental diagram for car and bicycle heterogeneous traffic. *Transportmetrica B: Transport Dynamics*, 10(1), 312–339. <https://doi.org/10.1080/21680566.2021.1994050>
- Jiang, R., Jia, B., & Wu, Q.-S. (2004). Stochastic multi-value cellular automata models for bicycle flow. *Journal of Physics A: Mathematical and General*, 37(6), 2063–2072. <https://doi.org/10.1088/0305-4470/37/6/007>
- Kaths, H., Keler, A., & Bogenberger, K. (2021). Calibrating the Wiedemann 99 Car-Following Model for Bicycle Traffic. *Sustainability*, 13(6), 3487. <https://doi.org/10.3390/su13063487>
- Li, Z., Ye, M., Li, Z., & Du, M. (2015). Some Operational Features in Bicycle Traffic Flow. *Transportation Research Record: Journal of the Transportation Research Board*, 2520(1), 18–24. <https://doi.org/10.3141/2520-03>
- Liang, X., Xie, M., & Jia, X. (2018). New Microscopic Dynamic Model for Bicyclists' Riding Strategies. *Journal of Transportation Engineering, Part A: Systems*, 144(8). <https://doi.org/10.1061/JTEPBS.0000148>
- Loder, A., Bressan, L., Wierbos, M. J., Becker, H., Emmonds, A., Obee, M., Knoop, V. L., Menendez, M., & Axhausen, K. W. (2021). How Many Cars in the City Are Too Many? Towards Finding the Optimal Modal Split for a Multi-Modal Urban Road Network. *Frontiers in Future Transportation*, 2. <https://doi.org/10.3389/ffutr.2021.665006>
- Lopez, P. A., Wiessner, E., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flotterod, Y.-P., Hilbrich, R., Lucken, L., Rummel, J., & Wagner, P. (2018). Microscopic Traffic Simulation using SUMO. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2575–2582. <https://doi.org/10.1109/ITSC.2018.8569938>
- PTV Group. (2023). PTV Vissim. <https://www.myptv.com/en/mobility-software/ptv-vissim>
- Pucher, J., & Buehler, R. (2017). Cycling towards a more sustainable transport future. *Transport Reviews*, 37(6), 689–694. <https://doi.org/10.1080/01441647.2017.1340234>
- Twaddle, H., Schendzielorz, T., & Fakler, O. (2014). Bicycles in Urban Areas. *Transportation Research Record: Journal of the Transportation Research Board*, 2434(1), 140–146. <https://doi.org/10.3141/2434-17>
- Wierbos, M. (2021). *Macroscopic Characteristics of Bicycle Traffic Flow* (Doctoral dissertation). Delft University of Technology.
- Zhao, Y., & Zhang, H. (2017). A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic. *Transportation Research Part B: Methodological*, 105, 315–327. <https://doi.org/10.1016/j.trb.2017.09.004>

Dynamics of the Ride-Sourcing Market: A Coevolutionary Model of Competition between Two-Sided Mobility Platforms

Farnoud Ghasemi*¹, Arkadiusz Drabicki², Rafał Kucharski³

¹ PhD Candidate, Mathematics and Computer Science, Jagiellonian University, Poland

² PhD Candidate, Department of Transport Systems, Cracow University of Technology, Poland

³ Assistant Professor, Jagiellonian University, Poland

SHORT SUMMARY

There is a fierce competition between two-sided mobility platforms (e.g., Uber and Lyft) fueled by massive subsidies, yet the underlying dynamics and interactions between the competing platforms are largely unknown. These platforms rely on the cross-side network effects to grow, they need to attract agents from both sides to kick-off: travellers are needed for drivers and drivers are needed for travellers. We use our coevolutionary model featured by the S-shaped learning curves to simulate the day-to-day dynamics of the ride-sourcing market at the microscopic level. We run three scenarios to illustrate the possible equilibria in the market. Our results underline how the correlation inside the ride-sourcing nest of the agents choice set significantly affects the platforms' market shares. While late entry to the market decreases the chance of platform success and possibly results in "winner-takes-all", heavy subsidies can keep the new platform in competition giving rise to "market sharing" regime.

Keywords: Two-sided mobility, Ride-sourcing, S-shaped learning, Platform competition, Agent-based simulation

1. INTRODUCTION

Ride-sourcing companies such as Uber and Lyft have achieved significant market share in a short time through the two-sided platform business model. The reason underlying such a tremendous potential to grow in two-sided markets is the power of network. The classic definition by Rochet and Tirole (2006), characterizes them as the markets in which one or several platforms enable interactions between end-users and try to get both sides on board by appropriately charging each side. The platforms associated with these markets rely on the critical mass required for their self-sustainable operations and the network effects to induce growth (Belleflamme & Peitz, 2016). Platforms apply various market entry strategies in the early adaptation phase to follow a desired growth pattern that includes different stages from their launch up to maturity.

Even though ride-sourcing platforms have the potential to grow rapidly, they fiercely compete over the common pool of travelers and drivers. Considering the so-called multi-homing characteristics of the market, in which users can move from one platform to another with ease, it becomes extremely challenging for the platform to gain and hold the market share. As for the platforms new to the market, they need to first induce the interactions between the decentralized supply and demand to reach market shares sufficient to trigger the cross-side network effects. On the other side, existing platforms with stable market share adjust their strategies in accordance with other platforms to avoid losing market share. These platform strategies are mostly controlled by subsidies and implemented through e.g., discounts, and incentives.

There is a wide body of research on the competition between the platforms. Relying on the game theory, Zhang and Nie (2021) study ride-sourcing market in which two platforms compete with each other, as well as with transit, and Ahmadinejad et al., (2019) examine the competition impact on ride-sourcing parties by adjusting the trip fare. Using analogous methodology, Siddiq and Taylor (2022) throw light on the importance of autonomous vehicles for platforms' profitability. In different approaches, Shoman and Moreno (2021), conduct a stated preference analysis to find the ride-sourcing impact on the modal split in city of Munich.

Study approach and contribution

Previous studies are either equilibrium-based or assume fixed demand and/or supply and they neglect the interactions between the parties driving the complex system evolution. Here, we illustrate with our experiments, an adequate framework to realistically model the platform competition in two-sided model with subsidizing strategies is missing.

Previously, we proposed a novel, microscopic co-evolutionary model which is capable of reproducing platform's growth mechanism day-to-day. The key element of the model is the S-shaped learning curves which enable the agents to adapt and stabilize their behavior and yet to remain sensitive to the changes in their environment (Ghasemi and Kucharski, 2022). In this study, we extend the previous model with platform competition considering the multi-homing characteristics of the system. We incorporate nested choice modelling to examine the correlation between platforms and the possible equilibria in the ride-sourcing market.

2. METHODOLOGY

We model two-sided mobility market with MaaSSim¹ agent-based simulator (Kucharski and Cats, 2022), extended here with a coevolutionary model to represent the day-to-day dynamics of two-sided mobility market. We simulate two classes of agents representing two sides of the system and a platform as an intermediate agent matching the demand to the supply. A pool of travelers and drivers, who are not formerly notified about our ride-sourcing platform, gradually become aware of the ride-sourcing. When an agent gets notified, he/she may decide to participate in the market – i.e., supply the demand as a driver or travel to his/her destination as the platform client. With the participation, agents start to learn and adapt their behavior through endogenous and exogenous factors.

Platform

Platform executes the strategy S_t with the control levers, namely: trip fare f_t , commission rate c_t , discount d_t and marketing m_t for each day t of the simulation.

$$S_t = \{f_t, c_t, d_t, m_t\} \tag{1}$$

Traveller

Each notified traveler r on day t selects between alternatives from the choice set $C_r = \{rs, pt\} = \{\{p_1, p_2\}, pt\}$ including public transport (pt) and two ride-sourcing (rs) platforms (p_1, p_2). While the utility of public transport is fixed (formulated with a typical access/egress, waiting times,

¹ <https://github.com/RafalKucharskiPK/MaaSSim>

transfers, etc.), the platforms' utility is composed of multiple components, each adjusted day-to-day (as detailed in the upcoming sections).

Driver

Analogous to travellers, each notified driver d makes a choice from the choice set $C_d = \{rs, pt\} = \{\{p_1, p_2\}, rw\}$ which includes: working for ride-sourcing (rs) platform one (p_1) and platform two (p_2), and working elsewhere, for a fixed reservation wage (rw). The utility of working for platform as a driver is composed of same components of travellers and adjusted day-to-day as detailed below.

Choice Utility

For any notified agent i , we propose the generic perceived utility (U) formulation composed of three components: experience (U^E), marketing (U^M) and word of mouth (U^{WOM}):

$$U_{i,t} = \beta_i^E \cdot U_{i,t-1}^E + \beta_i^M \cdot U_{i,t-1}^M + \beta_i^{WOM} \cdot U_{i,t-1}^{WOM} + ASC + \varepsilon_i \quad (2)$$

Agents every day (t) choose based on experiences collected until the previous day ($t-1$). Experienced utility is endogenous and comes directly from the simulation: drivers experience the actual incomes and operating cost, travelers experience travel time, waiting time and trip fare. Marketing is an exogenous factor being positive or negative (e.g., recent Uber scandal). Word-of-mouth is shared among agents over the social network. The β 's in the formula reflect the relative weights of respective utility components (ensuring that $\beta_i^E, \beta_i^M, \beta_i^{WOM} > 0$ and $\beta_i^E + \beta_i^M + \beta_i^{WOM} = 1$). The ASC captures the effect of unobserved factors on the perceived utility of alternatives and ε_i is the random utility error term. In such form, the utility is consistent with the discrete choice theory and can be applied e.g., in the logit model.

S-shaped learning and adaptation

The key element of the proposed model is the following adjustment mechanism which allows us to realistically represent the agents' dynamics specific to the platform growth. Agents learn and adapt their choice day-to-day based on the perceived utility components. Here, instead of exponential memory curve (used e.g., in de Ruijter et al., 2021), we follow Murre (2014) and propose a more adequate formulation of the so-called S-shaped learning curve in the context of urban mobility (Ghasemi and Kucharski, 2022). Fig. 1 provides a basic idea of our model.

The adjustment process can be seen as moving each of the utility components along the S-shaped curve with each of the utility components. Positive experience increases the experienced utility pushing the perception towards upper tail, and negative experience decreases it pushing the perception towards lower tail. The two extreme points (lower and upper endings) of curve represent absolutely negative and positive attitudes and learning can go both directions on any day. Triggered by the consecutive positive/negative experiences, learning proceeds slowly for the agents who already have sharp, extreme opinions, and is fast for the neutral agents. To this end, on the contrary to state-of-the-art models, we can stabilize the agents' behavior and, at the same time, remain sensitive to the system changes.

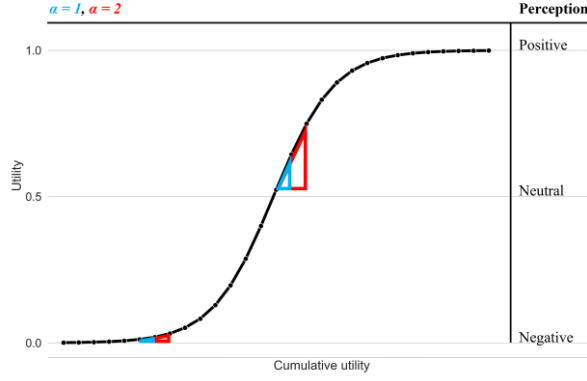


Figure 1: S-shaped curves used for the day-to-day learning process. The adjustment volume not only depends on sensitivity α but also the signal strength (Δu) and the position on the S-shaped curve.

Technically, we obtain the utility of respective components $c \in \{E, WOM, M\}$ after day t ($U_{i,t}^c$) as follows. First, we retrieve the cumulative utility on the previous day $t - 1$ by applying the inverse sigmoid function (eq. 3). Then, we update it with the difference coming from today t (eq. 4), i.e. the signal strength. We weight it with the learning speed parameter α (determining step size on the S-shaped curve). Eventually, to obtain the updated utility at the end of day t we use sigmoid (logistic) function with shape parameter β (5):

$$CU_{i,t-1}^c = \ln\left(\frac{1}{U_{i,t-1}^c} - 1\right) \quad (3)$$

$$CU_{i,t}^c = CU_{i,t-1}^c + \alpha \cdot \Delta u_{i,t}^c \quad (4)$$

$$U_{i,t}^c = \frac{1}{1 + \exp(\beta \cdot CU_{i,t}^c)} \quad (5)$$

The above formulation is generic to represent various kinds of learning new experiences and exposure to effects. For the purpose of this study, we introduce specific formulas for three components of utility.

For the experience we adjust the utility as follows. Experienced cumulative utility of drivers d on day t for platform p is updated with the relative difference between the reservation wage (RW_d) and the income experienced on that day (eq. 6). Similarly, traveler r adjusts his/her experienced cumulative utility on day t (eq. 7) according to relative difference between the experienced utility of the platform p (as a function of waiting time, travel time, and trip fare) and the public transport (pt).

$$\Delta u_{d,t}^{E,p} = \frac{RW_d - E_{d,t}}{RW_d} \quad (6)$$

$$\Delta u_{r,t}^{E,p} = \frac{U_r^{pt} - E_{r,t}}{U_r^{pt}} \quad (7)$$

The marketing spreads uniformly among all the agents (target clients) and accumulates in time over the period of the marketing campaign. While marketing is constant before and after the campaign, it produces a positive effect on each exposure. The chance of agent i to be exposed to the

marketing on day t depends on the campaign intensity (p_i^M , from $[0,1]$ range). We update the cumulative utility for marketing as follows:

$$\Delta u_{i,t}^{M,rs} = p_i^M (U_{i,t}^M - 1) \quad (8)$$

For the word-of-mouth, we assume pairwise interactions through the social network with agents, who share their perceived utility with each other. Analogically to the marketing, the WOM intensity ($p_{i,j}^{WOM}$) determines the likelihood of agent i to share his/her opinion with agent j on day t . Influenced by the exchange of views with their peers, agents adjust their cumulative utility of word-of-mouth as follows:

$$\Delta u_{i,t}^{WOM,rs} = p_{i,j}^{WOM} (U_{i,t}^{WOM} - U_{j,t}) \quad (9)$$

Participation probability

An agent starts considering a ride-sourcing platform in her mode choice set (C_i) only after being notified about it. Due to correlation between the ride-sourcing platforms, we apply the nested logit model. We assume the agent first selects between the alternative (pt/rw) and ride-sourcing, and when she selected ride-sourcing she choose among the competing platforms that she is notified about. The participation probability of notified agents is updated every day and depends on the perceived utility of alternatives as follows. The probability of choosing alternative k (eq. 13) is the product of probability of k inside the nest (eq. 10) and the probability of nest n (eq. 12) based on the expected maximum utility of nest (W^n). $I_{i,t}^k$ is a binary variable switching from zero to one when agent gets notified about alternative k . The scale parameters are θ (at the upper choice level) and θ_n (within the ride-sourcing nest) which allows us to calculate the correlation inside the nest ($\rho \in [0,1]$) as: $\rho = 1 - \frac{\theta_n}{\theta}$.

$$P_{i,t}^{k/n} = I_{i,t}^k \frac{\exp(\frac{u_{i,t}^k}{\theta_n})}{\sum_{k' \in K} \exp(\frac{u_{i,t}^{k'}}{\theta_n})} \quad (10)$$

$$W_{i,t}^n = \theta_n \cdot \log(\sum_{k' \in n} \exp(\frac{u_{i,t}^{k'}}{\theta_n})) \quad (11)$$

$$P_{i,t}^n = \frac{\exp(\frac{W_{i,t}^n}{\theta})}{\sum_{n' \in N} \exp(\frac{W_{i,t}^{n'}}{\theta})} \quad (12)$$

$$P_{i,t}^k = P_{i,t}^{k/n} \cdot P_{i,t}^n \quad (13)$$

Experimental design

We experiment on Amsterdam, with 2000 travelers and 200 drivers. The reservation wage of drivers is assumed 10.63[€/hour]. The operational costs (fuel, depreciation costs, etc.) of the drivers amount to 0.25 [€/km]. Each run simulates 4 hours of interactions between the parties. Vehicle speed is set to the flat 36 [km/h]. We consider the ride-sourcing fare of 1.2 [€/km] with a minimum of 2 [€] (based on the Uber price estimator). The utility weights of three main component are fixed as: $\beta^E = 0.7$, $\beta^{WOM} = 0.2$, and $\beta^M = 0.1$, while Marketing and WOM intensity are set to $p_i^M = p_{i,j}^{WOM} = 10\%$. We assumed the value of time 10.63 [€/hour] to compute the experienced utility of travellers.

3. RESULTS AND DISCUSSION

We first illustrate how the single platform competes against public transport in the period of one year (Figure 1). As agents get notified, they are initially reluctant to select the new travel mode. Yet, once travellers try out the platform and experience its benefits (40% discounts), they start to adapt and use ride-sourcing more frequently. This provides adequate income for the drivers. As the network gets denser on both demand and supply sides, it generates greater cross-side network effects. These effects provide both travellers and drivers with extra utilities in terms of lower waiting times and higher incomes, respectively. Thanks to this, the system grows and stabilizes around day 200.

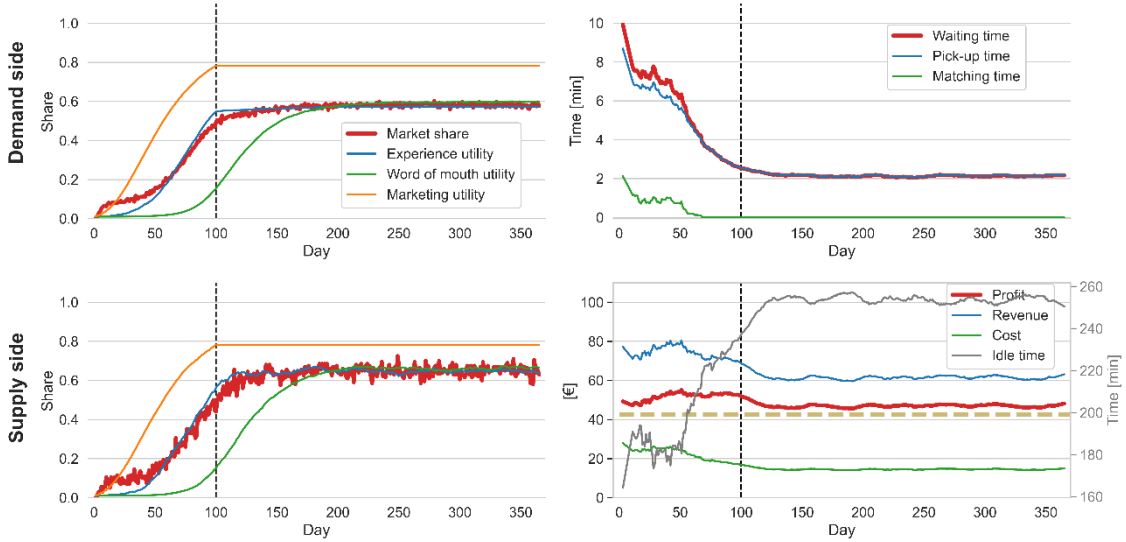


Figure 2: Single platform evolution with the baseline strategy. Platform applies marketing campaign and 40% discount on trip fares only in the initial 100 days (vertical dashed line), while commission rate is fixed to 10%.

Correlation between platform alternative and the market shares

Next, we introduce a second platform to the system and apply nested logit for the agents' mode choice. We investigate how the assumption of correlations in the nested choice model affects the equilibrium market shares. (fig. 3).

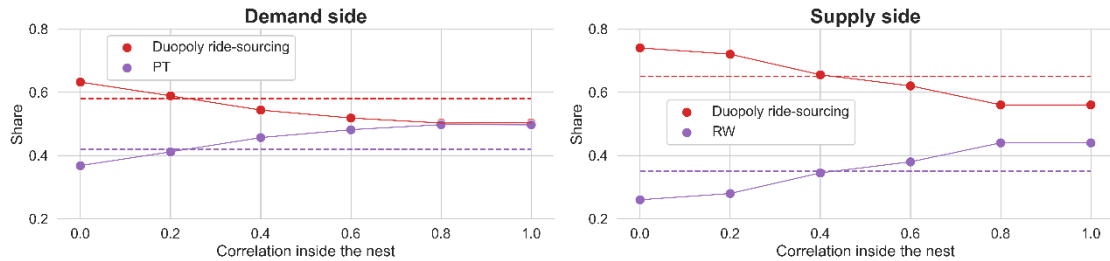


Figure 3: Ride-sourcing and PT market shares with different correlations between ride-sourcing alternatives. As correlation increases, total market share of ride-sourcing (both platforms) decreases. The dashed lines represent the ride-sourcing market share in the monopoly market with single platform.

Three scenarios and relevant equilibria

We fix the correlation rate inside the ride-sourcing choice nest at a moderate value of 0.4, and demonstrate the three scenarios with predetermined strategies as illustrated in fig. 3 to examine the emerging equilibria. In the first scenario, both platforms (P1, P2) apply baseline strategy, P2 launches 25 days later in the second scenario without any strategy change. In the third scenario, the P2 enter later, but aggressively, with the 80% discount starting form day 25.

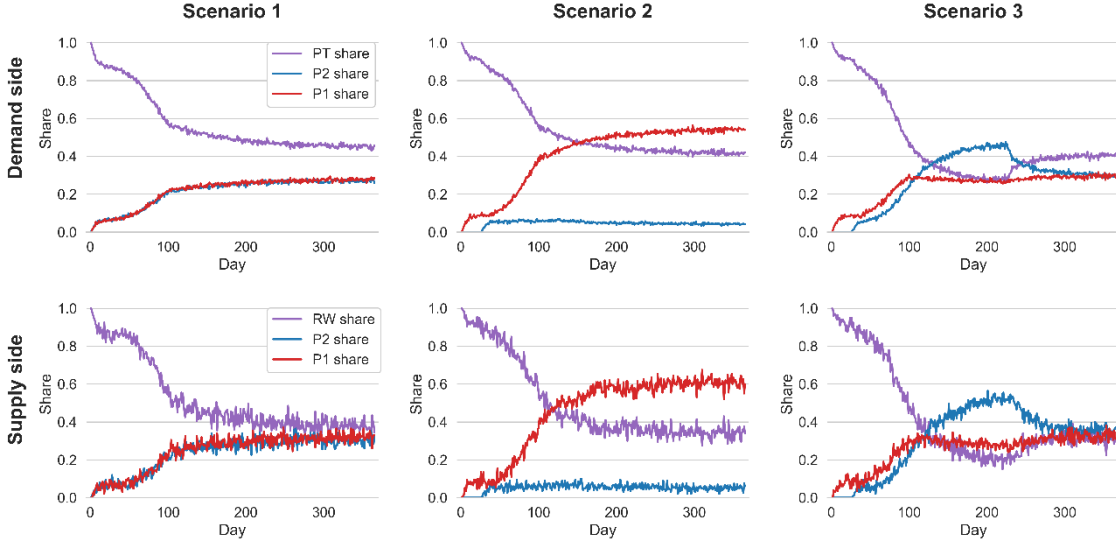


Figure 3: Three competition scenarios and resulting market shares.

Two platforms applying the same strategy, at the same time, end up with the same market shares (Scenario 1). However, late market entry with the same strategy results in failure (Scenario 2). This happens because agents have already started to use the early platform once P2 enters the market. Thus, P1 starts the cross-side network effects earlier, and as the utility and market share differences increase between two platforms, it becomes impossible for late platform to succeed. P2 requires an alternative strategy to compensate the late market penetration which means more subsidies on demand or/and supply side. In the Scenario 3, the late platform enters at the day 25 offering 80% discount for the next 200 days (instead of 100 days of 40% discount) on the demand side to overtake the early platform. In contrast to the second scenario, P2 reaches higher market share than P1 in the last scenario, at first (until day 225). Yet, as P2 terminates the discount its market share decreases and stabilizes (where it is supportable by network effects). Indeed, the market share bubble induced by disloyal agents, relying on discount, bursts with discount termination. While Scenarios 1 and 3 depict the market sharing regime, the second scenario resembles the winner-takes-all in the ride-sourcing market.

4. CONCLUSIONS

In this research, we shed light on the dynamics of ride-sourcing market in which two platforms compete with each other and the public transportation (for travellers) and reservation wage (for drivers). We use our day-to-day coevolutionary model featured by the S-shaped learning curves to capture the rise and fall of the system in MaaSSim. Our results underpin how correlation inside the ride-sourcing choice nest significantly affects the total market share of ride-sourcing platforms, which calls for further empirical studies. Assuming a moderate correlation rate, we

analysed the competition in three different scenarios. We found that platforms with late entry to the market require more subsidies to trigger the cross-side network effects. However, subsidies can induce market share bubble for the platforms which can easily burst with the termination of subsidies, i.e., the platform stabilizes later, on the market share supportable by the network effects. We conclude that ride-sourcing market reaches an equilibria in long term, and both the “winner-takes-all” and the “market sharing” are the possible competition outcomes. Nevertheless, the market remains sensitive and late-entry alternatives may still reach significant market shares, mastering the network effects.

ACKNOWLEDGEMENTS

This research is funded by National Science Centre in Poland program OPUS 19 (Grant Number 2020/37/B/HS4/01847) and by Jagiellonian University under the program Excellence Initiative: Research University (IDUB).

REFERENCES

- Ahmadinejad A., Nazerzadeh H., Saberi A., Skochdopole N., Sweeney, K. 2019. Competition in Ride-Hailing Markets. *SSRN Scholarly Paper*, No. 3461119.
- Belleflamme P., Peitz M. 2018. Chapter 11: Platforms and network effects. *Game Theory and 8 Industrial Organization*, Cheltenham.
- de Ruijter A., Cats O., Kucharski R., van Lint H. 2022. Evolution of labour supply in ridesourcing. *Transportmetrica B: Transport Dynamics*, Vol. 10, No. 1, p. 599–626.
- Ghasemi F., Kucharski R. 2022. Modelling the Rise and Fall of Two-Sided Mobility Markets with Microsimulation (*arXiv:2208.02496*). *arXiv*. <https://doi.org/10.48550/arXiv.2208.02496>
- Kucharski R., Cats O. 2022. Simulating Two-Sided Mobility Platforms with MaaSSim. *PLOS ONE*, Vol. 17, No. 6, p. e0269682.
- Murre J. 2014. S-Shaped Learning Curves. *Psychonomic Bulletin & Review*, Vol. 21, No. 2, p. 344–356.
- Rochet J.C., Tirole J. 2006. Two-Sided Markets: A Progress Report. *The RAND Journal of Economics*, Vol. 37, No. 3, 2006, pp. 645–667.
- Shoman M., Moreno A. T. 2021. Exploring Preferences for Transportation Modes in the City of Munich after the Recent Incorporation of Ride-Hailing Companies. *Transportation Research Record*, Vol. 2675, No. 5, p. 329–338.
- Siddiq A., Taylor T. A. 2022. Ride-Hailing Platforms: Competition and Autonomous Vehicles. *Manufacturing & Service Operations Management*, Vol. 24, No. 3, p. 1511–1528.
- Zhang K., Nie Y. 2021. Inter-platform competition in a regulated ride-hail market with pooling. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 151, p. 102327.

Cooperation between Ride-Hailing and Public Transportation with Tradable Credit Schemes

Louis Balzer*¹ and Ludovic Leclercq²

¹PhD student, Univ Gustave Eiffel, ENTPE, LICIT-ECO7, France

²Professor, Univ Gustave Eiffel, ENTPE, LICIT-ECO7, France

SHORT SUMMARY

Ride-Hailing (RH) companies have expanded significantly in urban areas in the past decade. However, they may compete with Public Transportation (PT) instead of completing them. This work proposes to use Tradable Credit Scheme (TCS), a quantity-based policy, to encourage RH drivers to operate in parts of the network not well served by PT. Credits are given to the RH drivers. Operating in some regions requires credits. RH drivers can trade credits between themselves. We use a trip-based Macroscopic Fundamental Diagram (MFD) to compute the dynamic and heterogeneous RH trips. Customers choose between different PT and RH alternatives. The RH drivers' decision to operate in a region is a balance between the potential revenue and the credit charge of this region. We evaluate the equilibrium of different TCS on a test case. TCS fosters multimodal trips, which combine PT and RH to complete the trips.

Keywords: Macroscopic Fundamental Diagram, Ride-hailing, Tradable Credit Scheme, Traffic flow theory, Transport economics and policy

1 INTRODUCTION

Ride-Hailing (RH) companies introduced new mobility alternatives in many cities (OECD, 2018). However, RH may negatively affect the transportation network, as they contribute to congestion (Erhardt et al., 2019) and compete with Public Transportation (PT) (Cats et al., 2022). Nevertheless, RH has the potential to complete the PT network.

When studying how RH services operate and compete with other modes, it is important to keep track of the transportation system dynamics, as congestion significantly impacts travel times and service quality. We must also consider the service's full spatial extent and reproduce the vehicles' trip patterns over the day. The Macroscopic Fundamental Diagram (MFD) has been proposed by Daganzo (2007). It defines the average mean speed as a function of the number of vehicles driving in the road network. The MFD concept permits the design of large-scale and low-computation dynamic simulations. In particular, the trip-based formulation (Mariotte et al., 2017; Lamotte & Geroliminis, 2018; Jin, 2020) keeps track of all users' and vehicles' moves, which makes it suitable for reproducing RH driver matching and pick-up.

Several recent contributions regarding RH services are founded on the MFD concept. Nourinejad & Ramezani (2020) study the equilibrium between offer, demand, and service pricing. They use the MFD framework to tune a model predictive controller. Beojone & Geroliminis (2021) encourage passengers to share their rides and park unmatched vehicles to reduce the impact of RH vehicles on congestion.

RH drivers operate in high-demand areas, which may already have a good PT network, as they want to increase their profits. However, the regulator wants them to operate in regions with low PT coverage to achieve system optimum. Traditional taxi license schemes aim to regulate and redistribute the number of operating taxis in a given zone in the long run. More flexible quantity-based tools like Tradable Credit Schemes (TCS) can also be applied. The regulator limits access to a shared resource (access to part of the network) by distributing a given total amount of credits. Their price is not fixed but results from the trading between users, introducing more flexibility. Some early concepts on quantity-based instruments to mitigate congestion have been proposed by Verhoef et al. (1997). Yang & Wang (2011) formulated a TCS to reduce the total travel time in an urban transportation system dealing with route choice. Balzer & Leclercq (2022) implemented TCS in a multimodal context to reduce the share of car drivers and decrease the total travel time and carbon emissions.

For now, TCS has only been proposed for demand management. Here, we want to extend the concept to the offer side. The goal is to encourage RH drivers to shift from the city center to the suburbs, where they can propose efficient first-/last-mile alternatives and complete the PT offer.

2 METHODOLOGY

The urban area is divided into N_R different regions. The regions are indexed by increasing order from the center to the outskirts. Each traveler chooses its travel mode $m \in M$ according to the associated costs. The alternatives are riding the PT or using the RH service. The road network is shared with users driving their personal cars. We assume the number of RH drivers N_D is constant. The regulator enforces a TCS to prevent unnecessary competition between RH vehicles and the PT in the city center where the transit offer is satisfying. Its strategy is nudging RH drivers to serve the travel demand in the outskirts to (i) promote PT in the city center (ii) promote multimodal trips where RH drivers permit travelers from the outskirts to ride an RH vehicle to a transit hub at the border of the city center and then use the PT for the remaining trip. Figure 1 presents a schematic representation of the different travel options for a traveler going from the suburbs to the city center.

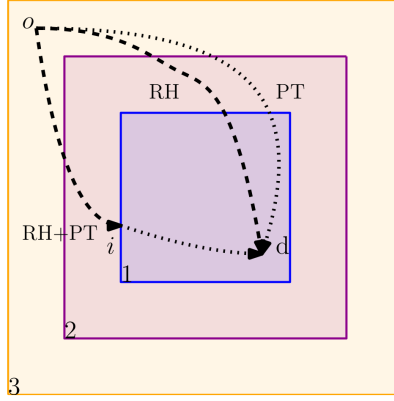


Figure 1: A trip between an origin o in region 3 (suburbs) and a destination d in region 1 (city center) has three alternatives: RH, PT, or RH till the border i and then PT.

We set a framework based on the trip-based MFD to study the effect of the TCS. It considers the congestion dynamics and the heterogeneity of the trips: each customer has its own departure time t_d , origin o , and destination d . RH drivers move only to pick up or drive a customer to its destination. The rest of the time, they park on the street and wait to pick up another customer. Those idle drivers do not contribute to the congestion. The PT mean speed V_{PT,r_o,r_d} depends on the regions of origin and destination. It is faster in the city center (subways) and slower in the suburbs (buses). We use the trip-based MFD to compute RH trips. The mean car speed V depends on the car accumulation n , i.e., the number of cars driving. It includes RH vehicles and also private cars. The actual arrival time t_a for a RH trip of distance l_{RH} is computed using the following relationship:

$$l_{pu} + l_{RH} = \int_{t=t_d}^{t_a} V(n(t))dt. \quad (1)$$

It accounts for the time the customer waits to be picked up, i.e., for the RH vehicle to travel the distance l_{pu} between the current position of the RH vehicle and the customer's origin.

At their departure times t_d , the travelers choose their travel alternatives according to the different *perceived* travel costs (RH or PT). The user travel costs from origin o to destination d are defined by the travel time and the service price:

$$C_{o,d,PT}(t_d) = \alpha L_{PT,o,d}/V_{PT,r_o,r_d} + f_{PT}; \quad (2)$$

$$C_{o,d,RH}(t_d) = \alpha (L_{pu,o,d}(t_d) + L_{RH,o,d})/V(t_d) + f_{RH}L_{RH,o,d}; \quad (3)$$

$$C_{o,d,RH+PT}(t_d) = \alpha (L_{pu,o,i}(t_d) + L_{RH,o,i})/V(t_d) + f_{RH}L_{o,i} + \alpha L_{PT,i,d}/V_{PT,r_i,r_d} + f_{PT}; \quad (4)$$

$$C_{o,d,PT+RH}(t_d) = \alpha L_{PT,o,i}/V_{PT,r_o,r_i} + f_{PT} + \alpha (L_{pu,i,d}(t_d) + L_{RH,i,d})/V(t_d) + f_{RH}L_{RH,i,d}. \quad (5)$$

α is the value of time. $L_{PT,o,d}$ is the PT trip length from o to d , and $L_{RH,o,d}$ the RH trip length. f_{PT} is the price of a unitary ticket. We assume the ticket price is independent of the trip. The

RH travel cost consists of the pick-up time, the travel time, and the RH charge. The RH charge is the distance-based fee f_{RH} multiplied by the trip length. The pick-up distance $L_{pu,o,d}$ depends on the current position, availability, and licenses of the RH drivers. For the RH+PT alternative, the travel cost is the sum of the RH travel cost until the border i of the destination region and then the PT travel cost from this border to the destination. The same applies to PT+RH in reverse: the traveler rides the PT and then takes an RH vehicle. Travelers starting at t_d from o to d choose the travel mode m with the probability $\psi_{o,d,m}(t_d)$, depending on the travel costs following the logit rule:

$$\psi_{o,d,m}(t_d) = \frac{e^{-\theta C_{o,d,m}(t_d)}}{\sum_{m' \in M} e^{-\theta C_{o,d,m'}(t_d)}}, \quad \forall m \in M = \{PT, RH, PT + RH, RH + PT\}. \quad (6)$$

Each driver gets κ credits for free from the regulator per day with the TCS. The drivers need to spend τ_r credits to buy a license to operate (i.e., pick-up or drop-off passengers) in the regions with an index higher or equal to r for a day. Since the regions are defined for TCS purposes, we assume $\tau_r < \tau_{r-1}$, $\forall r \in [2, N_R]$. Drivers can trade their credits on a specific market. The law of the offer and demand determines the credit price p . The regulator does not fix it a priori. It is budget-neutral, as all trades occur only between RH drivers.

We note x_r the number of drivers with a license for region r . They can operate in regions $r' \geq r$. For an RH trip from an origin in region 2 to a destination in region 1, the driver needs a license with the smallest index, i.e., 1. The RH drivers are ordered in the list according to their willingness to acquire licenses. It means the first x_1 drivers will acquire license 1, the next x_2 license 2, and the last x_{N_R} will only operate in region N_R .

We focus on the within-day process. The drivers' assignment x , i.e., the choice of operating regions, balances two markets: the RH market, where travelers buy RH services, and the credit market, where drivers trade credits. Figure 2 summarizes the different interactions. Travelers' mode choice

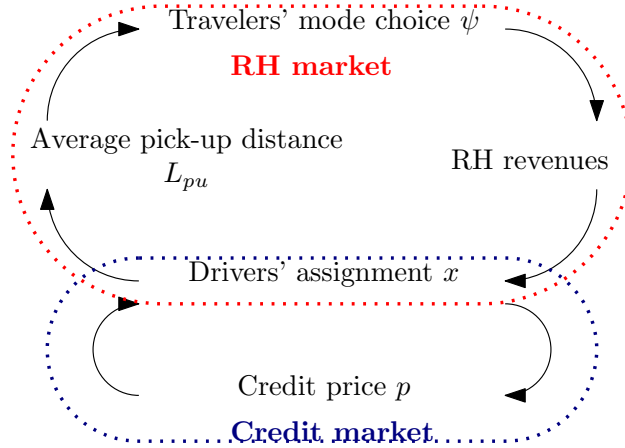


Figure 2: Interactions between drivers, travelers, and credit market.

impacts RH revenue for drivers, which, with the credit price, will change drivers' assignments. The average pick-up distance decreases with the number of drivers available for the trip. The pick-up distance affects the RH perceived costs, thus modifying mode choices.

The equilibrium of drivers' assignment x and credit price p are linked. The RH revenues from the trip requiring access to region r (but not $r - 1$) is

$$R_{RH,r} = \sum_{\text{RH trips with } \min(r_o, r_d)=r} f_{RH} L_{RH,o,d}. \quad (7)$$

It accounts for the RH parts combined trips, where o and d refer to the RH part.

To calculate the equilibrium assignment for drivers, we first define the marginal gain of adding access to region r for a driver as the difference between the RH market per driver and the price to access this market. It is the volume of fees paid by travelers using RH for a trip requiring access to region r but not $r - 1$, divided by the number of drivers with access to region r minus the increase in credit charge times the credit price:

$$G_r = \frac{R_{RH,r}}{\sum_{s \leq r} x_s} - p(\tau_r - \tau_{r+1}), \quad \forall r \in [1, N_R - 1], \quad (8)$$

A positive marginal gain for region r means switching from license $r - 1$ to r will increase the driver's profit. On the opposite, negative marginal gain means accessing the new market is smaller than the extra credit cost. The equilibrium for RH drivers is reached when they have no incentive to change their region access, i.e., when the marginal gains of the regions' access are zero:

$$G_r = 0, \forall r \in [1, N_R - 1]. \quad (9)$$

We do not look at G_{N_R} because every driver has access to region N_R , the further away from the center. We assume operating solely in region N_R does not require credits and is thus always possible. Equation (9) can be expressed as a fixed-point problem $(x, p) = \Gamma(x, p)$ with

$$\Gamma : (x, p) \mapsto \begin{pmatrix} \frac{R_{RH,1}(x)}{p(\tau_1 - \tau_2)} \\ \vdots \\ \frac{R_{RH,N_R-1}(x)}{p(\tau_{N_R-1} - \tau_{N_R})} \\ x_{N_R} \\ p \end{pmatrix}, \quad (10)$$

under the following constraints:

$$x_r \geq 0, \forall r \in [1, N_R]; \quad (11)$$

$$\sum_{r=1}^{N_R} x_r = N_D; \quad (12)$$

$$\sum_{r=1}^{N_R} x_r (\tau_r - \kappa) \leq 0; \quad (13)$$

$$p \sum_{r=1}^{N_R} x_r (\tau_r - \kappa) = 0; \quad (14)$$

$$p \geq 0. \quad (15)$$

The first is that the number of drivers per license is non-negative. The second is the conservation of the number of drivers. The third is the credit cap: the drivers cannot spend more credits than the distributed amount. The fourth is the market clearing condition: all credits are used, or their price is zero. The last one is that the credit price is non-negative. The previous three constraints are specific to the TCS.

Solving the fixed-point problem (Equation (10)) under the different constraints (11, 12, 13, 14, and 15) give the drivers' assignment and the credit price at equilibrium.

The challenge lies in the complex relationship between drivers' licenses and customers' mode choices through the trip-based MFD. We use Bayesian optimization to minimize the cost function J , the sum of the magnitudes of the fixed points errors:

$$J = \sum_{r=1}^{N_R-1} \left| x_r - \frac{R_{RH,r}(x)}{p(\tau_r - \tau_{r+1})} \right|. \quad (16)$$

We use the constraints to reduce the size of the minimization problem. We assume the price is non-zero. Otherwise, the TCS is non-effective, and the state of the system is the same as without TCS, where all drivers can operate in all regions, i.e., $x_1 = N_D$ and $x_r = 0, \forall r \in [2, N_R]$. Then the equality holds for the credit cap (Equation (13)). We combine it with the driver conservation (Equation (12)) to remove two variables. We choose to replace x_{N_R-1} and x_{N_R} with

$$x_{N_R-1} = \frac{N_D(\kappa - \tau_{N_R}) - \sum_{k=1}^{N_R-2} (\tau_k - \tau_{N_R}) x_k}{\tau_{N_R-1} - \tau_{N_R}}; \quad (17)$$

$$x_{N_R} = N_D - \sum_{r=1}^{N_R-1} x_r.$$

3 RESULTS

To illustrate our preliminary work on the proposed methodology, we set an example with $N_R = 3$ regions, $N_D = 150$ drivers, and 1000 travelers. Region 1 is the city center with high demand and

good PT coverage.

Thanks to the problem size reduction (Equation (17)), there are only two unknowns: x_1 and p , which makes the problem suitable for Bayesian optimization (BO). The agent-based model is run 20 times to remove the stochastic bias for the evaluation of the results. We use the open-source Python package *BayesianOptimization* (Nogueira, 2014) to compute the drivers' assignment and credit price at equilibrium.

Two different TCS are compared against the status quo (i.e., no RH regulation). Every driver gets $\kappa = 1$ credit. Region 3 is always free of charge. In the first TCS (TCS1), region 2 requires 1 credit, and region 1 requires 2 credits. The second scenario, TCS2, is more constraining: operating in region 1 requires 2 credits, and region 1 requires 4 credits. The corresponding equilibriums are compared in Table 1.

Table 1: Impacts of two TCS on all trips; trips within the region 1 (city center); trips from or to the region 1 (exclusive or); and trips outside the region 1.

	Status quo	TCS1	TCS2
Credit charge τ	[0,0,0]	[2,1,0]	[4,2,0]
Credit price p (EUR)	0	23	13
Drivers' assignment x	[150,0,0]	[58,34,58]	[26,23,101]
All trips			
RH revenue (EUR)	1454	1549	1266
Average travel time (min)	18.8	18.4	19.1
Trips within the city center			
PT (%)	74	78	89
RH (%)	26	22	11
Average travel time (min)	6.7	6.9	7.2
Trips from or (but not and) to the city center			
PT (%)	65	65	73
RH (%)	16	12	6
PT+RH (%)	10	12	12
RH+PT (%)	8	10	9
Average travel time (min)	17.6	17.7	18.3
Trips outside the city center			
PT (%)	59	51	55
RH (%)	35	37	30
PT+RH (%)	4	7	9
RH+PT (%)	2	5	7
Average travel time (min)	24.4	23.1	24.1

Without TCS, all drivers choose the license to operate in region 1 and above, as it allows access to the whole RH market without extra costs. With TCS1 and TCS2, the credit cap forces some drivers to not operate in regions 1 and 2. The credit price is lower with TCS2 than TCS1, but the TCS-related cost to operate in region 1 (and above) $p\tau_1$ is higher with TCS2 (52 EUR vs. 46 EUR with TCS1). The TCS decreases RH shares and increases PT ridership in the city center (region 1). For trips from or to the city center, the TCS decreases RH-only trips at the profit of combinations with PT. For trips outside of the city center, where the PT coverage is low, both the number of PT- and RH-only trips decreases, and the combinations of RH and PT increase. The average travel times are better for trips outside the city center by about 5%. The total RH revenue increases by 7% with TCS1. Forcing some drivers not to operate in region 1 reduces the pick-up distance in the suburbs and makes some customers choose RH for relatively long trips. It, however, decreases with TCS2, as it greatly restricts operations in the city center (region 1), which is a lucrative market because of the relatively high demand.

4 CONCLUSIONS

This work proposes a tradable credit scheme to foster cooperation between ride-hailing services and public transportation. The TCS nudges RH drivers to serve the suburbs and complete the PT offer to foster multimodal trips: RH in an area with sparse PT infrastructure (suburbs) and PT in a dense area (city center). The main effect for the customer is that increasing the number of RH drivers in the suburbs will decrease the average pick-up distance. We compute the drivers' operating regions and the credit price by solving the equilibrium with BO. We evaluate and compute the equilibrium using the trip-based MFD to calculate the trips. The results in a simplified scenario show the TCS nudge part of the travelers to combine RH and PT for their trips. It makes the PT more competitive for travelers outside the city center.

Further work will apply this methodology for a case study based on a real city.

ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement no. 953783 (DIT4TraM).

REFERENCES

- Balzer, L., & Leclercq, L. (2022, 6). Modal equilibrium of a tradable credit scheme with a trip-based MFD and logit-based decision-making. *Transportation Research Part C: Emerging Technologies*, 139, 103642. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0968090X22000857> doi: 10.1016/J.TRC.2022.103642
- Beojone, C. V., & Geroliminis, N. (2021, 3). On the inefficiency of ride-sourcing services towards urban congestion. *Transportation Research Part C: Emerging Technologies*, 124, 102890. doi: 10.1016/J.TRC.2020.102890
- Cats, O., Kucharski, R., Danda, S. R., & Yap, M. (2022, 1). Beyond the dichotomy: How ride-hailing competes with and complements public transport. *PLOS ONE*, 17(1), e0262496. Retrieved from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0262496> doi: 10.1371/JOURNAL.PONE.0262496
- Daganzo, C. F. (2007, 1). Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41(1), 49–62. doi: 10.1016/j.trb.2006.03.001
- Erhardt, G. D., Roy, S., Cooper, D., Sana, B., Chen, M., & Castiglione, J. (2019). Do transportation network companies decrease or increase congestion? *Science Advances*, 5(5). Retrieved from <https://www.science.org/doi/10.1126/sciadv.aau2670> doi: 10.1126/SCIADV.AAU2670/SUPPL{_}FILE/AAU2670{_}SM.PDF
- Jin, W.-L. (2020). Generalized bathtub model of network trip flows. *Transportation Research Part B*, 136, 138–157. doi: <https://doi.org/10.1016/j.trb.2020.04.002>
- Lamotte, R., & Geroliminis, N. (2018, 11). The morning commute in urban areas with heterogeneous trip lengths. *Transportation Research Part B: Methodological*, 117, 794–810. doi: 10.1016/j.trb.2017.08.023
- Mariotte, G., Leclercq, L., & Laval, J. A. (2017, 7). Macroscopic urban dynamics: Analytical and numerical comparisons of existing models. *Transportation Research Part B: Methodological*, 101, 245–267. doi: 10.1016/j.trb.2017.04.002
- Nogueira, F. (2014). *Bayesian Optimization: Open source constrained global optimization tool for Python*. Retrieved from <https://github.com/fmfn/BayesianOptimization>
- Nourinejad, M., & Ramezani, M. (2020, 2). Ride-Sourcing modeling and pricing in non-equilibrium two-sided markets. *Transportation Research Part B: Methodological*, 132, 340–357. doi: 10.1016/J.TRB.2019.05.019

- OECD. (2018). *Taxi, ride-sourcing and ride-sharing services - Background Note by the Secretariat* (Tech. Rep.). Retrieved from <http://www.oecd.org/daf/competition/taxis-and-ride-sharing-services.htm>
- Verhoef, E., Nijkamp, P., & Rietveld, P. (1997, 11). Tradeable permits: their potential in the regulation of road transport externalities. *Environment and Planning B: Planning and Design*, 24(4), 527–548. Retrieved from <http://epb.sagepub.com/lookup/doi/10.1068/b240527> doi: 10.1068/b240527
- Yang, H., & Wang, X. (2011). Managing network mobility with tradable credits. *Transportation Research Part B: Methodological*, 45(3), 580–594. doi: 10.1016/j.trb.2010.10.002

An auctioning process for large-scale ride-hailing vehicles repositioning

M. Seppecher*¹, L. Leclercq¹

¹ Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT-ECO7 UMR T9401, F-69675, Lyon
(France)

SHORT SUMMARY

On-demand mobility services are transforming urban mobility. They can provide individual and collective benefits when managed optimally, and their successful integration within the existing urban transport system can enhance its performance. In contrast, inadequate fleet management can inflict high pick-up waiting times and passenger drop-out rates. One of the main challenges for on-demand mobility service operators is to proactively rebalance their fleets to ensure that the spatial distribution of supply matches the demand. This paper proposes to address this problem with a distributed auctioning approach. We design an architecture that relies on local controllers interacting with idle vehicles, encouraging them to relocate to their service area. We conduct simulations on the city of Lyon in France, which reveal a substantial increase in the number of passengers served compared to a scenario without rebalancing.

Keywords: auctioning, fleet management, multi-agent modeling, on-demand services, ride-hailing, simulation.

1. INTRODUCTION

Over the last decades, novel mobility services have appeared in cities, such as ride-sourcing (including e-hailing and ride-splitting) and vehicle sharing. In particular, ride-sourcing companies have multiplied, competing with traditional taxi companies and providing travelers with a vast range of services. This new offer can meet an increasingly dynamic and non-regular mobility demand, unsatisfied by public transportation or personal car constraints. On one side, ride-sourcing services offer more flexible services than public transit, on-spot and on-demand pick-up, and no connections. On the other side, they can be less costly than private car ownership and provide satisfying solutions to parking issues. At a collective scale, the services can contribute to limiting car ownership and its externalities, such as land occupancy, soil sealing, or congestion.

However, efficiently managing this type of service requires handling several operational issues. Fleet rebalancing is one of them. It consists of reorganizing a vehicle fleet in space and time by dispatching idle vehicles towards high-demand areas to limit vehicle accumulation in attractive zones and ensure continuous and prompt service to passengers.

Rebalancing must preferably be proactive, *i.e.*, anticipate the future demand and reorganize the fleet accordingly. Numerous literature studies have looked at this management issue. Yet, most offer centralized management methods, raising questions regarding their robustness to scaling or communication failures (Alonso-Mora et al., 2017; Miao et al., 2017, 2015; Ramezani and Nourinejad, 2018). In this respect, distributed approaches are interesting alternatives. Recent works

have looked into fleet rebalancing through the lens of passengers and drivers-intended incentives, with pricing and information-sharing strategies or coverage control (Zhu et al., 2022). In this work, we explore this subject through the angle of auctioning. While auctioning approaches have been applied for developing (reactive) matching strategies (Manjunath et al., 2021; Nourinejad and Roorda, 2016; Wu et al., 2008), this is, to our knowledge the first attempt to extend its application to fleet large scale repositioning.

2. METHODOLOGY

The method we develop relies on a mesh of controllers that divide the urban network into an equal number of service areas. These controllers, which can be associated with physical infrastructures such as taxi stations and deposits, are considered at the service of a public authority (*e.g.*, local authority or transport agency). Their goal is to ensure that ride requests occurring within the boundaries of their service area are served with the minimum waiting time. For this purpose, they aim to attract idle vehicles within their perimeter by negotiating with them at regular intervals (*e.g.*, every 10 minutes) within a **two-sided matching market**.

With this frequency, the controllers are first in charge of forecasting the future demand (*i.e.*, the number of requests). The specific topic of demand forecasting is out of the scope of this paper, and we will assume that historical data allow modeling the future number of requests as a random variable $X = N(\mu_i^T, (\sigma_i^T)^2)$. This assumption is supported by recent research in demand prediction (Khalesian et al., 2022). To attract the required number of vehicles, local controllers publish within the matching market as many **relocating offers** as expected ride requests. These offers will allow vehicles to which they are assigned to relocate within the corresponding service area. Each relocation offer is characterized by:

1. The likelihood of the expected ride request. We define the likelihood p_k of the k^{th} expected ride request as the probability that at least k ride request occur during t . Therefore, we have:

$$\forall k \in \mathbb{N}, p_k = S_X(k) = P(X \geq k) \quad (1)$$

with S_X the survival function of X .

2. The expected revenue \hat{g}_i for picking up a passenger in service area i , which can be estimated based on historical data.

Then, the matching of vehicles with a relocation offer follows a **distributed Gale-Shapley algorithm** (Brito and Meseguer, 2006, 2005), often used to solve matching problems (marriage, student-college or resident-hospital matching). First, the features of relocating offers allow drivers to estimate their utility in applying to one or another relocation option. This utility is estimated as the expected net revenue, computed as the difference between expected incomes (expected revenue weighted by request likelihood) and rebalancing costs:

$$U_v(i, k) = p_k \hat{g}_i - c_v(i) \quad (2)$$

Drivers bid on the most useful relocating option and share their expected arrival time within the region with the local controller. Then, local controllers rank the received offer according to their utility. For each relocation offer, the controller accepts its preferred application and rejects the others. Rejected vehicles update their preference list and apply to their second most-preferred option. If a controller previously now receives a better application, it can reject the previously-matched vehicle and accept the new one. The rejected vehicle updates its preference list and applies to another relocation offer. This process goes on until all cars run out of interesting relocation

offers. In the end, vehicles matched relocate to their destination region, and unassigned vehicles remain idle at their current position. Although iterative, this process can be close to instantaneous, as drivers actually do not interfere in the process. We illustrate this communication protocol in Figure 1.



Figure 1: Communication protocol supporting the fleet rebalancing

Note that controllers can use several methods to evaluate the utility of the application of a vehicle. In the present paper, we use the following approach. A fictive occurrence time within the rebalancing period is assigned to each expected ride request. Then, the utility of a vehicle application is determined according to the delay the travel time the vehicle needs to join the service area would inflict on this expected passenger, given this fictive occurrence time. The utility function is triangular, maximal when the car arrives right on time, and decreases faster when the vehicle comes later than when it arrives in advance, as illustrated in Figure 2.

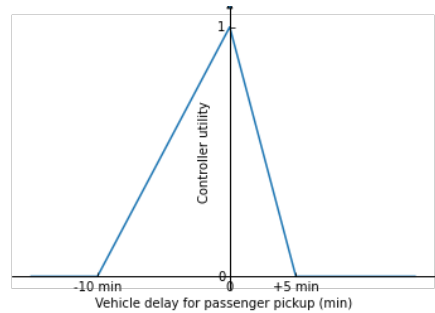


Figure 2: Controllers' utility

3. RESULTS AND DISCUSSION

Case study

We choose the city of Lyon, France, as a case study. The network we model covers 121 km² and includes both the city of Lyon and the city of Villeurbanne, located within a circular ring road. To conduct rapid simulations, we model the traffic on a simplified network of the city. The network only includes the primary and secondary urban roads and highways, as illustrated by Figure 2.a). The supply calibration and the demand scenarios used here have been calculated within the ERC Magnum Project (Mariotte et al., 2020). 15% of the inner flows are assigned to ride-hailing, while the remaining users are assumed to take their personal cars. The city is partitioned into 50 service areas, as illustrated in Figure 2.b). The simulations are conducted on the MnMS multi-agent simulation platform developed at Univ. Gustave Eiffel. This paper presents the results of simulations performed with a 4000-vehicle-large vehicle fleet and 10-minute-long passenger waiting time tolerance.

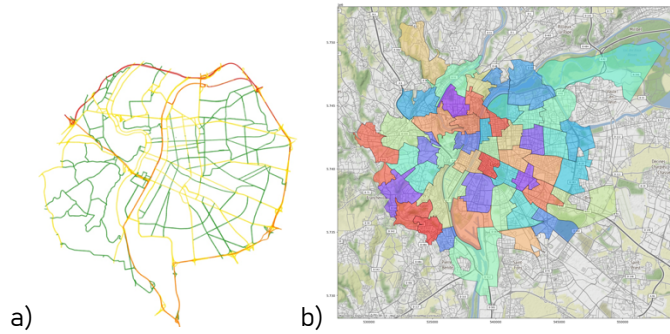


Figure 2. Simulation network. a) Road network. b) Service areas.

Results

We compare our strategy to a no-rebalancing scenario and evaluate performance based on several KPIs regarding service, users, drivers, and traffic. First, our analyses show that implementing our strategy over the city of Lyon allows increasing the number of passengers served by 9.88% (+1975 passengers) compared to the base scenario. This service increase is especially significant between 8:00 a.m. and 9:00 a.m., during the peak demand hour, as illustrated in Figure 3. Figure 4 shows the level of service improvement in space. Applying our rebalancing strategy especially allows for increasing the service in the western suburban and less connected areas of the city (+32% of demand served in some areas) while being slightly detrimental to the service in the city center and eastern neighborhoods. We observed that this overall service improvement comes with an increase in waiting time before pick-up of 1.39 minutes on average. This increase is explained by the decrease in the number of available vehicles, due, on the one hand, to the rise in the number of passengers served, on the other hand, to rebalancing vehicles being considered unavailable for matching. Exploring variants of this rebalancing strategy that allow rebalancing vehicles to pick up passengers should allow limiting this waiting time increase.

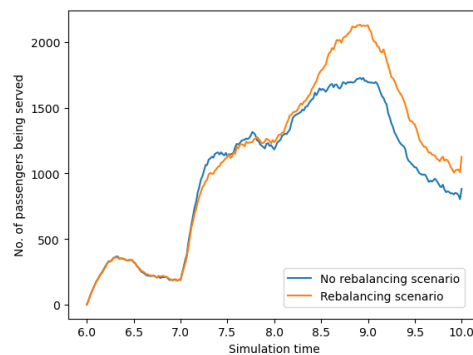


Figure 3: Number of users being served throughout simulation time.

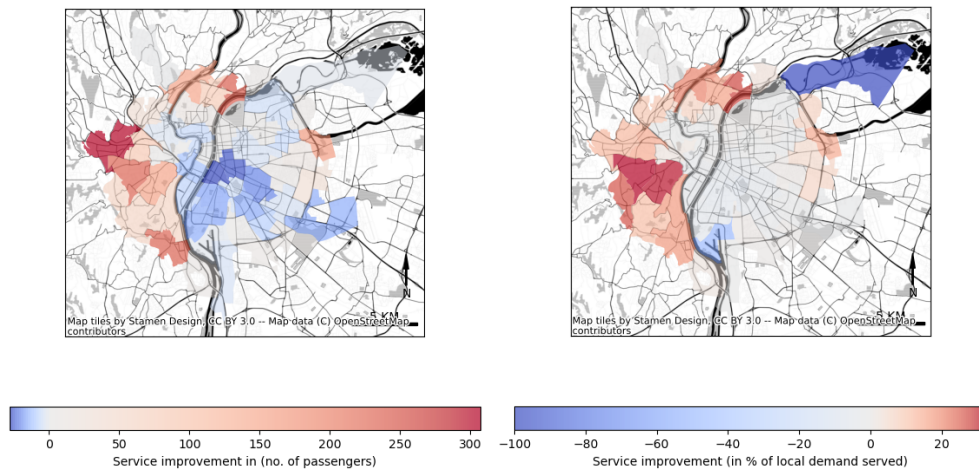


Figure 4: Spatial visualisation of the service improvement thanks to rebalancing.

4. CONCLUSIONS

In this paper, we propose an original fleet rebalancing strategy based on outsourcing rebalancing management to local controllers and implementing a negotiation process between them and the vehicles. Our method significantly impacts the number of passengers served, especially in suburban areas less connected to the city center. Although it also seems to increase the average waiting time of passengers, this increase is limited compared to the increase in the number of passengers served, and more flexible matching strategies will help to mitigate this effect.

As a continuation of this work, future works will focus on conducting advanced sensitivity analyses to fleet size, uncertainty levels, or riding fares. We will also explore different utility functions for local controllers and assess their impact on waiting time, amount of passengers served, or empty mileage.

In the mid-term, we will use this approach to develop local incentive strategies to encourage vehicles to relocate to service areas with lower accessibility or uncertain demand. We will also look at enriching the method to foster cooperation between local controllers rather than competition. Finally, this approach based on controllers external to the service could be relevant in managing the competition between different mobility services.

ACKNOWLEDGEMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement no. 953783 (DIT4TraM).

REFERENCES

Alonso-Mora, J., Wallar, A., Rus, D., 2017. Predictive routing for autonomous mobility-on-demand systems with ride-sharing, in: 2017 IEEE/RSJ International

- Conference on Intelligent Robots and Systems (IROS). Presented at the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3583–3590. <https://doi.org/10.1109/IROS.2017.8206203>
- Brito, I., Meseguer, P., 2006. Distributed Stable Matching Problems with Ties and Incomplete Lists, in: Benhamou, F. (Ed.), Principles and Practice of Constraint Programming - CP 2006, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 675–679. https://doi.org/10.1007/11889205_49
- Brito, I., Meseguer, P., 2005. Distributed Stable Matching Problems. Presented at the LNCS, pp. 152–166. https://doi.org/10.1007/11564751_14
- Khalesian, M., Furno, A., Leclercq, L., 2022. High-Resolution Traffic Demand Prediction, Based on Deep Learning and Hierarchical Reconciliation. *Transportation Research Part C: Emerging Technologies*. <https://doi.org/10.2139/ssrn.4250272>
- Manjunath, A., Raychoudhury, V., Saha, S., Kar, S., Kamath, A., 2021. CARE-Share: A Cooperative and Adaptive Strategy for Distributed Taxi Ride Sharing. *IEEE Transactions on Intelligent Transportation Systems* 1–17. <https://doi.org/10.1109/TITS.2021.3066439>
- Mariotte, G., Leclercq, L., Batista, S.F.A., Krug, J., Paipuri, M., 2020. Calibration and validation of multi-reservoir MFD models: A case study in Lyon. *Transportation Research Part B: Methodological* 136, 62–86. <https://doi.org/10.1016/j.trb.2020.03.006>
- Miao, F., Han, S., Lin, S., Stankovic, J.A., Huang, H., Zhang, D., Munir, S., He, T., Pappas, G.J., 2015. Taxi Dispatch with Real-Time Sensing Data in Metropolitan Areas: A Receding Horizon Control Approach. *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems* 100–109. <https://doi.org/10.1145/2735960.2735961>
- Miao, F., Han, S., Lin, S., Wang, Q., Stankovic, J., Hendawi, A., Zhang, D., He, T., Pappas, G.J., 2017. Data-Driven Robust Taxi Dispatch under Demand Uncertainties. [arXiv:1603.06263 \[cs\]](https://arxiv.org/abs/1603.06263).
- Nourinejad, M., Roorda, M.J., 2016. Agent based model for dynamic ridesharing. *Transportation Research Part C: Emerging Technologies* 64, 117–132. <https://doi.org/10.1016/j.trc.2015.07.016>
- Ramezani, M., Nourinejad, M., 2018. Dynamic modeling and control of taxi services in large-scale urban networks: A macroscopic approach. *Transportation Research Part C: Emerging Technologies, ISTTT22* 94, 203–219. <https://doi.org/10.1016/j.trc.2017.08.011>
- Wu, Y.H., Guan, L.J., Winter, S., 2008. Peer-to-Peer Shared Ride Systems, in: Nittel, S., Labrinidis, A., Stefanidis, A. (Eds.), *GeoSensor Networks: Second International Conference, GSN 2006, Boston, MA, USA, October 1-3, 2006, Revised Selected and Invited Papers, Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, pp. 252–270. https://doi.org/10.1007/978-3-540-79996-2_14
- Zhu, P., Sirmatel, I.I., Trecate, G.F., Geroliminis, N., 2022. Idle-vehicle Rebalancing Coverage Control for Ride-sourcing systems, in: *2022 European Control Conference (ECC)*. Presented at the 2022 European Control Conference (ECC), pp. 1970–1975. <https://doi.org/10.23919/ECC55457.2022.9838069>

Public Transport Crowding Valuation in a Post-Pandemic Era

Menno Yap^a, Howard Wong^b, Oded Cats^a

^a Delft University of Technology, department of Transport and Planning, Delft, the Netherlands

^b University College London, UCL Transport Institute, United Kingdom

ABSTRACT

The main contribution of this study is to derive the crowding valuation of public transport passengers in a post-pandemic era entirely based on observed, actual passenger route choices. We derive passengers' crowding valuation for the London metro network based on a revealed preference discrete choice model using maximum likelihood estimation. We find that after the passenger load on-board the metro reaches the seat capacity, the in-vehicle time valuation increases by 0.422 for each increase in the average number of standing passengers per square metre upon boarding. When comparing this result to a variety of crowding valuation studies conducted before the pandemic in London and elsewhere, we can conclude that public transport passengers value crowding more negatively since the pandemic. Our study results contribute to a better understanding on how on-board crowding in urban public transport is perceived in a European context since the outbreak of the COVID-19 pandemic.

Keywords: COVID-19; Crowding; Public Transport; Revealed Preference; Smart Card Data.

Word count (incl. abstract and references, excl. tables): 2,947

1. INTRODUCTION

In many urban public transport (PT) systems worldwide high passenger volumes result in high crowding levels on-board PT vehicles. Over the last two decades, many studies have been performed to public transport crowding valuation, by inferring the PT in-vehicle time crowding multiplier as a function of the on-board load factor or standing density (average number of standing passengers per m²). Initially, most of these studies relied on stated preference (SP) approaches where respondents were asked in (online) surveys to indicate which route or mode choice alternative they would choose based on hypothetical crowding scenarios (e.g. Batarce et al. 2016, Tirachini et al. 2017, Wardman and Whelan 2011, Li and Hensher 2011). In more recent years there is an increasing number of studies using revealed preference (RP) for this purpose. With the availability of large-scale passenger data from Automated Fare Collection (AFC) systems and/or Automated Passenger Count (APC) systems such as load-weigh systems, passengers' crowding valuation can be derived from empirically observed route and mode choice behaviour. RP based crowding studies have been applied to case studies in Singapore (Tirachini et al. 2016), Hong Kong (Hörcher et al. 2017), the Netherlands (Yap et al. 2020) and Washington, DC (Yap and Cats 2021).

All abovementioned studies estimate the perception of PT crowding based on data before the outbreak of the COVID-19 pandemic. One can expect that passengers are perceiving crowding more negatively since the start of the pandemic as crowded environments generally pose a higher risk of contracting COVID-19. It is thus of utmost importance to understand how PT passengers perceive on-board crowding in this post-pandemic era, as changes in crowding perception might

influence route and mode choice and might hamper a full demand recovery on PT routes being perceived as (over)crowded, imposing in effect new de-facto capacity limits. More recently, a few studies have been performed which assess passengers’ post-pandemic crowding perception based on stated preferences elicited from choice experiments (Bansal et al. 2022, Basnak et al. 2022, Flügel and Hulleberg 2022, Shelat et al. 2022). However, as of yet no studies have been performed which use observed passenger route choices from large-scale AFC and APC systems to re-establish public transport crowding perception in the aftermath of the pandemic based on actual passenger behaviour rather than based on stated behaviour in surveys or choice experiments.

The main contribution of our study is deriving the crowding valuation of public transport passengers in a post-pandemic era entirely based on observed, actual passenger route choices. The results of our revealed preference approach thereby add to the emerging evidence from studies which derive post-pandemic crowding perceptions from SP surveys (see **Table 1**). By relying on large-scale, empirical passenger demand data, we derive crowding valuations based on more than 20,000 observed passenger journeys in the London PT network.

Table 1. Study contribution

PT crowding studies	Stated Preference	Revealed Preference
Pre-pandemic	Li and Hensher (2011) Wardman and Whelan (2011) Batarce et al. (2016) Tirachini et al. (2017)	Tirachini et al. (2016) Hörcher et al. (2017) Yap and Cats (2021)
Post-pandemic	Bansal et al. (2022) Basnak et al. (2022) Flügel and Hulleberg (2022) Shelat et al. (2022)	This study

2. METHODOLOGY

Data input

As input for our study we use passenger demand and occupancy data derived from the urban PT network of the Greater London Area, which is under the authority of Transport for London (TfL). For metro journeys in London each row in the AFC data consists of the location and time of the first station entry and of the last station exit. For buses only the boarding stop, time and bus route are empirically available, whereas the alighting bus stop is inferred. The data format for metro and bus journeys is illustrated in **Table 2**. As metro crowding information is not directly available from AFC data, we rely on APC data obtained from load-weigh data for the lines where the rolling stock is equipped with load-weigh systems. This provides on-board passenger loads for each line segment by train and on average per 15-minute time interval. As London buses are not equipped with APC systems, bus crowding information can only be inferred. Therefore, we only focus on estimating the crowding valuation for metro journeys for which we can rely directly on APC data, while keeping the bus data in the passenger journey dataset.

Table 2. An illustration of the structure of the AFC dataset

Mode	Route	Start Time	Start Stopcode	End Time	End Stopcode
Metro	-	2022-06-15 08:01:12	778	2022-06-15 08:19:53	729
Bus	43	2022-06-17 16:44:05	BP3065	2022-06-17 16:59:22*	BP2336*

* *inferred, not empirically available*

In this study we estimate three different models:

- A pre-pandemic off-peak model based on 3-7 February 2020. We use this as an uncrowded pre-pandemic baseline model.
- A post-pandemic off-peak model based on 13-17 June 2022, used to assess whether base level in-vehicle time and waiting/walking time valuations have changed since the COVID-19 pandemic.
- A post-pandemic peak model based on the same period 13-17 June 2022. This model, focusing on AM and PM journeys, estimates the post-pandemic metro crowding valuation based on load-weight data which is available for this period.

During the selected post-pandemic period 13-17 June 2022 there were no COVID related restrictions in place anymore in London. Additionally, no capacity constraints, social distancing or mandatory face covering were in place when travelling by PT. This implies that June 2022 reflects a more steady-state situation in the post-pandemic era.

Choice set generation

To generate a choice set we apply the following criteria and filtering rules:

- Exclude incomplete and unrealistic journeys.
- Include metro journeys entirely made on lines for which load-weight data is available (Central and Victoria Line).
- Include metro journeys between station pairs with unambiguous routing, to reliably infer the in-vehicle time and waiting time corresponding to the route a passenger took.
- Include journeys made in the off-peak for the two uncrowded models (10-14h or 20-23h), and journeys in the AM peak (6-10h) or PM peak (15-19h) for the crowding model.
- Only include origin-destination pairs with a sufficient number of observations for at least two different observed paths, as we rely entirely on observed passenger route choices to derive crowding perceptions.
- For the crowding model, only include OD pairs with sufficient crowding levels (at least a load factor of 50%) for at least one of the paths.
- Exclude OD pairs where one route option is dominant over the other paths.

The resulting choice set inputs for all three models are summarised in **Table 3**.

Table 3. Choice set description

	Model 1 Pre-pandemic uncrowded model	Model 2 Post-pandemic uncrowded model	Model 3 Post-pandemic crowding model
Observations	50,494	46,400	20,970
Number of OD pairs	407	377	60
Number of paths	820	764	126
Average number of paths per OD pair	2.01	2.03	2.10
Average number of observations per OD pair	124	123	350

Model specification

We adopt a standard utility maximisation framework. To prevent biased estimates due to possible correlations between unobserved components of the different path alternatives $a_{od} \in A_{od}$, we explicitly account for overlap between paths using a path size correction factor as proposed by Ben-Akiva and Bierlaire (1999). Therefore, the total disutility of each path $U(V, r, \varepsilon)$ is composed of the structural, deterministic utility component V , a path size factor r and a random error term ε (**Eq.1**). The probability P_a for choosing each path a can then be calculated using the closed-form function shown in **Eq.2**.

$$U_{a_{od}} = V_{a_{od}} + \beta_{psl} \cdot r_{a_{od}} + \varepsilon_{a_{od}} \quad (1)$$

$$P_{a_{od}} = \frac{\exp(V_{a_{od}} + \beta_{psl} r_{a_{od}})}{\sum_{a_{od} \in A_{od}} \exp(V_{a_{od}} + \beta_{psl} r_{a_{od}})} \quad (2)$$

The structural part of the utility function V is a vector of observable route attributes with their corresponding weights as defined for the uncrowded off-peak models 1 and 2 (**Eq.3**) and for the crowding model (**Eq.4**). We specify mode-specific in-vehicle time coefficients β_{ivt}^b for bus and β_{ivt}^m for metro, so that potential mode-specific differences in in-vehicle time valuation can be captured. A generic waiting/walking out-of-vehicle time coefficient β_{wtt} is specified in the utility function, in such a way that β_{wtt} directly reflects the ratio between waiting/walking time and in-vehicle time valuation. We use the standing density on-board the metro d^m as a crowding metric, which reflects the average number of standing passengers per square metre as derived from load-weight data for each route segment per 15-minute time interval. The standing density equals zero if the passenger load q is smaller than the seat capacity κ – implying that all passengers can have a seat – and increases up to 4 standing passengers per m^2 when all surface available for standing θ has been used. In this study we test three different metrics for capturing the crowding perception associated with the standing density: the average standing density across all links of a passenger journey (**Eq.5**), the standing density at the first link of a passenger journey upon boarding (**Eq.6**), and the maximum standing density (over all links $e_i \in E_i$) at the busiest point of the passenger journey (**Eq.7**). This enables us to assess which formulation of standing density is most important for passenger's crowding valuation. The coefficient β_d^m is specified such that it reflects the in-vehicle time crowding multiplier as function of the standing density.

$$V = asc^b \cdot b + \beta_{ivt}^b \cdot t_{ivt}^b + \beta_{ivt}^b \cdot \beta_{wtt} \cdot t_{wtt}^b + asc^m \cdot m + \beta_{ivt}^m \cdot t_{ivt}^m + \beta_{ivt}^m \cdot \beta_{wtt} \cdot t_{wtt}^m \quad (3)$$

$$V = asc^b \cdot b + \beta_{ivt}^b \cdot t_{ivt}^b + \beta_{ivt}^b \cdot \beta_{wtt} \cdot t_{wtt}^b + asc^m \cdot m + \beta_{ivt}^m \cdot t_{ivt}^m \cdot (1 + (\beta_d^m \cdot d^m)) + \beta_{ivt}^m \cdot \beta_{wtt} \cdot t_{wtt}^m \quad (4)$$

$$d_i^{avg} = \max\left(\frac{\sum_{e_i \in E_i} \frac{q_{e_i} - \kappa_{e_i}}{\theta_{e_i}}}{|E_i|}, 0\right) \quad (5)$$

$$d_i^{first} = \max\left(\frac{q_{e_1} - \kappa_{e_1}}{\theta_{e_1}}, 0\right) \quad (6)$$

$$d_i^{max} = \max\left(\max\left(\frac{q_{e_i} - \kappa_{e_i}}{\theta_{e_i}}, 0\right)\right) \quad \forall e_i \in E_i \quad (7)$$

3. RESULTS AND DISCUSSION

Results

Maximum likelihood estimation is performed to infer the coefficients which best explain the observed passenger route choices for the three different models. The Newton algorithm is used as iterative method to solve this non-linear optimisation problem. From the model estimation summary shown in **Table 4**, it can be seen that the Rho-square-bar of crowding model 3 is 37% higher compared to the Rho-square-bar of uncrowded post-pandemic model 2. Model estimation results are presented in **Table 5**. The signs of all coefficients are plausible and in line with a-priori expectations and findings reported by previous studies. As we don't have access to information on the panel structure of the data, we report the robust t-statistic and robust p-value as sandwich estimator with the aim of preventing an overestimation of the model coefficients. The absolute value of the robust t-value is larger than 1.96 for all estimated coefficients, which confirms that our results are statistically significant.

Table 4. Model estimation summary

	Model 1 Pre-pandemic uncrowded model	Model 2 Post-pandemic uncrowded model	Model 3 Post-pandemic crowding model
Observations	50,494	46,400	20,970
Number of estimated parameters	6	6	6
Initial log-likelihood	-35,339	-33,177	-12,209
Final log-likelihood	-28,182	-25,936	-8,551
Rho-square	0.203	0.218	0.300
Rho-square-bar	0.202	0.218	0.299
Akaike Information Criterion (AIC)	56,377	51,884	17,115
Bayesian Information Criterion (BIC)	56,430	51,937	17,162

Table 5. Model estimation results

	Model 1 Pre-pandemic uncrowded model	Model 2 Post-pandemic uncrowded model	Model 3 Post-pandemic crowding model
Coefficients	Value (robust t-value)	Value (robust t-value)	Value (robust t-value)
asc^b – alternative specific constant bus	-0.677** (-21.0)	-0.792** (-21.2)	-0.635** (-16.1)
asc^m – alternative specific constant metro	+0.677** (+21.0)	+0.792** (+21.2)	+0.635** (+16.1)
β_{ivt}^b – in-vehicle time bus	-0.0653** (-23.0)	-0.0458** (-17.2)	-0.0399** (-14.9)
β_{ivt}^m – in-vehicle time metro	-0.0520** (-13.9)	-0.0388** (-9.07)	-0.0220** (-12.5)
β_{wtt} – ratio wait/walk time : in-vehicle time	+1.94** (+17.6)	+1.93** (+11.5)	+1.93 (fixed) ¹
β_{psl} – path-sized logit factor	-0.438** (-3.54)	-0.757** (-9.18)	-0.573** (-5.95)
β_d^m – standing density metro			+0.422* (+2.21)

*robust t-values in parentheses. * robust $p < 0.05$; ** robust $p < 0.01$*

¹ Fixed for the ratio wait/walk time : in-vehicle time as found in uncrowded post-pandemic model 2 estimated for the same time period

Discussion on uncrowded models

Based on the ratio between the metro and bus in-vehicle time coefficients $\beta_{ivt}^m : \beta_{ivt}^b$ of the uncrowded pre-pandemic **model 1**, we find that on average uncrowded in-vehicle time on-board a metro is perceived 20% less negatively than uncrowded bus in-vehicle time. The same ratio for

post-pandemic uncrowded **model 2** shows that metro in-vehicle time is now on average valued 15% less negatively compared to bus in-vehicle time. Whilst this still confirms a generic passenger preference for metro over bus regarding in-vehicle time, this result suggests that the relative attractiveness of the metro compared to bus has decreased somewhat in terms of in-vehicle time. A possible explanation is that since the COVID-19 outbreak passengers value travelling in enclosed, underground environments such as a metro system more negatively than pre-pandemic, as these might be perceived as areas with higher infection risks. In contrast, bus travel on the surface with frequent door openings at stops and the possibility for passengers to open windows can be perceived as a travel mode providing better ventilation and thus reducing COVID-19 infection risks.

β_{wtt} , the coefficient which reflects the ratio between waiting/walking time and uncrowded in-vehicle time, equals 1.94 for the pre-pandemic model. This implies that on average passengers value one minute of out-of-vehicle (walking or waiting) time as almost two minutes of in-vehicle time. In the post-pandemic model we see that on average out-of-vehicle time is perceived 1.93 times more negatively compared to uncrowded in-vehicle time. As β_{wtt} remains almost unchanged between the pre-pandemic and post-pandemic off-peak models, we can conclude that PT waiting/walking time valuation relative to in-vehicle time did not change since the COVID-19 pandemic.

Discussion on crowding model

For the post-pandemic crowding model the estimated metro crowding coefficient β_d^m is significant at a 95% significance level, with the robust t-statistic of 2.21 being larger than 1.96. The value of this coefficient implies that after the passenger load on-board the metro reaches the seat capacity, the in-vehicle time valuation increases by 0.422 for each increase in the average number of standing passengers per square metre. When we linearly extrapolate the estimated crowding coefficient – as observed crowding levels averaged per 15-minute interval in our choice set did not exceed 3 standing passengers per m^2 – we can estimate that the in-vehicle time multiplier would be equal to 2.69 when a train operates at full capacity (assumed at 4 standing passengers per m^2). The model using the crowding level upon boarding (d^{first}) was the only model resulting in a statistically significant standing density crowding coefficient. This suggests that the PT crowding level upon boarding best captures passengers' crowding valuation. An explanation for this is that the crowding level upon boarding is related to the passenger's seat probability, as this is an important determinant of whether a passenger will be able to have a seat during the entire journey.

In **Figure 1** we compare the in-vehicle time crowding curve as derived from our model to previous studies. Compared to the three pre-pandemic RP studies performed based on large-scale AFC data (Singapore, Hong Kong, Washington DC) we can conclude that our post-pandemic crowding multiplier found for London is substantially higher. The same conclusion is reached when comparing the SP results between a pre-pandemic and post-pandemic study conducted in Santiago de Chile. Specifically for London we refer to two pre-pandemic studies on crowding valuation. The first one is a RP study performed in the 1990s by Transport for London. The resulting crowding multiplier of 2.32 at 4 standing passengers per m^2 is notably higher than other pre-pandemic studies, although this study has been performed several years ago using a different methodology than more recent RP studies. We can derive a more recent average pre-pandemic crowding multiplier for London using the SP based coefficients estimated for seated and standing passengers by Whelan (2009), which results in an average pre-pandemic in-vehicle time multiplier of 1.77. Our equivalent RP based estimated crowding multiplier for London in the post-pandemic era of 2.69 provides strong evidence that PT passengers value metro crowding substantially more negatively

in London since the COVID-19 outbreak compared to both pre-pandemic studies in London, despite their differences in methodology. The crowding valuation found in our study is comparable to the post-pandemic crowding valuation derived from SP research for Santiago de Chile by Basnak (2022), which gives confidence in the magnitude of our estimated crowding coefficient.

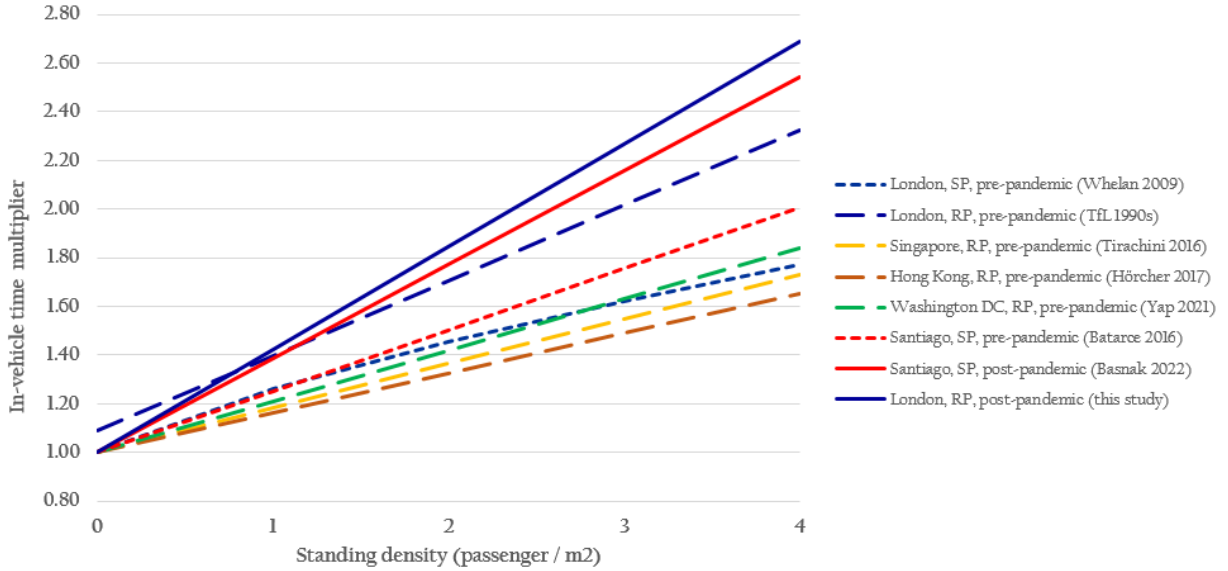


Figure 1. In-vehicle time crowding multiplier as function of standing density

4. CONCLUSIONS

Based on the three estimated discrete choice models we can formulate three main conclusions. First, the average post-pandemic out-of-vehicle time valuation remains unchanged at almost twice the uncrowded in-vehicle time valuation. Second, whilst our study results confirm that there is a generic passenger preference for metro over bus regarding in-vehicle time, we find that the relative attractiveness of metro compared to bus has decreased somewhat post-pandemic in terms of in-vehicle time. This possibly echoes a more negative perception of metro travelling in a more enclosed, underground environment compared to bus travel. Third, our crowding model estimation results show that passengers' average in-vehicle time valuation increases by 0.422 for each increase in the average number of standing passengers per square metre. In contrast, this same value equals 0.22 as average across the six studies to pre-pandemic crowding valuation as reported in **Figure 1**. Compared to the results of these SP and RP studies conducted before the pandemic in London and elsewhere we thus clearly see a steeper slope of the post-pandemic crowding curve as found in our study, based on which we can conclude that PT passengers value crowding more negatively since the COVID-19 pandemic.

REFERENCES

Bansal, P., Kessels, R., Krueger, R., Graham, D.J. (2022). Preferences for using the London Underground during the COVID-19 pandemic. *Transportation Research Part A*, 160, 45-60.

- Basnak, P., Giesen, R., Munoz, J.C. (2022). Estimation of crowding factors for public transport during the COVID-19 pandemic in Santiago, Chile. *Transportation Research Part A*, 159, 140-156.
- Batarce, M., Munoz, J.C., Ortuzar, J. (2016). Valuing crowding in public transport: Implications for cost-benefit analysis. *Transportation Research Part A*, 91, 358-378.
- Ben-Akiva, M., Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. In *Handbook of Transportation Science*, edited by R. W. Hall. Boston, MA: Springer. doi:10.1007/978-1-4615-5203-1_2.
- Flügel, S., Hulleberg, N. (2022). Aversion to in-vehicle crowding before, during and after the COVID-19 pandemic. *Findings*, <https://doi.org/10.32866/001c.37641>.
- Hörcher, D., Graham, D.J., Anderson, R.J. (2017). Crowding cost estimation with large scale smart card and vehicle location data. *Transportation Research Part B*, 95, 105-125.
- Li, Z., Hensher, D. (2011). Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. *Transport Policy*, 18, 880-887.
- Shelat, S., Van de Wiel, T., Molin, E., Van Lint, J.W.C., Cats, O. (2022). Analysing the impact of COVID-19 risk perceptions on route choice behaviour in train networks. *PLoS ONE*, 17: e0264805.
- Tirachini, A., Hurtubia, R., Dekker, T., Daziano, R.A. (2017). Estimation of crowding discomfort in public transport: Results from Santiago de Chile. *Transportation Research Part A*, 103, 311-326.
- Tirachini, A., Sun, L., Erath, A., Chakirov, A. (2016). Valuation of sitting and standing in metro trains using revealed preferences. *Transport Policy*, 47, 94-104.
- Wardman, M., Whelan, G. (2011). Twenty years of rail crowding valuation studies: evidence and lessons from British experience. *Transport Reviews*, 31, 379-398.
- Yap, M.D., Cats, O. (2021). Taking the path less travelled: Valuation of denied boarding in crowded public transport systems. *Transportation Research Part A*, 147, 1-13.

A data-driven dynamic demand hotspots forecasting framework for on-demand meal delivery platforms

Jingyi Cheng*¹ and Shadi Sharif Azadeh²

¹Transport & Planning, Delft University of Technology, the Netherlands

SHORT SUMMARY

Speed and reliability are the keys to high quality on-demand meal delivery service. The rebalancing of couriers locations according to future demand remains an operational challenge in the industry. This study proposes an adaptive framework to identify and predict the near-future demand hotspots, utilizing the semi real-time predictive information as input. This framework provides demand insights to assist meal delivery platforms in making operational forward-looking resource-demand rebalancing decisions in real time, such as fleet management and demand management. To generate fast and accurate demand forecasting, we incorporate time series features and data-driven machine learning methods to create an adaptive forecasting approach. We create a dynamic demand hotspot clustering algorithm which takes predictive and geographic information as input. In the case study, our predictive forecasting model outperforms the time series and deep learning benchmarks in deterministic forecasting. The hotspots clustering performance is improved by using probabilistic predictive input.

Keywords: short-term demand forecasting; on-demand meal delivery; data-driven methods

1 INTRODUCTION

The on-demand meal delivery business has created a global market valued at over 150 billion dollars. However, competition among on-demand meal delivery service platforms (OMDPs), such as DoorDash, Uber, and Grubhub, is intense. Good delivery service quality is essential for maintaining high customer satisfaction, which improves customer loyalty in the long run. Delivery speed and reliability are the key factors in determining service quality. To generate higher profits while providing high-quality service, the platform should operate its limited courier resource efficiently. Various studies have been conducted to enhance the delivery efficiency of OMDPs. Reyes et al. (2018) and Yildiz & Savelsbergh (2019) study the meal delivery routing problem (MDRP) to minimize the travel time. An anticipatory customer assignment strategy is proposed by Ulmer et al. (2021) to study the stochastic dynamic pickup and delivery problem.

To operate efficiently, on-demand meal delivery service platforms should also act proactively in responding to the dynamics of demand in the city. With a limited number of hired couriers, platforms can strategically prioritize assigning couriers to areas with higher predicted demand in the near future. By doing so, more couriers will be available in high-demand areas, reducing the average waiting time for order pickups. To support real-time operational decisions, accurate and fast-to-generate short-term demand predictions of the city are necessary. However, to the best of our knowledge, little attention has been paid to short-term demand forecasting problem of OMDPs. Hess et al. (2021) compare the performance of mainstream forecasting methods in generating deterministic demand predictions per hour. In their case study, an exponential smoothing based model achieves the best performance when a rich amount of historical data is available, while random forecast regression outperforms the rest when the training data is limited.

This study has two main objectives. Firstly, we combine the benefits of both parametric time series and data-driven methods to create a fast-to-generate forecasting approach for the demands arising from different parts of the city in the next 15-minute. Secondly, we propose an adaptive framework to identify and predict the next demand hotspots, utilizing the semi real-time predictive information as inputs. Addressing these two objectives, we propose the dynamic demand hotspots forecasting framework, which can be used by the platforms to proactively optimize their fleet and demand management decisions in real-time.

2 METHODOLOGY

In this study, we propose the novel dynamic demand hotspots forecasting framework to assist the rebalancing decision-making of on-demand meal delivery platforms by predicting the areas with higher demand in the next 15-minute interval. Our framework consists of two major steps. Firstly, we generate demand predictions for different areas of the city. Then, we generate demand hotspot clustering of the city utilizing the predictive outcomes and the geographic information of these areas.

Compared to the forecasting for a longer time interval, short-term demand predictions are more volatile and prone to fluctuations caused by latent events. Parametric time series models are good at interpreting recent signals for forecasting, but they require manual effort for selecting suitable parameters during training. On the other hand, non-parametric data-driven machine learning models are able to capture the non-linear patterns in data and interpret the interactions between features. To combine the advantages of both sides, we propose an adaptive forecasting approach that recurrently applies lagged-dependent features to generate predictions, based on data-driven machine learning models. Inspired by autoregression analysis, the lagged-dependent features are the previous values of demand, e.g. y_{t-i}^α meaning the demand of area α from i time steps ago. In our model, each time step is a 15-minute interval. To distinguish the regular hourly seasonality and the unexpected temporal fluctuations of demand, we choose to include four lagged-dependent features $y_{t-1}^\alpha, y_{t-2}^\alpha, y_{t-3}^\alpha, y_{t-4}^\alpha$, covering the demand information from the recent one hour. And we adopt random forest regression (RF) and eXtreme Gradient Boosting (XGBoost) as the baseline machine learning models. Their adaptive versions we create are called LD-RF and LD-XGBoost respectively.

Although deterministic demand forecasting is the focus of many demand forecasting literature, the determination of future hotspots might be benefit from probabilistic predictive information, which provides distributional insight of demand. The probabilistic predictive information feature \hat{Y}_i we adopt in the dynamic demand hotspot clustering framework is a weighted average of the 25%, 50% and 75% quantile prediction vectors, $\hat{Y}_i = 0.25 \cdot \hat{Y}_{i,0.25} + 0.5 \cdot \hat{Y}_{i,0.50} + 0.25 \cdot \hat{Y}_{i,0.75}$. The quantile prediction vectors are provided by quantile regression forest (QRF) by Meinshausen & Ridgeway (2006) or its adaptive version LD-QRF in this study.

Besides of the predictions, we also include the geographic features latitude and longitude of the area centers as input to the clustering algorithm. Constrained K-means clustering proposed by Bradley et al. (2000) is utilized to generate clusters of areas with a condition of no less than three zones per cluster, where the optimal number of clusters is chosen between 3 to 6 based on the highest mean silhouette coefficient.

3 RESULTS AND DISCUSSION

Data analysis and feature engineering

Our case study uses the meal order placement data of a city, which is collected from April 1st, 2020 to September 14th, 2020. Each instance contains the placement time, pickup and delivery destination locations. The locations are hashed into hexagonal zones using Uber’s H3 geospatial indexing system at a resolution level of 8. The average area per zone is 0.737 km^2 . We denote the location of the restaurant as the pickup zone, and the location of delivery destination as the destination zone. Our data covers 20 pickup zones and 50 destination zones.

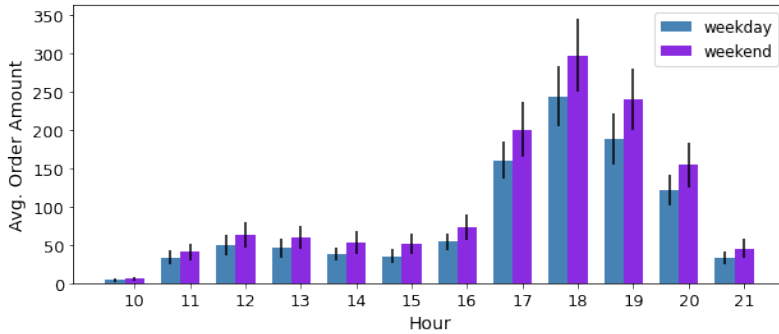


Figure 1: The average total number of received orders per hour for weekdays (i.e. Monday to Thursday) and weekends (i.e. Friday to Sunday), the error bars represent standard deviations.

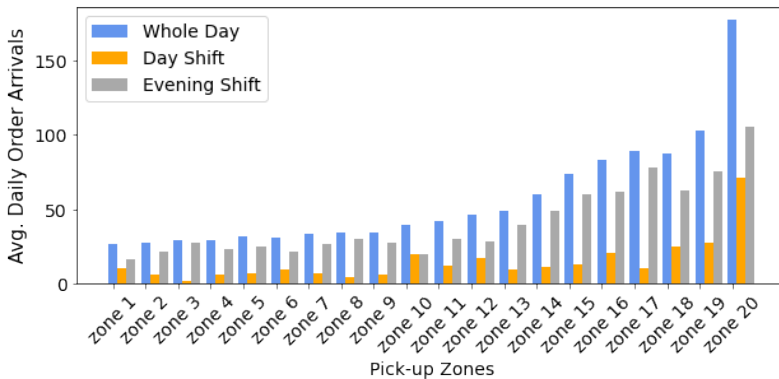


Figure 2: The daily average total number of received orders for each pickup zone at *whole day*, *day shift*, and *evening shift* respectively.

According to our data analysis, all days of the week exhibit a dual-peak demand pattern, which shows the demand level is much higher around dinner time compared to lunch time. Also, Friday, Saturday and Sunday attract more orders than other days in general. To distinguish this demand difference we visualise the average number of orders received per hour during the week and on weekends in Figure 1. We consider Friday to Sunday together as weekend days since the demand on Friday evening is much closer to the patterns of weekend days. A higher demand level is showed among the weekend group than the weekday group. Inspired by the within-day demand pattern, we can further split a day into the *day shift* (10:30-17:00) and the *evening shift* (17:00-21:30). To explore the demand level difference among pickup zones, we compare their daily average number of orders received in Figure 2. We observe a variation in the relative demand levels between the day and the evening shifts, indicating a change of comparative demand levels of the city throughout a day.

In this case study, we are interested in predicting the near future demands in the city, i.e. the aggregated amount of orders coming in the next 15 minutes from all the restaurants in each pickup zone. The per 15-minute demands are the aggregated number orders for each pickup zone according to the 44 different 15-minute time windows each day. We denote this target variable *order demand* of pickup zone α at the i^{th} time interval as y_i^α . To capture the within-day and within-week double-seasonal demand patterns in data, we include the temporal features hour and day of a week. A binary variable is adopted to indicate whether the date is a national holiday. Considering that weather also affects the interests of customers in online meal ordering, we include three exogenous weather features, namely the average temperature, precipitation amount and wind speed. The weather features are solely time-dependent, implying that the values measured at the same time are consistent across different zones.

Experimental design, evaluation metrics and baseline models

For model training, we use the first 21 weeks' data (6468 samples) from the processed dataset. And the predictions are generated for the last 1 week's data (308 samples) in a one-step-ahead

rolling window fashion. The predictions of each pickup zone are generated in parallel at each step by the corresponding predictors.

The first forecasting experiment involves performing deterministic demand forecasting for the next 15-minute. To gain an overview of model performance, we report the accuracy of each subset of data by properties ‘types of shift’ and ‘partition of week’. In addition to reporting the mean prediction errors averaged from predictions for all the pickup zones, we are also interested in comparing the individual prediction errors measured per pickup zone to inspect whether certain type of pickup zones are consistently better predicted by a kind of predictor. We apply the root-mean-squared error (RMSE) and mean absolute error (MAE) as the evaluation metrics for deterministic forecasting. To investigate the gains in deterministic forecasting using the data-driven lagged-dependent models, namely lagged-dependent regression forest (LD-RF) and lagged-dependent XGBoost (LD-XGBoost), we include the vanilla regression forest (RF) and XGBoost models as the baselines. Moreover, we also include Trigonometric, Box–Cox, Auto-Regressive-Moving-Average, Trend, and Seasonality (TBATS) model proposed by De Livera et al. (2011) as the conventional time series benchmark for its superiority in capturing complex seasonal patterns and convenience in automatic hyperparameters tuning function. Additionally, we include a two-layer Long short-term memory networks (LSTMs) as the deep learning benchmark, given its similarity in recurrent feature processing to our lagged-dependent methods.

The second forecasting experiment concerns identifying the demand hotspots of the city in the next 15-minute interval using the adaptive clustering framework we propose. The actual and predicted within-cluster median demand values are the medians taken from the clusters generated by actual demand and predictions inputs respectively. The deviation between the actual and predicted within-cluster medians serves as an indicator of predictive clustering performance. Again, RMSE and MAE are applied as error measurements.

Results of Case Study

The deterministic prediction accuracy of different models are reported in Table 1, measured by the averaged RMSE and MAE. The performance of the vanilla data-driven machine learning models well as their lagged dependent versions are rather close to each other, although XGBoost consistently predicts most accurately for the forecasting of day shift. And all data-driven models outperform the benchmarks TBATS and LSTMs by some margin in terms of prediction errors.

To confirm the statistical significance of prediction errors between models, we perform Diebold-Mariano tests for all pairs of models. Only the forecasting performance difference between RF and XGBoost is tested to be insignificant.

Table 1: The one-week point forecast results of various models trained with 21 weeks’ data, measured by MAE and RMSE.

PERIOD	MODEL	WHOLE WEEK		WEEKDAY		WEEKEND	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
<i>Whole Day</i>	TBATS	1.762	1.395	1.653	1.373	1.851	1.438
	LSTM	1.975	1.330	1.506	0.961	2.190	1.515
	RF	1.145	0.830	1.110	0.851	1.229	0.895
	LD-RF	1.155	0.838	1.092	0.835	1.237	0.906
	XGBoost	1.145	0.829	1.103	0.843	1.235	0.902
	LD-XGBoost	1.174	0.859	1.106	0.849	1.255	0.934
<i>Day Shift</i>	TBATS	1.514	1.272	1.613	1.424	1.500	1.257
	LSTM	1.029	0.698	0.551	0.370	1.169	0.815
	RF	0.770	0.582	0.664	0.566	0.847	0.655
	LD-RF	0.780	0.591	0.656	0.556	0.865	0.669
	XGBoost	0.769	0.582	0.654	0.555	0.847	0.658
	LD-XGBoost	0.796	0.611	0.680	0.575	0.885	0.697
<i>Evening Shift</i>	TBATS	2.040	1.572	1.675	1.336	2.205	1.700
	LSTM	2.799	2.244	2.163	1.719	3.087	2.527
	RF	1.514	1.185	1.489	1.238	1.607	1.242
	LD-RF	1.523	1.197	1.466	1.218	1.604	1.249
	XGBoost	1.515	1.185	1.484	1.234	1.619	1.256
	LD-XGBoost	1.549	1.216	1.478	1.220	1.627	1.275

Figure 3 shows the average RMSE and MAE per pickup zone obtained from different forecasting methods. As we analyzed in Figure 2, the average number of orders received per day increase as we move from pickup zone 1 to zone 20. The rise of demand level seems to be positively correlated to the average forecasting errors. And the forecasting performance of benchmarks TBATS and LSTMs are much poorer compared to that of the tree-based methods when applied to the high demand pickup zones.

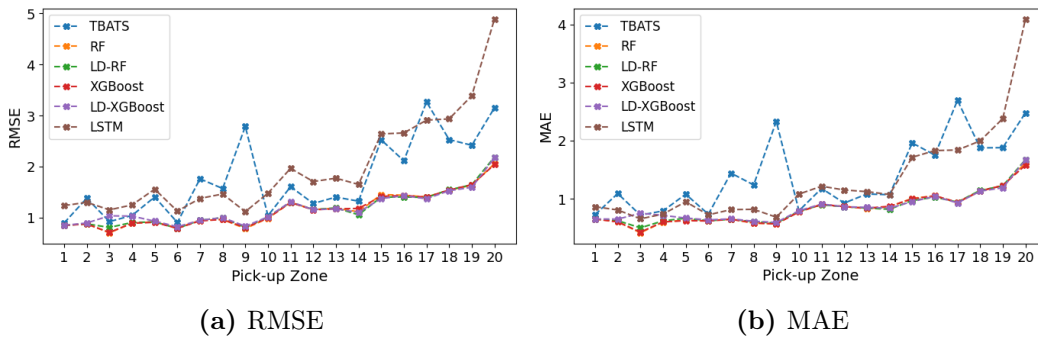


Figure 3: The RMSE and MAE of different models measured for each pickup zone at the general (i.e. whole day and whole week) level.

From the deterministic forecasting experiment, RF, LD-RF, XGBoost and LD-XGBoost have shown good performance. Therefore, we continue to apply these four predictors and utilize their deterministic predictions as part of the input for dynamic demand hotspots clustering. Additionally, we include the quantile forecasting approaches QRF and LD-QRF. The vanilla data-driven models are treated as the benchmarks here to evaluate the benefit of using lagged-dependent models as well as applying probabilistic predictive information as input.

Figure 4 shows an example visualization of demand hotspots clustering, where the grids are colored based on the within-cluster average actual/predicted demands. In this example, adaptive predictors

LD-RF and LD-QRF manage to identify the central hotspot, although predicted within-cluster demands are all less than actual.

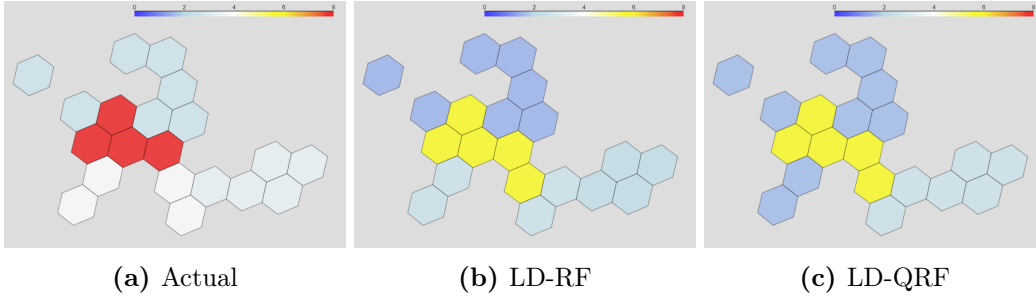


Figure 4: Visualization of clustered hotspots for 19:00-19:15 on Monday, September 14th using (a) actual demands, (b) predicted deterministic demands by LD-RF, and (c) predicted quantile demands by LD-QRF respectively.

The RMSE and MAE evaluations of the real and predicted within-cluster median difference are presented in Table 2. On the whole day level, the best RMSE and MAE are obtained by the adaptive predictors using lagged-dependent features, namely the LD-RF, LD-XGBoost and LD-QRF. Also the accuracy measured by MAE suggests the probabilistic predictors to perform better. It means using probabilistic predictive information generate less errors on average, although it is punished by some occasional larger errors.

Table 2: The RMSE and MAE of one-week’s predicted within-cluster median demand by dynamic clustering method using deterministic and probabilistic predictions from models trained by 21-week data.

PERIOD	MODEL	WHOLE WEEK		WEEKDAY		WEEKEND	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
<i>Whole Day</i>	RF	0.791	0.810	0.729	0.770	0.867	0.863
	XGBoost	0.795	0.808	0.730	0.762	0.874	0.869
	LD-RF	0.784	0.794	0.721	0.761	0.861	0.838
	LD-XGBoost	0.788	0.801	0.730	0.759	0.859	0.857
	QRF	0.831	0.716	0.781	0.660	0.893	0.791
	LD-QRF	0.820	0.699	0.767	0.648	0.886	0.766
<i>Day Shift</i>	RF	0.657	0.601	0.647	0.551	0.671	0.668
	XGBoost	0.660	0.599	0.648	0.543	0.677	0.673
	LD-RF	0.660	0.601	0.652	0.552	0.670	0.667
	LD-XGBoost	0.666	0.614	0.659	0.554	0.675	0.694
	QRF	0.718	0.483	0.729	0.412	0.713	0.577
	LD-QRF	0.712	0.484	0.727	0.422	0.691	0.568
<i>Evening Shift</i>	RF	0.952	1.112	0.834	1.086	1.090	1.145
	XGBoost	0.956	1.109	0.834	1.078	1.099	1.151
	LD-RF	0.935	1.072	0.811	1.063	1.080	1.084
	LD-XGBoost	0.936	1.071	0.822	1.055	1.070	1.092
	QRF	0.971	1.054	0.851	1.020	1.112	1.099
	LD-QRF	0.954	1.008	0.820	0.975	1.109	1.052

Again, we perform Diebold-Mariano tests on the within-cluster median predicting outcomes for each pair of models covered in the second experiment. Results show that the forecasting performance between RF and XGBoost, and between XGBoost and LD-XGBoost are not significantly distinguishable from one another. The statistical test is significant for other pairs.

4 CONCLUSIONS

This study proposes a dynamic demand hotspots forecasting framework that is able to assist operational decision making for on-demand meal delivery platforms. Through a case study, we show that the performance is improved by using the adaptive data-driven forecasting methods we propose, and the probabilistic predictive feature we create from the quantile demand predictions as input. A further study could focus on generating real-time resource-demand rebalancing decisions by incorporating the predicted insight provided by our framework.

REFERENCES

- Bradley, P. S., Bennett, K. P., & Demiriz, A. (2000). Constrained k-means clustering. *Microsoft Research, Redmond*, 20(0), 0.
- De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), 1513–1527.
- Hess, A., Spinler, S., & Winkenbach, M. (2021). Real-time demand forecasting for an urban delivery platform. *Transportation Research Part E: Logistics and Transportation Review*, 145, 102147.
- Meinshausen, N., & Ridgeway, G. (2006). Quantile regression forests. *Journal of Machine Learning Research*, 7(6).
- Reyes, D., Erera, A., Savelsbergh, M., Sahasrabudhe, S., & O’Neil, R. (2018). The meal delivery routing problem. *Optimization Online*.
- Ulmer, M. W., Thomas, B. W., Campbell, A. M., & Woyak, N. (2021). The restaurant meal delivery problem: Dynamic pickup and delivery with deadlines and random ready times. *Transportation Science*, 55(1), 75–100.
- Yildiz, B., & Savelsbergh, M. (2019). Provably high-quality solutions for the meal delivery routing problem. *Transportation Science*, 53(5), 1372–1388.

Incorporating Domain Knowledge in Deep Neural Networks for Mode Choice Analysis

Shadi Haj Yahia*, Omar Mansour, Tomer Toledo

Faculty of Civil and Environmental Engineering,
Technion – Israel Institute of Technology, Haifa 32000, Israel

Emails: shadi8@campus.technion.ac.il

omar@technion.ac.il

toledo@technion.ac.il

SHORT SUMMARY

Discrete choice models (DCM) are widely used in travel demand analysis to understand and predict choice behaviors. However, a priori specification of the utility functions is required for model estimation, leading to subjectivity and potential inaccuracies. Machine learning (ML) approaches have emerged as a promising solution but lack interpretability and may not capture expected relationships. This study proposes a framework that supports the development of interpretable models that incorporate domain knowledge and prior beliefs. The framework includes pseudo data samples and a loss function to measure relationship fulfillment. This approach combines the flexibility of ML structures with econometrics and interpretable behavioral analysis, improving model interpretability. The proposed framework's potential is demonstrated through a case study, providing a promising avenue for the advancement of data-driven approaches in DCM.

Keywords: Deep neural networks, discrete choice models, domain knowledge, interpretability

1. INTRODUCTION

Discrete choice models (DCMs) are used in travel demand analysis to understand individuals' decision-making processes. Most DCMs are formulated as random utility models (RUMs) that assume individuals make decisions based on maximizing utility. However, specifying a plausible utility model that captures these complexities is a challenging task (Torres et al., 2011). Recently, data-driven approaches using machine learning (ML) methods have emerged as a promising avenue to overcome the limitations of RUM specifications. Deep neural networks (DNNs) are an increasingly popular data-driven approach that has shown higher prediction accuracy in many tasks.

Unlike RUM, DNN models require essentially no a priori beliefs about the nature of the true underlying relationships among variables. DNN models can find complex non-linear specifications, and their high flexibility means that the role of the analyst is minimized. However, their “black box” form limits their interpretability, and the extracted relationships may not be consistent with domain knowledge (Van Cranenburgh et al., 2021; Wang et al., 2020b).

To address these limitations, some studies combine RUM and ML. For example, Sifringer et al., (2020, 2018) added a DNN-learned utility term to the traditional interpretable RUM utility function. This improves the model's fit to the data but the unbounded DNN term may dominate and prevent interpretation. The decision on which variables enter each part is subjective.

Another approach that was proposed by Wang et al., (2020a) is to use an alternative-specific utility deep neural network (ASU-DNN) architecture, which maintains separate utility functions for each alternative that depend only on its own attributes, resembling RUM. The model is more interpretable compared to fully connected DNNs and achieved comparable or better fit to the data. However, it still might suffer from unreasonable relationships among explanatory variables and choices.

Current methods for interpretability lack control over the relationships among variables and choices, making them inconsistent with domain knowledge and limiting their application in predicting new policies (Alwosheel et al., 2021, 2019). This study proposes a framework that incorporates domain knowledge through constraints and a loss function to penalize violations. The proposed approach preserves flexibility and can be implemented on any model architecture, providing control over the model's behavior for better travel choice predictions.

2. METHODOLOGY

The idea behind incorporating domain knowledge in DNNs involves augmenting the data given to the model and modifying the loss function that the model optimizes. To achieve this, additional data, termed pseudo data, is generated to hold the targeted knowledge that the model is expected to capture. The loss function is then formulated to include terms that use this data, in combination with the original model loss function, such as the negative log-likelihood. The additional loss terms measure the extent to which the trained model is consistent with the domain knowledge.

The overall framework for incorporating domain knowledge into DNNs is shown in Figure 1 and is independent of the model structure, allowing for seamless integration with existing DNN architectures. The model is trained on two sets of inputs: the originally available observed data and domain knowledge, which is mathematically formulated as a set of constraints on the outcomes of the trained model. The observed data represents the available dataset collected, including socio-economic characteristics of decision makers, attributes of the alternatives, and choices. The domain knowledge represents the knowledge that the modeler wants to incorporate into the model and expects to be captured (e.g., directions of sensitivities).

In this work, the modeler's prior expectations are related to signs of the effects of an alternative's attributes on its own utility. For example, these may be negative effects of mode travel times and costs on the utilities of these modes. In this case, the model is constrained to learn a monotonically decreasing probability of choosing an alternative with respect to its travel time and cost and, consequently, monotonically increasing probabilities of choosing the remaining alternatives.

Consider a training set consists of N samples $\{(x_i, y_i)\}, i = 1, \dots, N$, where x_i is a feature vector in $x \in \mathbb{R}^D$, and y_i is the discrete choice among \mathcal{C} alternatives, $y_i \in \{1, \dots, \mathcal{C}\}$. Let $p_c(x_i)$ be the probability of choosing alternative c given input x_i , and $x_i[m]$ is the value of feature m in the feature vector. The estimated model is considered to be monotonically increasing in p_c with respect to feature m if $p_c(x_i) \geq p_c(x_j)$ for any two feature vectors x_i, x_j , such that $x_i[m] \geq x_j[m]$ and $x_i[h] = x_j[h]$, for all $h \in \mathcal{D} \setminus m$. The opposite applies for decreasing monotonicity. The rest of the components are described as follows.

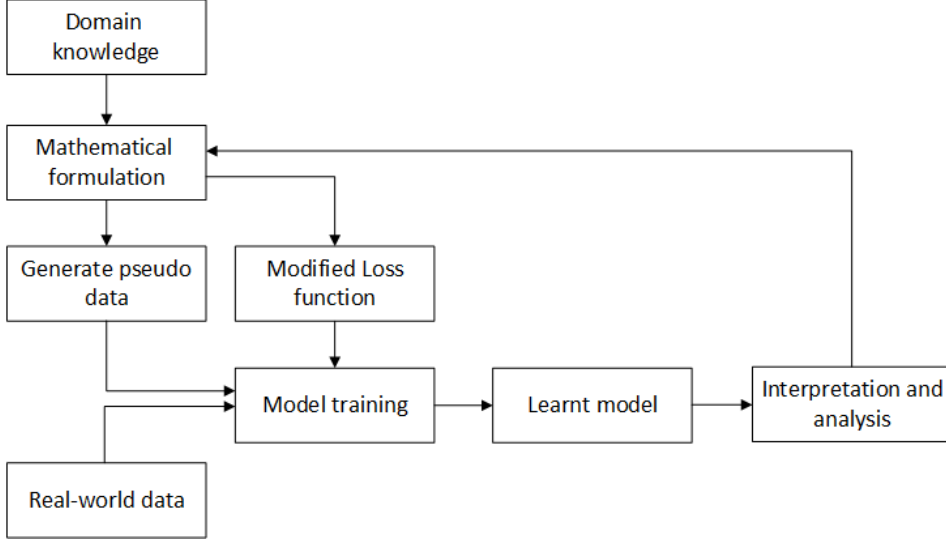


Figure 1. Overall framework for incorporating domain knowledge

Generating Pseudo Data

Following monotonicity constraints above, pseudo data can be generated as pairs of samples to numerically approximate the probabilities' derivatives that are constrained. For each monotonicity constraint with respect to a feature m , K pseudo samples are generated uniformly along the region values of that feature $x_{k,1}^*$. Each pseudo sample is then paired with another pseudo one, such that the second pseudo sample has a positive incremental change applied to feature m . The relationship required for an increasing monotonicity constraint of probability of choosing alternative c with respect to feature m is $p_c(x_{k,2}^*) - p_c(x_{k,1}^*) \geq 0$.

The pseudo data does not require labels (i.e., chosen alternatives), as they are only used for capturing domain knowledge, not for predicting the chosen alternative. This ability to generate pseudo samples enhances the model in three ways:

1. When the dataset is small, the pseudo dataset helps increase the dataset size to learn the model's parameters.
2. When the input feature region is imperfectly covered, the pseudo data helps fill gaps and enforce the model to learn along the full range of possible values.
3. Generating pseudo data outside the range of current values for specific features helps enforcing better learning, hence enabling extrapolation in the outer regions (i.e., unseen scenarios).

Loss Function

The loss function includes two components: prediction loss and domain knowledge loss. The prediction loss quantifies the accuracy of predictions and can be calculated for example using the negative log-likelihood (\mathcal{L}_{NLL}) method commonly used in RUMs. This calculation is performed only for samples with observed choices and is represented by the following formula:

$$\mathcal{L}_{NLL} = - \sum_{i=1}^N \sum_{c \in \mathcal{C}} g_{i,c} \cdot \log(p_{i,c}) \quad (1)$$

Where $g_{i,c}$ equals 1 if alternative c is chosen by individual i and 0 otherwise.

The domain knowledge loss measures the violation of monotonicity constraints on the probability of choosing alternative c with respect to feature m . This is determined using pseudo sample pairs that estimate the derivatives of the probabilities, represented by the following formula:

$$\mathcal{L}_{c,m} = \sum_{k=1}^K \max\left(0, d_{c,m} \cdot \frac{p_c(x_{k,2}^*) - p_c(x_{k,1}^*)}{\Delta x_m^*}\right) \quad (2)$$

Where $d_{c,m}$ equals 1 if the probability of choosing alternative c with respect to feature m should be increasing and -1 otherwise.

If it is assumed that when the probability of choosing alternative c with respect to feature m is in one direction, the probability of choosing other alternatives should be in the opposite direction, the total loss to be minimized can be expressed as follows:

$$\min \mathcal{L}_{total} = \mathcal{L}_{NLL} + \sum_{m \in M} \sum_{c \in C} w_{c,m} \cdot \mathcal{L}_{c,m} \quad (3)$$

Where M represents the indices of the features that constrain the probabilities, and $w_{c,m}$ represents the weight of each constraint violation penalty.

Model Training

The training process is illustrated in Figure 2. Observed data, represented as vector \mathbf{x} , and a vector of pseudo sample pairs $\mathbf{x}^* = \{(x_{1,1}^*, x_{1,2}^*), \dots, (x_{k,1}^*, x_{k,1}^*)\}$ are fed into the model. The total loss, calculated as a combination of the prediction loss from the observed samples \mathbf{x} and the domain knowledge loss from the pseudo data \mathbf{x}^* , is minimized using the backpropagation technique. This process continues iteratively until convergence is achieved.

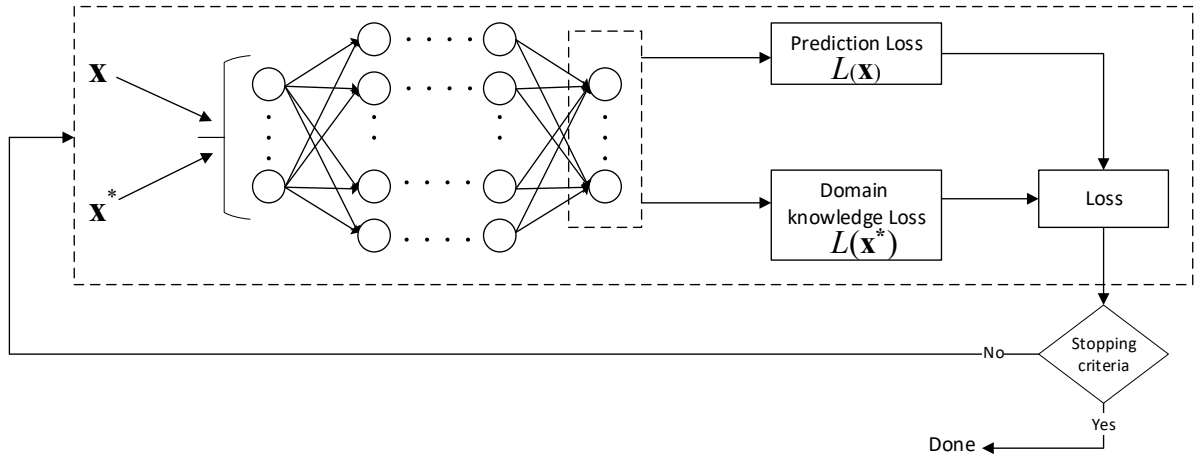


Figure 2. Model training process

3. RESULTS AND DISCUSSION

The methodology outlined above was applied to a mode choice dataset to assess the potential of incorporating domain knowledge in a DNN model and examine the impact of such knowledge on the resulting economic information.

Dataset

The experiment was based on the Swissmetro dataset, which is a publicly available stated preference survey collected in Switzerland in 1998 (Bierlaire et al., 2001). Participants were asked

to provide information regarding their preferred mode of transportation between the new Swissmetro (SM) mode, car, and train. Travel time and cost were considered as the key descriptive variables for each alternative mode. Observations with missing alternatives or outliers were removed. The dataset was then divided into training, validation, and testing sets in the ratio of 60:20:20.

Experimental Design

The proposed methodology was implemented on two model architectures: DNN and ASU-DNN. The DNN model was an off-the-shelf model, while the ASU-DNN model was proposed by Wang et al., (2020a) and calculates alternative-specific utilities. Both models were estimated in both an unconstrained and a constrained (i.e., with domain knowledge) version. The constrained models are referred to as C-DNN and C-ASU-DNN, respectively. In addition, a Multinomial Logit (MNL) model was also estimated for comparison.

The domain knowledge incorporated in the constrained models includes negative own-sensitivities of choice probability to travel time and cost and positive cross-sensitivities. All constraints are incorporated simultaneously. The models' negative log-likelihood and prediction accuracy were measured on each of the datasets. Predicted market shares were also calculated for each model. Choice probabilities with respect to each feature were then presented to demonstrate the fulfillment of domain knowledge.

Results

Prediction performance

Table 1 presents the negative log-likelihood (NLL) and accuracy of each estimated model. The results indicate that the DNN model provides the best NLL and accuracy, thanks to its high ability of empirical fit to data. The ASU-DNN also demonstrates good fit to data. When domain knowledge is introduced, constrained models become less flexible and achieve lower fit to data compared to unconstrained ones. This is expected since the introduction of constraints to the models limits the search space for the optimal fit and might restrict the flexibility of the model. Nonetheless, the decrease in accuracy in testing is only 2.1%.

Table 1. Negative log-likelihood (NLL) and prediction accuracy

Model	Training		Testing	
	NLL	Acc [%]	NLL	Acc [%]
DNN	3182	70.5	1133	69.2
Constrained DNN	3336	68.7	1189	67.1
ASU-DNN	3438	68.3	1188	67.7
Constrained ASU-DNN	3577	66.7	1235	65.6
MNL	3508	67.9	1209	66.2

Market shares

While prediction accuracy relates to predicting choices at the level of individuals, transportation policy planners are mainly interested in prediction at the market level. Table 2 shows the predicted market shares by the different models and the root mean square error (RMSE) in each model. The

constrained models provide better market shares in terms of RMSE in the training set compared to the unconstrained models but perform worse in the testing set. The DNN and ASU-DNN models outperform the MNL model in terms of RMSE in the testing set, which was guaranteed to provide exact market shares in training. Although the constrained models have worse performance than the unconstrained models, the RMSE values are within a range of 2.4% and are not much different from the observed shares in the sample.

Table 2. Market shares of travel modes

Training set						
	DNN	C-DNN	ASU-DNN	C-ASU-DNN	MNL	Observed
Train	5.5%	6.0%	7.6%	4.6%	6.2%	6.2%
SM	55.0%	58.3%	54.7%	56.3%	56.9%	56.9%
Car	39.5%	35.7%	37.7%	39.1%	36.9%	36.9%
RMSE	1.9%	1.1%	1.6%	1.6%	0%	
Testing set						
	DNN	C-DNN	ASU-DNN	C-ASU-DNN	MNL	Observed
Train	5.4%	6.1%	7.5%	4.5%	6.2%	7.4%
SM	55.9%	58.7%	55.2%	56.8%	57.7%	55.4%
Car	38.7%	35.1%	37.4%	38.6%	36.1%	37.2%
RMSE	1.5%	2.4%	0.2%	2.0%	1.6%	

Choice probabilities

To demonstrate the consistency with expected domain knowledge, choice probability functions provided by the different models were calculated as a function of each of the six variables. They were calculated using the partial dependence plots (PDP) method which calculates choice probabilities for every possible value of the variable for each observation (Friedman, 2001). Three of them are illustrated in Figure 3-5. In remaining three, the constrained models satisfied the constraints while unconstrained ones did not. They are not presented due to paper length constraints.

The estimated coefficients in the MNL are with the expected sign (i.e., negative coefficients of travel time and cost in all utility functions), therefore, the directions of choice probabilities are consistent with domain knowledge as can be seen in Figure 3-5. However, choice probabilities may not always be consistent with domain knowledge when derived from unconstrained models, even in ASU-DNN where utilities are calculated independently from the others following RUM. This inconsistency could be restrained when domain knowledge is incorporated into the models.

For example, Figure 3 presents choice probabilities as a function of train travel time. It is expected that SM and car shares would increase at the expense of the decrease in train shares, as train travel time increases. Figure 3(a) shows that DNN, while the most accurate, reveals unreasonable decreasing of SM choice probability. This finding is unreasonable since train becomes less attractive, and SM shares should not be negatively affected. This knowledge is

considered in C-DNN, and choice probabilities become more reasonable as illustrated in Figure 3(b). The rest of the models behave as expected.

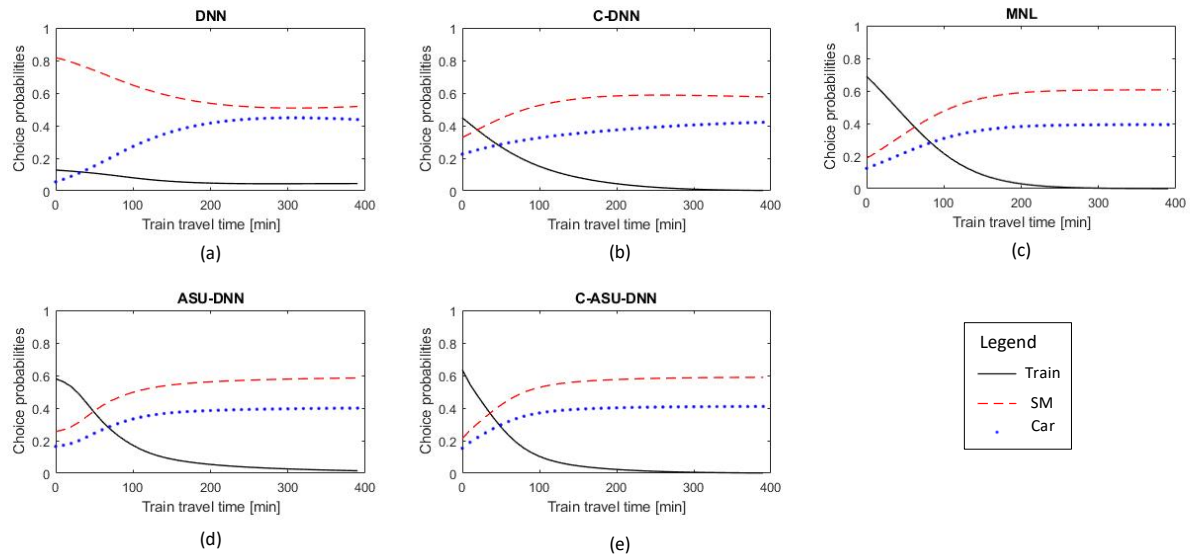


Figure 3. Alternatives' choice probabilities as a function of train travel time

In Figure 4, choice probabilities are calculated as a function of train cost. It is expected that SM and car shares would increase at the expense of the decrease in train shares, as train cost increases. However, DNN fulfills this expectation only up to about 150 CHF train cost, as shown in Figure 4(a). At this point, SM shares increase drastically at the expense of car shares, which start decreasing. This is unexpected since increased train cost must not negatively affect car shares. At worst, car shares would not change (i.e., would not increase), and train users would shift only to SM. Another unexpected finding can be found in the ASU-DNN model in Figure 4(d). Around a train cost of 150 CHF, all choice probabilities switch directions. Both models were corrected by incorporating knowledge, as shown in Figure 4(b) and Figure 4(e) for C-DNN and C-ASU-DNN, respectively.

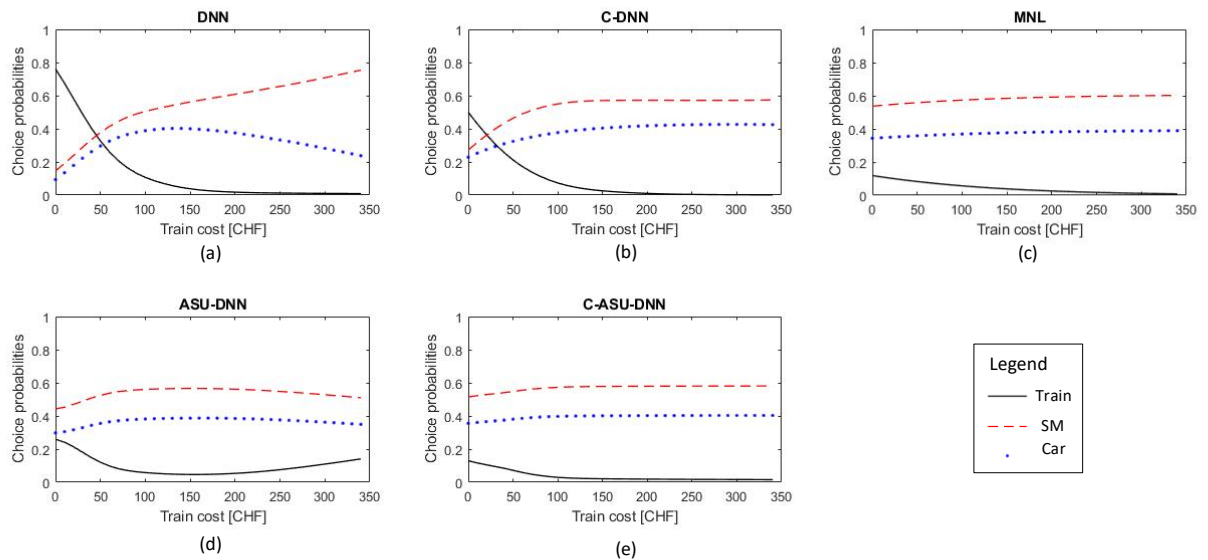


Figure 4. Alternatives' choice probabilities as a function of train cost

Figure 5 presents choice probabilities as a function of SM cost, where train and car shares are expected to increase at the expense of decrease in SM shares. In Figure 5(a), DNN fails to fulfill this knowledge at costs above 170 CHF, where car shares start decreasing, as if higher SM cost makes using the car less attractive. Although domain knowledge was incorporated, C-DNN was not corrected. In ASU-DNN, however, all choice probabilities switch directions at 300 CHF SM cost, as illustrated in Figure 5(d), which had been overcome by incorporating knowledge, as shown in Figure 5(e).

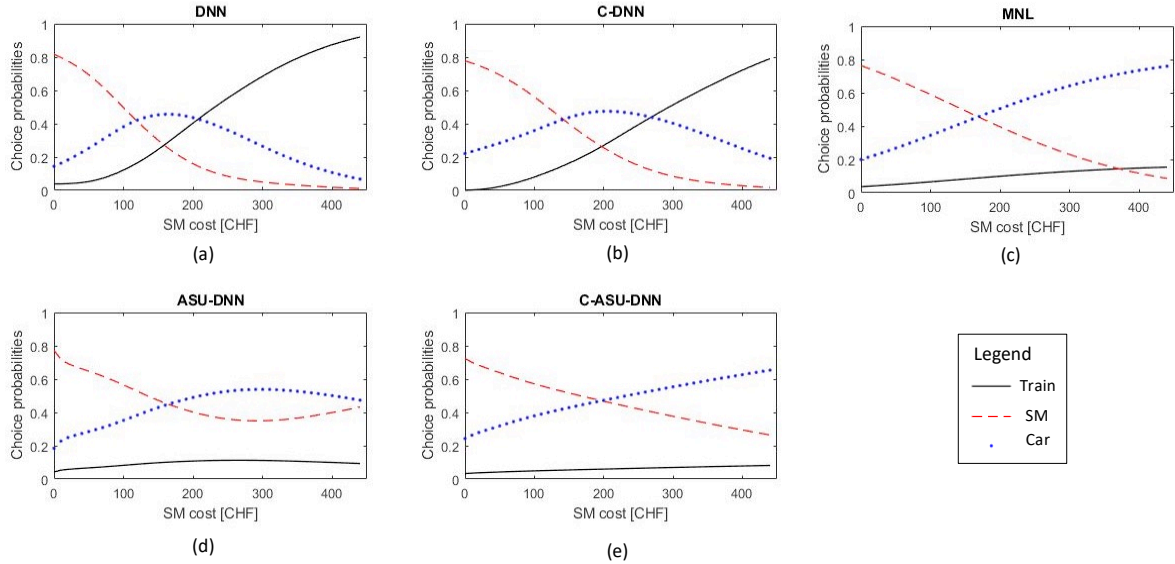


Figure 5. Alternatives' choice probabilities as a function of SM cost

In conclusion, while accurate, unconstrained models that rely solely on data may produce unreasonable interpretations of choice probabilities, making them unsuitable for use in policy-making processes. The results obtained through the proposed methodology of incorporating domain knowledge into these models demonstrate the potential of achieving more interpretable results while still relying on data in a controllable manner. In C-DNN, only one constraint out of 18 was not fulfilled (i.e., increasing car choice probability as a function of SM cost, Figure 5 (b)), whereas all constraints were fulfilled in C-ASU-DNN. While constraints may not always be fully satisfied, they can significantly enhance the models' consistency with domain knowledge, making them more useful for choice analysis and planning purposes.

4. CONCLUSIONS

This study addresses the limitations of uncontrollable DNN application in discrete choice analysis. Incorporating domain knowledge into DNN is crucial for its interpretation and usability. The proposed framework enhances model consistency with domain knowledge by introducing constraints, making it easy to implement on different architectures. The case study on Swissmetro dataset demonstrates a tradeoff between accuracy and interpretability, showing promising results in combining domain knowledge with DNN models for choice analysis. Future research could explore the proposed framework's generalizability to other discrete choice modeling problems and datasets. The proposed framework could also be extended to incorporate other types of domain knowledge, such as prior distributions on model parameters or constraints on the functional form of the model.

REFERENCES

- Alwosheel A., van Cranenburgh S., and Chorus C. G. (2021). “Why did you predict that? Towards explainable artificial neural networks for travel demand analysis”. *Transportation Research Part C: Emerging Technologies*, 128, 103143.
- Alwosheel A., Van Cranenburgh S., and Chorus C. G. (2019). ““Computer says no’ is not enough: Using prototypical examples to diagnose artificial neural networks for discrete choice analysis”. *Journal of choice modelling*, 33, 100186.
- Bierlaire M., Axhausen K. and Abay G. (2001). “The acceptance of modal innovation: The case of Swissmetro”. In ‘Proceedings of the Swiss Transport Research Conference’, Ascona, Switzerland.
- Friedman J. H. (2001). “Greedy function approximation: a gradient boosting machine”. *Annals of statistics*, 1189-1232.
- Sifringer B., Lurkin V., and Alahi A. (2018). “Let me not lie: Learning multinomial logit”. arXiv preprint arXiv:1812.09747.
- Sifringer B., Lurkin V., and Alahi A. (2020). “Enhancing discrete choice models with representation learning”. *Transportation Research Part B: Methodological*, 140, 236-261.
- Torres C., Hanley N., and Riera A. (2011). “How wrong can you be? Implications of incorrect utility function specification for welfare measurement in choice experiments”. *Journal of Environmental Economics and Management*, 62(1), 111-121.
- Van Cranenburgh S., Wang S., Vij A., Pereira F., and Walker J. (2021). “Choice modelling in the age of machine learning”. arXiv preprint arXiv:2101.11948.
- Wang S., Mo B., and Zhao J. (2020a). “Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions”. *Transportation Research Part C: Emerging Technologies*, 112, 234-251.
- Wang S., Wang Q., and Zhao J. (2020b). “Deep neural networks for choice analysis: Extracting complete economic information for interpretation”. *Transportation Research Part C: Emerging Technologies*, 118, 102701.

Modeling the Demand for Bicycle Parking Facilities

David Kohlrantz*¹, Tobias Kuhnimhof²

¹ Chair and Institute of Urban and Transport Planning, RWTH Aachen University, Germany

² Chair and Institute of Urban and Transport Planning, RWTH Aachen University, Germany

SUMMARY (147 WORDS)

Improving opportunities for bicycle parking is essential for promoting cycling. However, there is a lack of approaches for predicting the demand for bicycle parking based on the facility type and the facility's location. Considering both during planning could help improve bicycle parking according to user needs. This is particularly applicable when cyclists face several parking options, such as on university campuses, as in our case study. The paper presents a stated preference-based model, which was additionally calibrated using bicycle parking count data.

Considering facility types improves the model fit substantially. Furthermore, the stated preference-based, original model underestimates the sensitivity to walking distances between facilities and buildings. When cyclists can choose between multiple parking facilities, it is critical to consider walking distances to realistically predict the demand for bicycle parking facilities. This confirms previous findings, that positioning parking facilities close to destinations is essential for attractive parking infrastructure.

Keywords: Bicycle parking, cycling, cycling behavior, demand modeling

Conference topic: Cycling and walking behavior and design

SHORT PAPER (2,842 WORDS WITHOUT TABLES)

1. INTRODUCTION

Improving bicycle parking infrastructure is, in addition to measures for moving bicycle traffic, essential to promote cycling (Heinen and Buehler, 2019). Even though many previous studies analyzed bicycle parking behavior and preferences, research does not yet cover modeling the demand for single facilities considering facility type and position. We present a model for bicycle parking facility demand at RWTH Aachen University, one of the largest technical universities in Germany (45.000 students, 8.000 employees). We model bicycle parking behavior based on a stated preference experiment among RWTH students and staff, focusing on privately owned bicycles. Specifically, we analyze to which degree the following factors are relevant for the prediction of bicycle parking demand:

1. Type of parking facility
2. Cycling detour (additional cycling distance to access the parking facility compared to parking the bicycle directly at the destination building entrance)
3. Walking distance between parking facility and destination (building entrance)

First, we review previous studies regarding bicycle parking before we describe our method, present our results, discuss them, and draw a conclusion.

2. REVIEW

Promoting cycling is one approach to increase the sustainability of mobility, especially in dense urban areas. One way is the improvement of bicycle parking facilities. For a general review of bicycle parking preferences and behavior, particularly at the workplace, we refer to Heinen and Buehler (2019).

Studies show that improving parking facilities increases the probability of commuting by bicycle. Several studies found a strong impact (Bueno et al., 2017; Hunt & Abraham, 2007; Noland & Kunreuther, 1995), while others only estimated a low or even statistically insignificant one (Handy & Xing, 2011; Stinson & Bhat, 2004). Furthermore, research showed that cyclists prefer sheds over parking racks (Lusk et al., 2014; Moskovitz & Wheeler, 2011; Yuan et al., 2017).

Less literature focuses on the influence of parking facility location-related factors. E.g., Molin and Maat (2015) found that the utility of bicycle parking facilities decreases when walking time increases. Papers and guidelines recommend short distances between parking facilities and buildings because users otherwise do 'fly parking' at facilities not intended for bicycle parking (Dufour, 2010; FGSV, 2012; Gamman et al., 2004; Larsen, 2015).

Previous models predicting the parking demand do not focus on single facilities and their attributes as in our paper, e.g.:

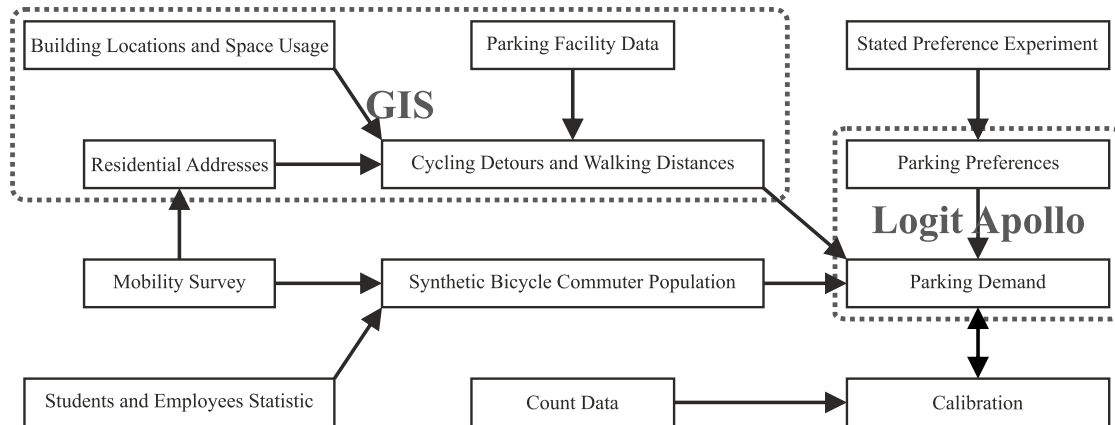
- Xu et al. (2012) developed a model for a university campus based on time series and attraction rates per building.
- Pfaffenbichler and Brezina (2016) analyzed the demand for public bicycle parking facilities in Vienna based on mode share and differentiating city districts, but not single facilities.
- Veillette et al. (2018) modeled the demand for bicycle parking on the grid cell level for Québec city.

3. METHODOLOGY

Figure 1 shows our approach to model bicycle parking choices. Firstly, we generated a synthetic university student and employee population commuting by bicycle based on mobility data and RWTH statistics. Hence, our total demand for bicycle parking includes the number of relevant students and employees per building.

Secondly, we calculated cycling detours and walking distances from parking facilities to the building entrances of their destination using a GIS. Thirdly, we used a mixed logit model to analyze a stated preference experiment. Fourthly, we applied the model to the synthetic population of bicycle commuters in order to predict their choice of bicycle parking and, thus, the total demand for bicycle parking per facility. Fifthly, we compared the predicted bicycle parking occupancy with bicycle parking count data. Finally, based on discrepancies between predicted and measured occupancy rates, we returned to the model and calibrated model parameters to reflect real bicycle parking behavior more adequately.

Figure 1: Model overview (own illustration)



Synthetic student and employee bicycle commuter population

In order to apply our model to the RWTH campus, we required a synthetic bicycle commuter population including a) group affiliation (students, professors, scientific employees, administrative and technical staff (ATS)), b) geographic direction of residential location (i.e., origin of commute trip) and c) building of work or study place (i.e., destination of commute trip). We used the results of a university mobility survey (n = 3,841) mailed out to all students and employees to generate the cycling commuters and assigned them by space usage data to buildings.

Stated preference experiment

For our analysis, we use the results of a web-based stated preference experiment conducted among RWTH students and employees in July 2022 (n = 2960). In this experiment, participants had to choose one of the following alternatives to park their bicycle:

- indoor parking in the building of their place of work respectively study (if possible in the status quo)
- a traffic sign pole representing 'fly parking'
- uncovered parking rack
- covered parking rack
- bicycle parking station

These alternatives were associated with varying cycling detours, walking distances, and prices, enabling us to analyze the attributes' influence with a mixed logit model using the R package Apollo (Hess & Palma, 2022).

Table 1 shows the models' coefficients for different facility types, taking the interactions with the resale value of the bicycle (RV) and group affiliation (scientific employees (reference category), students, professors, and ATS) into account.

Overall, the results show that – while there are differences between the various groups – the type of parking facility, whether it is covered or not, and the walking distance matter to cyclists. Our later findings in this paper will support these results, where we compare predicted and real bicycle parking on the RWTH university campus. However, the stated preference model also indicates that cycling detours significantly influence the probability of choosing a parking facility. Later on, our application to the campus will not confirm this finding.

Table 1: Coefficients mixed logit model

		Est.	Std. err.	t-ratio	p-value
<i>Indoor parking</i>	$\mu(\beta)$	-2.940	0.299	-9.828	<2E-12
	$\sigma(\beta)$	5.146	0.160	32.237	<2E-12
<i>Indoor parking</i> _{Student}	β	-2.419	0.271	-8.929	<2E-12
<i>Indoor parking</i> _{ATS}	β	1.784	0.375	4.758	1.96E-06
<i>Indoor parking</i> _{RV > 500 €}	β	1.740	0.300	5.808	6.32E-09
<i>Indoor parking</i> _{RV > 1,000 €}	β	1.304	0.441	2.955	0.003
<i>Indoor parking</i> _{No designated space}	β	-0.965	0.287	-3.359	7.82E-04
<i>Indoor parking</i> _{Forbidden in building}	β	-0.894	0.272	-3.282	0.001
<i>Indoor parking</i> _{Forbidden at department}	β	-0.936	0.420	-2.230	0.026
<i>Pole of a traffic sign</i>	$\mu(\beta)$	-2.032	0.075	-26.953	<2E-12
	$\sigma(\beta)$	1.945	0.072	26.972	<2E-12
<i>Uncovered bicycle parking rack</i>	$\mu(\beta)$		fixed		
	$\sigma(\beta)$	1.381	0.065	21.104	<2E-12
<i>Covered bicycle parking rack</i>	$\mu(\beta)$	0.656	0.066	9.899	<2E-12
	$\sigma(\beta)$	-1.547	0.065	-23.936	<2E-12
<i>Covered bicycle parking rack</i> _{RV > 500 €}	β	0.874	0.104	8.368	<2E-12
<i>Bicycle parking station</i>	$\mu(\beta)$	0.876	0.164	5.349	8.86E-08
	$\sigma(\beta)$	2.864	0.085	33.552	<2E-12
<i>Bicycle parking station</i> _{Student}	β	-0.495	0.181	-2.733	0.006
<i>Bicycle parking station</i> _{ATS}	β	-0.620	0.333	-1.861	0.063
<i>Bicycle parking station</i> _{RV > 500 €}	β	1.489	0.199	7.488	6.99E-14
<i>Bicycle parking station</i> _{RV > 1,000 €}	β	1.258	0.254	4.961	7.03E-07
<i>Bicycle parking station</i> _{Dist. to RWTH [km]}	β	0.045	0.018	2.551	0.011
<i>Cycling detour [m]</i>	β	-0.006	3.21E-04	-19.104	<2E-12
<i>Cycling detour</i> _{Student} [m]	β	-0.002	3.93E-04	-5.840	5.23E-09
<i>Cycling detour</i> _{Professor} [m]	β	-0.002	0.001	-2.868	0.004
<i>Cycling detour</i> _{ATS} [m]	β	0.001	0.001	1.950	0.051
<i>Walking distance [m]</i>	β	-0.016	4.23E-04	-38.871	<2E-12
<i>Walking distance</i> _{Student} [m]	β	-0.002	0.001	-4.380	1.19E-05
<i>Walking distance</i> _{Professor} [m]	β	0.004	0.001	3.147	0.002
<i>Walking distance</i> _{ATS} [m]	β	0.006	0.001	8.588	<2E-12

Parking facility data

In the stated preference experiment, we only included u-racks, also known as Sheffield racks, allowing for locking the bike frame to the stand. However, several facilities on the campus only allow locking the front wheel ('front racks'). To consider a higher theft risk, we applied the

coefficient for 'pole of a traffic sign' to them; for covered front racks, we used the coefficients for 'covered bicycle parking rack' on top.

We also aggregated the demand for close together located parking facilities of the same type. As a result, the number of analyzed facilities decreased from 163 to 99 shown in Table 2.

Table 2: Bicycle parking facilities in our analysis

	Front rack		Bicycle parking rack		Bicycle parking station
	Uncovered	Covered	Uncovered	Covered	
Number of aggregated facilities	15	2	73	8	1
Total capacity (bicycle parking spaces)	650	15	3814	452	543

To calculate the cycling detour to reach each facility, we measured the beeline distance between each geographic center of the residential addresses (trip origin) and parking facilities. Then, we calculated the distance between the trip origin and the building entrance (trip destination). The difference between them defines the (positive or negative) cycling detour. (This simple approach led to unrealistic results for eleven buildings and eight parking facilities, and we manually measured cycling detours with aerial images for these instances.) We used the beeline distance between the parking facility and the building entrance to determine the walking distance. Further, we assumed that poles of traffic signs (i.e., 'fly parking') at a walking distance of 60 m are available at each building.

Count data

We counted the occupancy of parking facilities at RWTH primarily on Thursday, 28.04.2022, in the morning (10-12) and afternoon (13-15), during a week with changing weather conditions. The counting phase took place shortly after the expiration of COVID-19 restrictions when many employees were still working from home, and many lectures and exercises were still web-based. Furthermore, our model does not consider temporal overlap. Therefore, we scaled our predicted demand by the quotient of counted and predicted demand, around one-third. Because the morning and afternoon results are similar, we only show the results for the afternoon.

Calibration

We used the count data to assess to which degree our prediction realistically reflects bicycle parking behavior. In the process, we calibrated the model by multiplying the cycling detour and the walking distance by factors ($F_{Cycling\ detour}$, $F_{Walking\ distance}$). We analyzed values between 0 and 5 regarding the correlation and the root mean square error (RMSE). The model names represent the factors, e.g., C_1W_3 means that $F_{Cycling\ detour} = 1$ and $F_{Walking\ distance} = 3$.

4. RESULTS

Apart from several models with different $F_{Walking\ distance}$ and $F_{Cycling\ detour}$, we evaluated a base model, as shown in Table 3. The base model assigns all demand generated by buildings to the closest bicycle parking facility, already explaining more than half of our demand differences between facilities. While the model based on the stated preference experiment using the calculated

beelines once has a lower correlation than the base model, increasing the $F_{Walking\ distance}$ improves the correlation to the counts up to two-thirds and reduces the RMSE. However, the calibration of the $F_{Cycling\ detour}$ showed no significant contribution for the prediction quality of our model.

Table 3: Model accuracy

Model	Base-model	C ₁ W ₀	C ₁ W ₁	C ₁ W ₂	C ₁ W ₃	C ₁ W ₄	C ₁ W ₅	C ₀ W ₃	C ₂ W ₃	C ₃ W ₃	C ₄ W ₃	C ₅ W ₃
$F_{Cycling\ detour}$	-	1	1	1	1	1	1	0	2	3	4	5
$F_{Walking\ distance}$	∞	0	1	2	3	4	5	3	3	3	3	3
Correl.	0.55	0.18	0.51	0.63	0.66	0.66	0.65	0.66	0.66	0.65	0.63	0.61
RMSE	22	33	22	19	18	19	19	18	19	10	19	20

In the following, we analyze the predictions for the C₁W₃ model. As Table 4 shows, the model overestimates the demand for covered parking racks and, in contrast, substantially underestimates the demand for front racks.

Table 4: Counted and modeled demand after calibration (C₁W₃)

	Indoor parking	Pole of a traffic sign	Front rack		Parking rack		Bicycle parking station
			Uncovered	Covered	Uncovered	Covered	
Predicted	806	747	72	7	1,536	286	32
Counted	-	-	132	11	1,504	201	85
Ratio	-	-	0.55	0.64	1.02	1.42	0.38

Regarding the facilities' geographic locations, Figure 3 shows some underpredictions due to missing demand data for specific buildings based on a lack of space usage data. For some places, an overprediction of the demand is also explainable. For example, the RWTH guest houses have a diverging mobility behavior compared to other university buildings, causing less demand than predicted. Furthermore, some locations with overpredicted demand have other facilities of the same or another facility type nearby. Taking only the closest facility of each type into account might cause prediction inaccuracies, and additional aggregation would be one solution.

Figure 2: Predicted and counted demand per parking facility (C₁W₃)

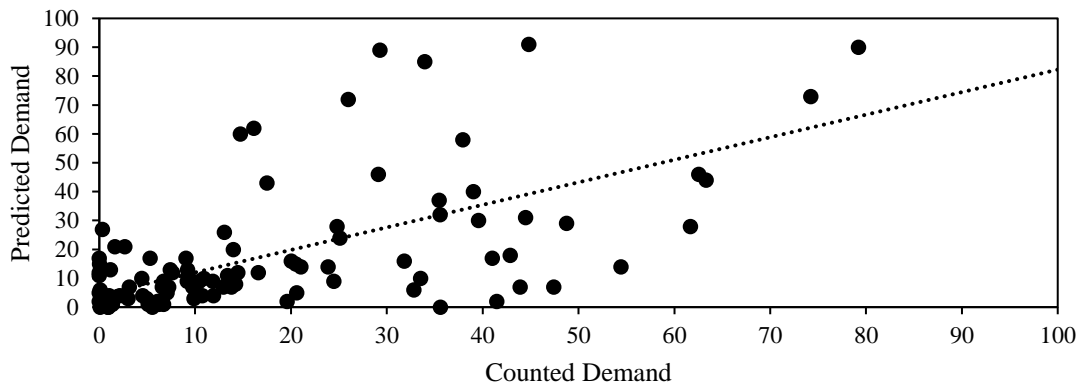
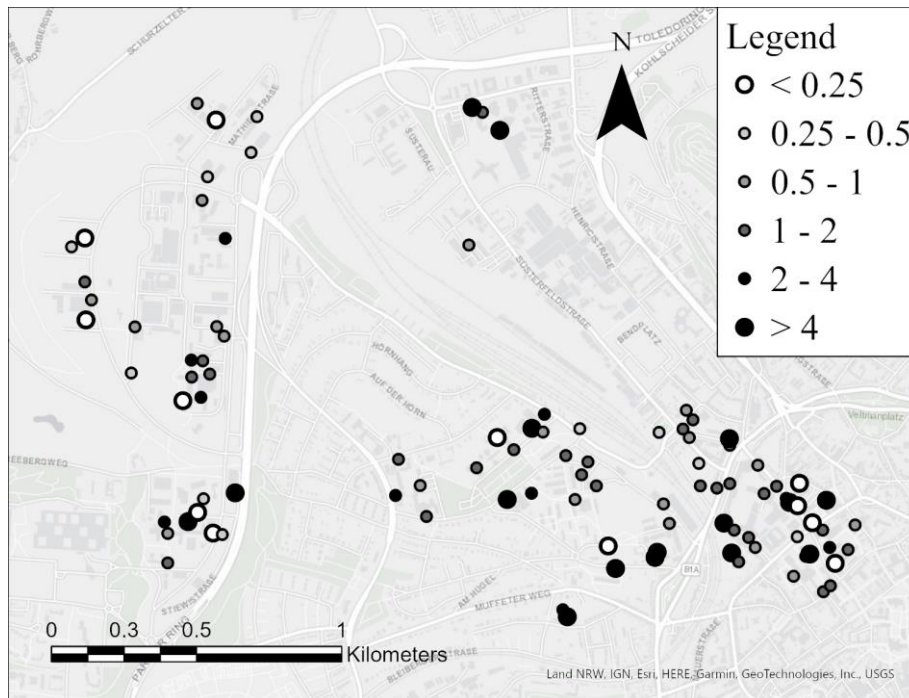


Figure 3: Ratio of predicted and counted demand per parking facility (C_1W_3)



5. DISCUSSION

Our findings show that the facility type is relevant for predicting the demand for bicycle parking because the consideration increases the correlation to 0.66. However, that almost three-fourths were uncovered parking racks might explain this limited increase in the correlation. In many cases, preferences slightly influence the parking choice because cyclists do not have the opportunity to choose between different facility types realistically.

According to our findings, the inclusion of cycling detours does not contribute to improving parking demand prediction. One explanation is that, in line with existing guidelines, most bicycle parking facilities are placed between the points of access to the respective property from public streets and building entrances (FGSV, 2012). Consequently, cycling detours hardly ever carry any weight, as cyclists usually arrive from the direction in which the parking facility is located. Nevertheless, we believe that our stated preference experiment's coefficients are proper. However, calculating exact cycling detours is complex, and their effect is limited compared to other factors such as walking distances and facility type. Finally, as close to the destination-located parking facilities have short cycling detours and walking distances in most cases, the correlation between both factors explains why considering cycling detours does not contribute substantially to the explanatory power of the prediction.

On the contrary, walking distances between parking facilities and trip destinations turned out to be more influential than our stated preference experiment suggested. One reason is that walking distances in the real-world built environment are longer than the beeline. While we assumed this would lead to a factor with a maximum of 1.5, we estimated a value of 3, optimizing our model. On average, real-world cyclists may not appreciate good parking facilities as much as our experiment participants. We deem it likely that survey participation was biased towards individuals with heightened interest in better parking facilities; one possible reason is that they

own expensive (e-)bikes. Therefore, survey participants may have been more willing to walk for better parking. This self-selection bias might also explain why the stated preference experiment overestimates the demand for low-quality front racks. However, our results underline the importance of good positioning of parking facilities. Otherwise, cyclists choose other options as street furniture, initially not designed for that purpose (FGSV, 2012; Gamman et al., 2004).

6. CONCLUSIONS

Our findings show that including the type of parking facility and the walking distances to building entrances improves the prediction of bicycle parking demand relative to a model solely based on the shortest distance to building entrances substantially. In this study, a stated preference experiment provided the user preferences constituting the basis for such an improved model. Bicycle parking counts contributed real-world bicycle parking data, which we used to perform a reality check of our model and to calibrate it to real-world circumstances.

There were some issues in our data that need to be taken into account. Firstly, our RWTH university mobility survey results indicate a higher demand for bicycle parking than counted. Abating COVID lockdown phenomena (e.g., a high proportion of people working and studying from home) and selectivity in our mobility survey may have contributed to this. Secondly, our results indicate a discrepancy between stated preference data-based parking preferences and real-world parking behavior. Hence, for further improvement of the approach presented in this paper, updated data collection and an in-depth investigation of the data discrepancies would be desirable. Another improvement of the approach would be the inclusion of effects induced because of occupancy, i.e., the question of to which degree demand for bicycle parking is diverted to other facilities if the desired facility is (completely) occupied. Due to low occupancy during our data collection period, we could not include these effects.

Nevertheless, we are confident that our model represents a substantial improvement in predicting the demand for bicycle parking compared to preexisting approaches. It is acknowledged that different groups of cyclists have diverging requirements regarding parking infrastructure concerning facility type and proximity to the destination. Simplistic models are not able to take account of that. Moreover, a model such as ours, which accounts for various attributes, may also be used to design a multi-optional bicycle parking facility layout that maximizes user benefit. Therefore, we believe that approaches such as ours will be increasingly needed to assess changes in the bicycle parking infrastructure and optimize extension strategies for bicycle parking in the context of promoting bicycle travel, e.g., regarding growing shares of e-bikes.

ACKNOWLEDGEMENTS

The authors confirm their contribution to the paper: research, analysis, and interpretation: David Kohlrautz. Supervision and advisory: Tobias Kuhnimhof. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- Bueno, P. C., Gomez, J., Peters, J. R., & Vassallo, J. M. (2017). Understanding the effects of transit benefits on employees' travel behavior: Evidence from the New York-New Jersey region. *Transportation Research Part a: Policy and Practice*, 99, 1–13. <https://doi.org/10.1016/j.tra.2017.02.009>

- Dufour, D. (2010). *PRESTO Cycling Policy Guide: Cycling Infrastructure*. Ligtermoet & Partners. https://transport.ec.europa.eu/system/files/2018-12/presto_cycling_policy_guide_infrastructure.pdf
- FGSV. (2012). *Hinweise zum Fahrradparken*. Köln. Forschungsgesellschaft für Straßen- und Verkehrswesen.
- Gamman, L., Thorpe, A., & Willcocks, M. (2004). Bike Off! Tracking the Design Terrains of Cycle Parking: Reviewing Use, Misuse and Abuse. *Crime Prevention and Community Safety*, 6(4), 19–36. <https://doi.org/10.1057/palgrave.cpcs.8140199>
- Handy, S. L., & Xing, Y. (2011). Factors Correlated with Bicycle Commuting: A Study in Six Small U.S. Cities. *International Journal of Sustainable Transportation*, 5(2), 91–110. <https://doi.org/10.1080/15568310903514789>
- Heinen, E., & Buehler, R. (2019). Bicycle parking: a systematic review of scientific literature on parking behaviour, parking preferences, and their influence on cycling and travel behaviour. *Transport Reviews*, 39(5), 630–656. <https://doi.org/10.1080/01441647.2019.1590477>
- Hess, S., & Palma, D. (2022). *Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application* [User Manual].
- Hunt, J., & Abraham, J. (2007). Influences on bicycle use. *Transportation*, 34, 453–470. <https://doi.org/10.1007/s11116-006-9109-1>
- Larsen, J. (2015). Bicycle Parking and Locking: Ethnography of Designs and Practices. *Mobilities*. Advance online publication. <https://doi.org/10.1080/17450101.2014.993534>
- Lusk, A. C., Wen, X., & Zhou, L. (2014). Gender and used/preferred differences of bicycle routes, parking, intersection signals, and bicycle type: Professional middle class preferences in Hangzhou, China. *Journal of Transport & Health*, 1(2), 124–133. <https://doi.org/10.1016/j.jth.2014.04.001>
- Molin, E., & Maat, K. (2015). Bicycle parking demand at railway stations: Capturing price-walking trade offs. *Research in Transportation Economics*, 53, 3–12. <https://doi.org/10.1016/j.retrec.2015.10.014>
- Moskovitz, D. A., & Wheeler, N. (2011). Bicycle Parking Analysis Using Time Series Photography. https://davidamoskovitz.files.wordpress.com/2010/08/bicycle_parking_analysis.pdf
- Noland, R. B., & Kunreuther, H. (1995). Short-run and long-run policies for increasing bicycle transportation for daily commuter trips. *Transport Policy*, 2(1), 67–79. [https://doi.org/10.1016/0967-070X\(95\)93248-W](https://doi.org/10.1016/0967-070X(95)93248-W)
- Pfaffenbichler, P. C., & Brezina, T. (2016). Estimating bicycle parking demand with limited data availability. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability*, 169(2), 76–84. <https://doi.org/10.1680/ensu.14.00063>
- Stinson, M. A., & Bhat, C. R. (2004). Frequency of Bicycle Commuting: Internet-Based Survey Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 1878(1), 122–130. <https://doi.org/10.3141/1878-15>
- Veillette, M.-P., Grisé, E., & El-Geneidy, A. (2018). Park' n' Roll: Identifying and Prioritizing Locations for New Bicycle Parking in Québec City, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(36), 73–82. <https://doi.org/10.1177/0361198118776522>
- Xu, J., Zhang, Z., & Rong, J. (2012). The Forecasting Model of Bicycle Parking Demand on Campus Teaching and Office District. *Procedia - Social and Behavioral Sciences*, 43, 550–557. <https://doi.org/10.1016/j.sbspro.2012.04.128>
- Yuan, C., Sun, Y., Lv, J., & Lusk, A. C. (2017). Cycle Tracks and Parking Environments in China: Learning from College Students at Peking University. *International Journal of Environmental Research and Public Health*, 14(8). <https://doi.org/10.3390/ijerph14080930>

Dynamic location for charging operations of shared free-floating e-scooters

Sara Momen^{*1}, Bart van Arem¹, and Shadi Sharif Azadeh¹

¹Department of Transport and Planning, Delft University of Technology, Netherlands

SHORT SUMMARY

Shared electric scooters (e-scooters) have recently become a popular mode of micromobility solution and this rapid growth causes significant operational challenges. One of the challenges micromobility companies face is collecting idle low-charge fleets from different corners of the city. The collector is usually a truck that needs to make a tour in town to collect these e-scooters. One solution can be for the operator to dynamically define special zones where e-scooters with low-charge are collected. In this paper, we propose a two-layered approach to tackle the problem. We propose a dynamic programming approach to investigate the required number and locations of designated low-charge drop-off points for the first layer of the problem. A simulated origin-destination data set of 30 e-scooters on TU Delft campus area is used for a time horizon of 8 working days and a time discretization level of 15 minutes. The results suggested consolidating low-charge e-scooters in three low-charge drop-off locations considering the state of charge (SOC) and distance of e-scooters from the low-charge drop-off zones. Then, we discuss the potential implications of our findings on recharging operations and spatial efficiency that are defined dynamically.

Keywords: Shared electric micro-mobility, dynamic programming, data-driven zoning, real-time operational decision, fleet rebalancing.

1 INTRODUCTION

A free-floating shared mobility system gives the user the flexibility to pick up and drop off the vehicle anywhere in the operating area. Operational issues are brought on by this flexibility, particularly in vehicle charging operations. A dynamic route design for the charging truck or overnight charging can be used to empower shared micromobility systems, but both of these methods are either very expensive for the system's performance or cause computational complexity Osorio et al. (2021).

In free-floating shared micromobility, the majority of the studies focus on dynamic routing problems and consider locations of gathering points as given for collecting vehicles or rebalancing Luo et al. (2022). In Mahmoodian et al. (2022) a dynamic hubbing strategy was addressed to satisfy the first demand of the following day and improve the efficiency of the shared bike rebalancing scheme. However, their main focus is on the rebalancing plan for simulated hubs. A unified overnight charging and vehicle rebalancing approach is suggested by Osorio et al. (2021) to carry out these tasks for a shared e-scooter system more effectively. They provide a charging-and-routing plan that concurrently determines an effective pickup and drop-off strategy while taking into account the SOC of all onboard e-scooters. Although they recommend breaking up a large zone into smaller zones in order to lessen the computational complexity for future studies, it has neither been implemented nor dynamic data-driven zoning discussed.

One potential approach to defining and addressing such problems is through the utilization of time-space network modeling techniques, as demonstrated in the context of logistic networks by Akyüz et al. (2023). However, it should be noted that this approach can be computationally intensive and time-consuming. As an alternative, dynamic programming modeling may offer significant advantages in terms of optimization performance, as demonstrated in the work of Al-Kanj et al. (2020). Unlike the current literature, we consider a dynamic and data-driven approach to decide on consolidating the low-charge e-scooters and in the second layer of the network charging them with a moving charger on spot. Specifically, in our proposed method as soon as the first zone is selected for the first low-charge e-scooter to drop off, this zone becomes the point of interest for

all future e-scooters whose charge drops in this zone and its neighboring zones. While considering the service level indicator for the rider which is the distance deviation that the selected low-charge drop-off zone causes from the rider's destination.

2 METHODOLOGY

To investigate the designated low-charge drop-off areas, a dynamic programming model based on the destination and SOC of the low-charge e-scooter is designed. We considered real-world practices for e-scooters that become unavailable after their battery level dips below 18% (Felyx.com). Therefore, we set a robust battery level threshold of 25% to define the low-charge e-scooter. The decision epochs or time steps are modeled in discrete time $t = \{0, 1, 2, \dots, 32\}$ with 15 minutes duration over 8 working hours. If the decision epoch is at time t , then all information arriving between $t - 1$ and t is collected and assigned at time t . $t = 0$ indicates the beginning of the operation. An e-scooter is shown with i index where $i \in \{1, \dots, N\}$ and N refers to the total number of low-charge e-scooters. To improve the computational complexity of the model, the operating area is discretized into several same-size hexagons. Each hexagon is considered as a potential low-charge drop-off location. The zones are shown j where $j \in \{1, \dots, Z\}$ and Z represents the total number of hexagons.

Using the language of dynamic resource management, e-scooters, and potential low-charge drop-off locations are resources (R). The physical system state vector is then given by $S_t^R = (R_{ti}, R_{tj})$, in which R_{ti} and R_{tj} indicate the states of e-scooters and potential low-charge drop-off areas at time t , respectively. The state vector of e-scooters over time is defined below.

$$R_{ti} = \begin{pmatrix} i_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} \text{current location of e-scooter} \\ \text{SOC} \end{pmatrix}$$

Where the current location of the e-scooter equals to user's destination and corresponds to the latitude and longitude coordinates of where the low-charge e-scooter is located at time t and searching for a low-charge drop-off location. Also, the state vector of potential low-charge drop-off locations changes over time in terms of the number of low-charge e-scooters parked at the low-charge drop-off location and is defined as following

$$R_{tj} = \begin{pmatrix} j_1 \\ j_2 \end{pmatrix} = \begin{pmatrix} \text{zone index} \\ \text{sum of low-charge e-scooters} \\ \text{parked in the zone till time } t \end{pmatrix}$$

The action of the dynamic model is a binary variable indicating whether the potential low-charge drop-off location with attribute R_{tj} should be offered to a low-charge e-scooter with attribute vector R_{ti} at time t .

$$x_{ij}^t = \begin{cases} 1 & \text{if e-scooter } i \text{ is sent to low-charge drop-off location } j \text{ at time } t \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The decision variable x_{ij}^t must satisfy the following constraints:

$$\sum_j^Z x_{ij}^t = 1 \quad \forall t, i \quad (2)$$

$$\sum_i^N \sum_t^T x_{ij'}^t \leq \sum_i^N \sum_t^T M(1 - x_{ij}^t) \quad \forall (j, j') \in \text{adjacent zones} \quad (3)$$

$$x_{ij}^t = \{0, 1\} \quad \forall t, i, j \quad (4)$$

Equation (2) indicates that each low-charge e-scooter should only be assigned to one zone at each time step. In equation 3, we define the neighboring zone context. In zoning the operating area, we obtained adjacent sets for each zone. If the grid distance between zones is less than one grid, those zones are instant neighbors (they have common boundaries) and are in the same adjacent set.

Equation 3 states that if zone j becomes a drop-off location, then all of the zones in its adjacent set cannot be an option (M refers to a big number). Equation (4) shows the variable domain. In addition, we consider the feasibility of trips from the final destination of e-scooters and selected low-charge drop-off locations based on the distance between these two locations at time t (d_{ij}^t). If this distance equals traveling more than three zones the trip becomes infeasible. Similarly, the insufficiency of the SOC to perform the trip makes the trip infeasible.

Transition function

The transition function represents how the system evolves over time. We model the transition function deterministically. E-scooters' current location or user's destination can be obtained through GPS information. However, in this study, we simulate the OD trips based. The SOC of e-scooters is calculated using the OD matrix by eq. (5), under the assumption that each e-scooter at most can perform one trip during each time step.

$$\text{SOC}_i^t = \text{SOC}_i^{t-1} - (\delta * \text{travelled distance of e-scooter } i \text{ between time } (t-1) \text{ and } t) \quad (5)$$

where δ indicates the rate of discharge per kilometer. The value of this parameter is calculated by considering the e-scooter battery Volt, the maximum distance, and the maximum speed it can perform. Regarding changes in zone states, the attribute of the drop-off zones will change as the number of low-charge e-scooters parked in the zone till time t (i.e. j_2) changes. $\sum_{t=1}^{t-1} q_t(j)$ captures the number of parked low-charge e-scooters at zone j up to time t . Equation (6) specifies that at the beginning of the planning horizon, there is no low-charge e-scooter parked in the system, and equation 7 calculates the number of parked low-charge e-scooters at time t in zone j based on the state of the system in the previous time step ($t-1$), it shows the summation of all parked low-charge e-scooters till $t-1$ and the number of e-scooters that are sent to the low-charge drop-off location at time t . This parameter is a post-decision parameter and will be updated after the decision is made at time t . The model uses the value of the parameter for the previous time step to make a decision at time t .

$$q_j^t = 0 \quad t = 0, \forall j \quad (6)$$

$$q_j^t = q_j^{t-1} + \sum_i^N x_{ij}^t \quad \forall t > 0, j \quad (7)$$

When the e-scooter is not low-charge, its SOC will be updated through equation 5 based on its trip between time $t-1$ and t . Therefore, We denote the transition function by

$$S_t = S^M(S_{t-1}, X_t) \quad (8)$$

where $S^M()$ governs the transition from pre-decision state S_{t-1} to pre-decision state S_t , and X_t is a decision vector containing x_{ij}^t decisions. $S^M()$ is a general statement for all of the transition functions we mention in this section.

Objective function

Each decision in the system produces a contribution c_{ij}^t to the system, especially to the routing costs of the charging truck. Collecting low-charge e-scooters in designated areas within the operating zone intends to save the system's routing costs for recharging operations. Hence, the objective function of the model aims at assigning low-charge e-scooter i to the nearest feasible low-charge drop-off zone j , if there is any. We aim at maximizing the reward that might obtain by the consolidation of low-charge e-scooters by equation 9.

$$c_{ij}^t = \begin{cases} \sum_{t=0} q_j^{t-1} - d_{ij}^t & \text{if } x_{ij}^t = 1 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

Where the contribution is defined based on the number of parked low-charge e-scooters till time $t-1$ at location j (q_j^{t-1}) and the service level indicator which assures that the distance between

the low-charge drop-off location j and the final destination of the e-scooter i trip should be small (d_{ij}^t). We assume linear contribution, as shown in equation 10.

$$C_t(S_t, x_t) = \sum_i^N \sum_j^M c_{ij}^t x_{ij}^t \quad (10)$$

Policy maps a state to an action/ decision. The optimal policy ($\pi \in \Pi$) maximizes the sum of contributions over all time periods as shown in equation 11.

$$F_t^*(S_t) = \max_{\pi \in \Pi} \left(\sum_{t=0}^T C(S_t, X_t^\pi(S_t)) | S_{t-1} \right) \quad (11)$$

where $X_t^\pi(S_t)$ is a function that determines x_{ij}^t given S_t when we are taking policy π , and Π is a set of decision functions or policies. To find the best policy, we intend to maximize the reward function subject to the above-mentioned constraints.

3 RESULTS AND DISCUSSION

We used simulated data of 30 e-scooters operating on the campus of TU Delft in one working day (8 hours), which was divided into 32 time slots as we considered every 15 minutes as a time step, to solve a toy problem as a proof of concept for our model. The rate of discharge is set at 8% per km. At the start of the planning period, it is expected that e-scooters are fully charged. Each of the 37 regular hexagons, each with a side length of 60 m, that make up the operating area can serve as a low-charge drop-off location.

The outcomes of using the model on this set of data are shown in Table 1, where 12 low-charge e-scooters emerged in 12 separate zones. The first column indicates the time steps in which low-charge e-scooters show up. The second column contains e-scooters' specific ids. The third column shows the final destination of e-scooters trips. The term 'outside' in this column specified that the destination of the e-scooter trip is outside of the operating zone. Therefore, it should be left within the operating area. Column 5 indicates the final selected consolidated low-charge drop-off zones. Also, the SOC of low-charge e-scooter before and after their assignment to the drop-off zone are shown in columns 4 and 6, respectively.

The optimality gap of the obtained results is 0%, which tells at each time step, the model finds the optimal solution. At the end of the planning horizon, selected zones are zone numbers 14, 20, and 3 with 4, 5, and 3 low-charge e-scooter parked in these zones, respectively. The mean occurred deviation from the destination is 114 m, and 41 m and 194 m are the minimum and maximum values of the deviations from the riders' final destinations, respectively.

The model will determine the initial e-scooter(s) that showed up as low-charge, calculate the distance between the location of the e-scooter and its surrounding zones, and then provide the rider with a drop-off zone that is close to her destination while also following the neighboring zone constraint equation 3. For instance, in Table 1, the first low-charge e-scooters appeared at time step 19, in zones 9 and 27. If allowed to be left at the destinations, these zones have 4 neighboring zones in common (i.e., {10, 12, 29, 30}). Also, the required travel distances for e-scooters to reach the centers of the destination zones are 40 m, and 50 m for vehicles 3 and 25, respectively. Whereas, selected drop-off locations i.e. zones 14 and 20 have fewer common neighboring zones (i.e., {24, 29, 30}) and the distances between the locations of e-scooters and the centers of the selected zones are not much larger (70 m for both e-scooters). Therefore, zones 14 and 20 are selected as the first low-charge drop-off zones, and the number of parked low-charge e-scooters in these zones will increase, which means they become more relevant to be a drop-off zone for the e-scooters appear in the next time steps. Using the same logic, the model sends the e-scooter to zone 14 at time step 20.

At time 21, one low-charge e-scooter appears in zone 3, which is farther than the permitted distance (service level assurance) from the selected drop-off zones 14 and 27, and {31, 15, 28, 23, 19} zones are in a neighboring distance of the selected drop-off locations so they can not be a potential drop-off location, then it is allowed to be left at its destination (as its distance from potential zone 6, 19, 21, 13, and 18 are greater than the distance to the centroid of zone 3). For the rest of the

time steps, the low-charge e-scooters are consolidated in these three drop-off locations following the same reasoning. It is worth noting that, we also assure that e-scooters are able to perform the trip to low-charge drop-off zone with sufficient SOC. Figure 1a shows the zoning structure of the operating area. In Figure 1b the selected low-charge drop-off zones are shown in green circles, and the zones which are in adjacent sets of the selected drop-off zones are shown in black color. Zones in blue color are neither a low-charge e-scooter destination nor under the coverage area of the selected zones.

Regarding recharging operations, in the upper level of the recharging operation problem, a routing problem arises, where the charging truck’s traveling distance needs to be minimized. The routing problem involves determining the optimal path that the charging truck should take to recharge the e-scooters, taking into account the locations of the consolidated low-charge e-scooters and the charging truck’s constraints. The objective of this routing problem is to minimize the total traveling distance of the charging truck while satisfying the demand for e-scooter charging. By applying this method, we are shrinking the size of the network for the charging truck by decreasing the visiting nodes. In the proposed example, the charging truck has 3 nodes to visit to recharge the low-charge e-scooters instead of visiting them scattered in 12 nodes. Then, the upper level of the problem, which is the routing problem of the recharging truck and will be studied in a further step of this research.

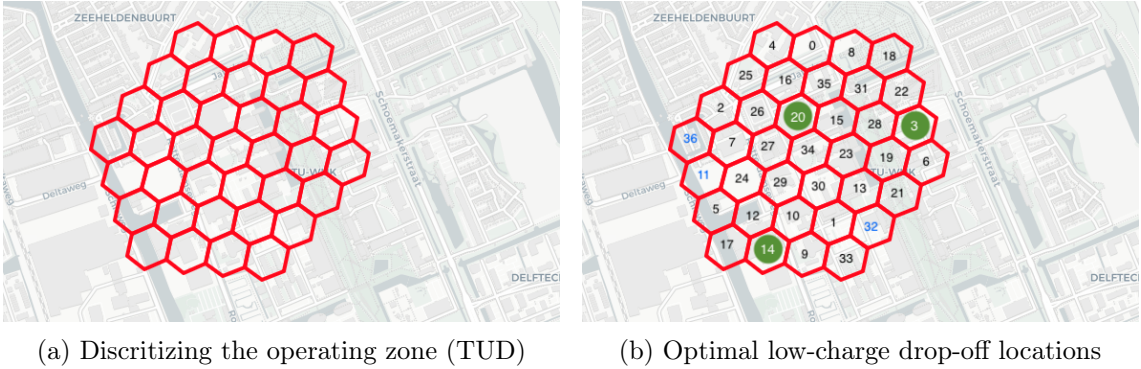


Figure 1: An overview of zoning and optimal low-charge drop-off location

4 CONCLUSIONS

This study addressed one of the operational challenges faced by shared electric scooters (e-scooter). In this study, we propose a data-driven approach to dynamically define designated drop-off areas for low-charge e-scooters by capturing the spatial and temporal characteristics of the e-scooters and potential drop-off zones. A simulated data set of 30 e-scooters in TU Delft campus area is used. The results suggest that the low-charge e-scooters can be consolidated in three drop-off locations among 37 potential zones over 8 working hours. By adopting this approach, the average deviation in distance of e-scooters from their intended endpoint is found to be 114 meters, a distance that is generally considered to be within the range of what is typically considered suitable for walking distance.

One of the significant advantages of our approach is the optimization of spatial usage by e-scooters, which is particularly crucial for low-charge e-scooters. Furthermore, our model ensures that the travel of the e-scooter to the suggested drop-off zone is feasible in terms of the deviation from the destination and its state of charge after performing the trip. Moreover, defining consolidated low-charge drop-off zones out of a discrete space of operating area can have a considerable impact on minimizing the traveling distance of the charging truck and decreasing the computational complexity of solving the model in continuous space. Overall, the approach of limiting visiting nodes to recharge low-charge e-scooters through consolidating can lead to a reduction in operational costs and improved system performance in shared micromobility systems. The subsequent routing problem of the charging truck can further optimize the recharging operations and enhance the overall efficiency of the system.

Table 1: Overview of the studied sample and selected drop-off zones

Time step	e-scooter id	low-charge e-scooter location	The SOC before trip to drop-off zone (%)	Selected drop-off zone	The SOC after trip to drop-off zone (%)
19	3	9	24.4	14	23.2
	25	27	24.0	20	22.9
20	16	12	24.5	14	22.8
21	6	3	24.4	3	23.9
22	23	outside	24.4	14	21.4
23	0	9	24.4	14	21.3
24	10	outside	24.6	3	22.3
	19	20	24.8	20	24.4
	22	7	24.3	20	21.3
25	20	23	24.8	20	22.9
27	29	0	24.8	20	22.9
30	17	19	24.4	3	23.4

REFERENCES

- Akyüz, M. H., Dekker, R., & Azadeh, S. S. (2023). Partial and complete replanning of an intermodal logistic system under disruptions. *Transportation Research Part E: Logistics and Transportation Review*, 169, 102968.
- Al-Kanj, L., Nascimento, J., & Powell, W. B. (2020). Approximate dynamic programming for planning a ride-hailing system using autonomous fleets of electric vehicles. *European Journal of Operational Research*, 284(3), 1088–1106.
- Luo, X., Li, L., Zhao, L., & Lin, J. (2022). Dynamic intra-cell repositioning in free-floating bike-sharing systems using approximate dynamic programming. *Transportation Science*, 56(4), 799–826.
- Mahmoodian, V., Zhang, Y., & Charkhgard, H. (2022). Hybrid rebalancing with dynamic hubbing for free-floating bike sharing systems. *International Journal of Transportation Science and Technology*, 11(3), 636–652.
- Osorio, J., Lei, C., & Ouyang, Y. (2021). Optimal rebalancing and on-board charging of shared electric scooters. *Transportation Research Part B: Methodological*, 147, 197–219.

The impact of social networks and coordinated destination choice on the spread of epidemics using Episim

Joanna Ji^{*1}, Qin Zhang², Dr. Ana Tsui Moreno³, and Dr. Rolf Moeckel⁴

^{1,2,3,4}School of Engineering and Design, Technical University of Munich, Germany

SHORT SUMMARY

Person-to-person contact is fundamental to the spread of epidemics. Human mobility is important for understanding the pattern of person-to-person contact. Therefore, understanding human travel behavior is crucial to the understanding of the geographic spread of infectious diseases. Transportation models built to give details on human movement and contact can then be used to simulate epidemic spread. Such has been done in the epidemic model Episim. In addition, there is also research showing people's social networks have strong influence on their travel behavior, such as destination choice. Therefore this study sets out to verify whether adding social network and coordinated destination choice to the epidemic model has impact on the spatio-temporal transmission progression of epidemics. Results show that though the total number of infections do not change, the addition of a social network and coordinated destination help capture during what kind of activities infection events are taking place. Coordinated travel with social networks contribute to a more rapid spread in the beginning. Moreover, results emphasize that social networks should be integrated in conjuncture with joint-travel to better capture social travel behavior and in turn epidemic spread.

Keywords: agent-based modeling, epidemic modeling, social networks, transportation network modeling

1 INTRODUCTION

Human mobility, like many other aspects of human behavior, is socially driven. In recent years, the transportation modeling field has increasingly tried to capture the undeniable link between social connections and travel behavior Kim et al. (2018). Various elements of travel behavior, including the creation of travel activities, choice of destination and mode of transportation, amongst others, are affected by social connections (Kim et al., 2018). Person-to-person contact is crucial to epidemic spread models. Since social networks can strongly influence travel behavior and contact patterns, it is beneficial to add in social networks when using travel demand models to simulate disease spread.

Previous researchers have used agent-based epidemic models to better understand and predict the spread of infectious diseases in Singapore (Sun et al., 2014; Mo et al., 2021), the Twin Cities (Bóta et al., 2017; Hajdu et al., 2020) or Berlin (Müller et al., 2020). Epidemic transition models varied from the simplest Susceptible-Infected model (SI), in which agents are exposed only while traveling (Hajdu et al., 2020), to the more complex Susceptible-Exposed-Infections-Removed (SEIR) model with exposure while traveling and at destinations (Müller et al., 2020; Mo et al., 2021). On the other hand, travel demand generation was rather simplistic, as it replicated transit trips (Bóta et al., 2017; Hajdu et al., 2020) or smart-card bus trips (Sun et al., 2014; Mo et al., 2021). A notable exception is Müller et al. (2020), who replicated activity-based trajectories from mobile phone data and used traffic and transit assignment in MATSim. Furthermore, Müller et al. (2020) validated the results of Episim with COVID-19 hospital cases in Berlin.

Social connections could potentially affect many aspects of travel behavior and by extension travel demand modeling. Socially connected people may share destinations, which has been seen in the studies by van den Berg et al. (2010) and Moore et al. (2013). Moreover, Ronald et al. (2012), Ma et al. (2011) and Dubernet & Axhausen (2015) presented an agent based system which integrate joint decision-making mechanisms based on rule based simulations of a bargaining process, but none have integrated and implemented a social network in a complete travel demand model framework.

In addition, none of the studies have explicitly considered the impact of social networks in epidemic

spread. This research seeks to fill the gap by adding in a synthesized social network to the process. As a first step, we focus on adding coordinated destination in travel demand generation, and infection rules influenced by social connections in Episim to see the impact of social network on infection spread patterns and rates.

2 METHODOLOGY

In order to simulate the epidemic spread, we generate travel demand with the open-source microscopic transportation orchestrator (MITO) (Moeckel et al., 2020), which we later assign to the road and transit networks using the Multi-Agent Transport Simulation (MATSim) (Horni et al., 2016). The event files from MATSim are later fed into Episim (Müller et al., 2020) that runs for an entire year.

The base input data for any agent-based model is the synthetic population, specifically, the synthetic population for the Munich metropolitan area. It comprises 4.5 million persons in 2.3 million households, but lacks social networks and group quarters (Moreno & Moeckel, 2018). For this research, a social network and nursing homes are implemented. The social network is built for the Munich metropolitan area. We also utilize the social network in conjunction with rule-based coordinated destination in the travel demand generation to account for joint destination.

This section summarizes the generation of the social network, the incorporation of the social network into MITO for coordinated destination, and the incorporation of the social network into Episim.

Geolocation-based social network

We generate a geolocation-based, ego-centric social network for the study area. They are so-called personal social networks or egocentric networks, consisting of an ego or focus point, connected to alters, e.g. family members, co-workers, etc. Due to the lack of relevant collected data on social network structures in the area, the network is currently built on characteristics present in the synthetic population. As such, the generated network reflects a degree of homophily and reciprocity.

Homophily describes how like attracts like, that people tend to be socially connected with those similar to them. Reciprocity is how social connections are bi-directional. If A is friends with B, then B is friends with A. However, it does not take into account the transitivity property of social networks, which means that if A is a friend of B, and B a friend of C, then A would also have a higher chance of being friends with C.

Type of social tie	Edge build criterion	Average degree
Household	From same household	2.2
Neighborhood	Share dwelling location and dwelling type	2.5
Education	Attend same school, same age	9.5
Work	Share job location and job type	5.5
Nursing home	Share nursing home location	10

Figure 1: Network type and build criterion

The social network is currently built on five kinds of relationships, household, neighborhood, education, work and nursing home. Those who reside in the same household are connected to each other. Those who share the same dwelling location and a similar dwelling type, such as single-family-unit, are socially connected. Those who attend the same school and are of the same age are presumed to be socially connected. Those who share a job location and job type have a chance of being socially connected. Residents in the same nursing home also forge social connections with each other. For neighborhood, education, work and nursing home locations, social connections are built using a small-world network algorithm with a maximum clique size of 10. Social connections forged from household, neighborhood, education or work categories are built with the criteria indicated in Figure 1. Figure 1 also shows the average degree forged in each type of social tie using these criteria. For example, a person attending school is on average connected to 9.5 others who attend the school with them. A person’s total social connections comprise of connections from all five of these relationship types.

Travel demand model and rule-based coordinated destination

The second part of this study uses an agent-based travel demand modeling suite to generate daily movements of individuals. To achieve this, the synthetic population with the geolocation-based social network were fed into MITO (Moeckel et al., 2020). MITO is an agent-based travel demand model that uses econometric statistical models to estimate trip generation, trip distribution, mode choice, and time of day for each individual in the synthetic population. After that, MATSim (Horni et al., 2016) is used as a dynamic traffic assignment model to assign car trips to the road network and simulate public transport trips on the transit system (Swiss Federal Railways, 2020). The individual movements are estimated for a typical 24-hour day for the following trip purposes: home-based work (HBW), home-based education (HBE), home-based shop (HBS), home-based recreation (HBR), home-based other (HBO), non home-based work (NHBW) and non home-based other (NHBO).

This study extended MITO by adding the rule-based coordinated destination choice. In Figure 2 the modifications to the existing MITO model sequence is shown. After the Trip Distribution step, we add an Arrival Time Choice in which we compare the trip list of an agent with the trip lists of those in their social network. Compatible trips are defined as trips that have arrival times within six hours of each other, and are of the same purpose. We also give a hierarchy to coordinated trips depending on social network type. Agents prioritize coordinating with household members, then with coworkers or schoolmates and lastly with those from the ‘neighbor’ social connection type. If compatible trips are found, we then proceed to the next step, Destination Coordination. In this step, the destination and arrival time of the trip belonging to the agent’s social connection is set to be the same as the agent’s.

After any changes of destination have been made to the trip, the mode choice is ran without any modifications. The departure time can then be calculated based on the chosen mode and projected travel time.

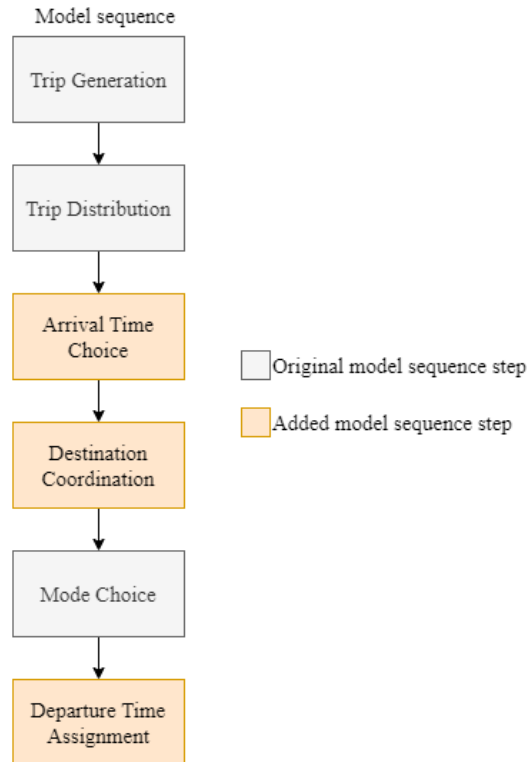


Figure 2: MITO modified model sequence

Epidemic spread model Episim

Episim is an infection dynamics model build on top of a person’s movement trajectories as developed by the Sebastian Müller et al. at the Technical University of Berlin (Müller et al., 2020). It allows testing of intervention policies such as home-office mandates, mask-wearing mandates, etc.

Episim is comprised of several models, the contact model, infection model, and disease progression model (Müller et al., 2020). The contact model defines who comes into contact with whom. Persons at the same location or facility can come into contact and infect each other. These facilities could either be in transit or at an activity location such as home or work. When two persons come into contact, a probability of infection is calculated using the infection model, which gives this probability based on contact intensity, contact duration, viral shedding and intake. If a person becomes infected, the disease progression model then gives the probability that this newly exposed and infected person progresses to the next stage of the disease. The exposed person can become infectious, recover, or get worse. The simulation runs for a year or until no more infections occur. For more details on Episim, please refer to (Müller et al., 2020).

To see the effect of social networks on epidemic spread, we extended the Episim model to account for social networks. Firstly, the contact model was modified. This contact model looks at agents when they leave a facility. Instead of randomly selecting other agents who are at the facility at the same time, we increased the likelihood of being selected if they belong to the same social network. Secondly, we changed the infection model parameters to consider that persons in the same social network may reduce social distancing, and therefore, the viral load may be increased. This is accomplished by increasing the contact intensity factor. For contacts from within the agent’s social network, the contact intensity is multiplied by a factor of 10. Short of observed data, this factor is an exogenous assumption that ensures a higher infection rate within social networks. The current contact intensities and infection probabilities are set according to those specified in (Müller et al., 2020) for the COVID 19 virus.

Scenarios

By varying the addition of social network and coordinated destination choice, we look at a total of four scenarios, as seen in Figure 3.

Episim rules MITO rules	Randomly selected maximum of three contacts	Selection priority given to social network, maximum three contacts
No coordinated destination	Base	Base.SocialNetwork
With coordinated destination	Coordinated	Coordinated.SocialNetwork

Figure 3: Scenario description

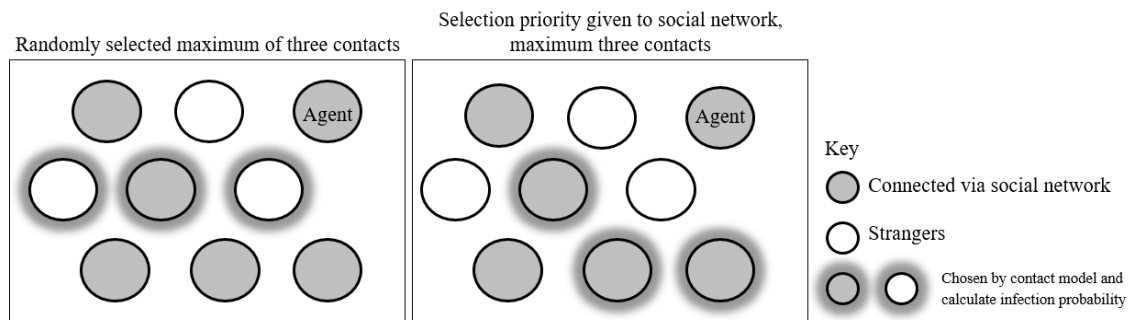


Figure 4: Episim contact rules for different scenarios regarding social network

In the *Base* scenario, for an agent at a certain location/facility, a maximum of three contacts are randomly selected from those who are at the same location and the same time as the agent. This number is the base assumed setting in the calibrated Episim Berlin scenario (Müller et al., 2020). The infection probability is calculated between the agent and each of these three contacts. This is graphically represented in Figure 4, in the left box.

In the *Base.SocialNetwork* scenario, similarly, a maximum of three contacts are selected, but selection priority is given to those within the agent’s social network. If there are less than three contacts at the location in the agent’s network, then random agents are selected until there are three contacts. This is seen in the right box in Figure 4.

The above scenarios are then repeated, but with a travel demand that now takes into account coordinated destination choice based the social network. To reduce model runtime, all scenarios are ran for a 5% scaled-down population for computational time savings. The social network is generated after scaling down. Episim results are reported after upscaling factors to 100%.

3 RESULTS AND DISCUSSION

Epidemic spread with and without social networks, with and without coordinated destination are compared. The main hypothesis is that the total number of infected persons would not vary, as agents perform the same number of activities. But the spatial and temporal distribution is expected to be affected by the presence of social networks and coordinated destination choice. For example, we would be able to better capture the outbreak in nursing homes or large employment centers; reducing contact at such hotspots may be more sensitive to interventions than limiting social contacts in general.

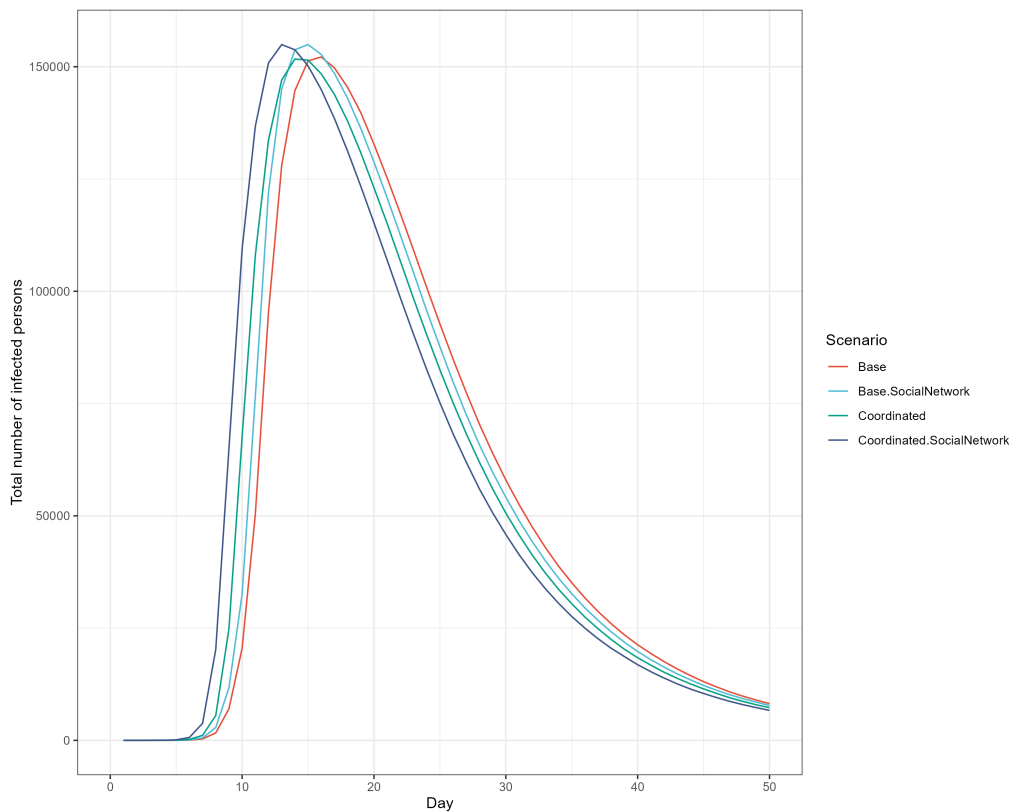


Figure 5: Number of infections from day 1 to day 50

In the Figure 5, the Episim total infected curve for the Munich region for each of the four scenarios is shown. As expected, the total number of infections do not vary much between each scenario. The trajectory and the peak of the graph are only slightly staggered. We zoom in on the first 14 days of the epidemic outbreak in Figure 6.

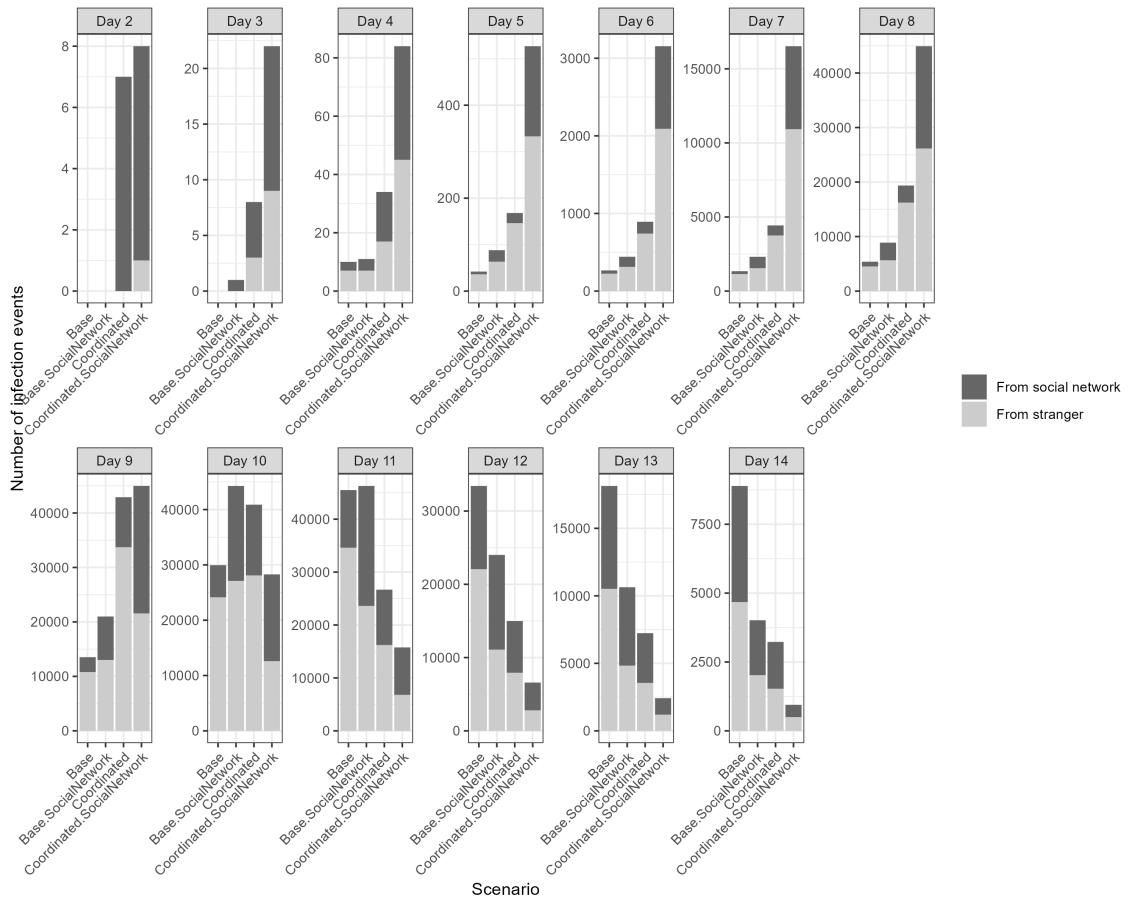


Figure 6: Daily number of infections

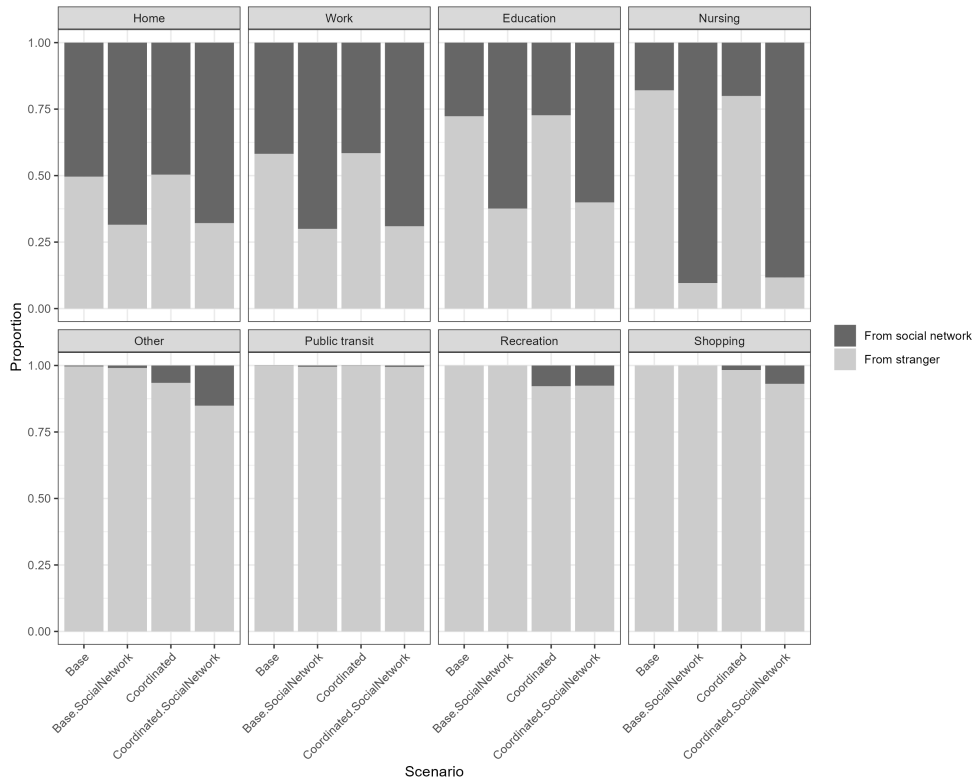


Figure 7: Number of infections by infection location and social network status

Figure 6 shows the day-by-day number of infection events for each scenario. In addition, the figure shows the proportion of infection events between socially connected agents (dark gray) and between strangers (light gray). At this more micro-temporal scale, we see that coordinated destination scenarios have a quicker start compared to the base scenarios, with infections in Coordinated.SocialNetwork scenario spreading most rapidly. Within the Base and Coordinated scenarios, the scenario with social-network contact rules also spread faster than scenarios without. This figure also shows that the early disease transmissions tend to be from social contacts. Once the epidemic is more wide spread, infection events between strangers begin to make up the larger proportion. Infections from the social network scenarios with social network contact rules have a higher share of being infected by someone within their social network.

We then break down the infection events by activity type. Figure 7 shows the proportion of infection events from the social network and from strangers. The proportion of infection from social network ties increased for Coordinated Destination scenarios compared to the Base scenarios in the Other, Recreation and Shopping activities. The addition of coordinated destination choice captures how non-essential leisure activities may be conducted together with friends, family and acquaintances, and captures the disease transmission that may happen as a result. Without coordinated destination, agents would seldom meet others in their social network, whereas activities like home, work, education and being in a nursing home usually guarantees that agents are in the same location as someone from their social network.

Figure 8 shows the percentage of infection from social network contacts per total number of infections. The addition of social networks has increased the share of infections that come from social network ties. The Coordinated.SocialNetwork scenario has the highest proportion of social network related infection events.

Scenario	Base		Coordinated	
	Random contacts rule	Social network contact rule	Random contacts rule	Social network contact rule
Home	23.3%	33.3%	22.8%	32.2%
Work	4.0%	7.2%	3.9%	6.9%
Education	1.6%	3.3%	1.6%	3.2%
Nursing	0.2%	0.8%	0.2%	0.8%
Other	0.1%	0.3%	2.3%	5.1%
Public transit	0.0%	0.0%	0.0%	0.0%
Recreation	0.0%	0.0%	0.1%	0.1%
Shopping	0.0%	0.0%	0.0%	0.0%
Total	29.2%	44.9%	30.9%	48.3%

Figure 8: Percentage of infection from social network contacts per total number of infections

The percentage of infection from social network contacts per number of infections varies by activity purpose. For example, in Base.SocialNetwork and Coordinated.SocialNetwork scenarios, social network in nursing homes only account for 0.75% of total infections, because the population of nursing homes is rather small. However, social network ties account for a big proportion of infections occurring in nursing homes, as seen in Figure 6.

4 CONCLUSIONS

We combined a simple synthesized social network with an agent-based travel demand model and epidemic spread model to see possible effects on epidemic spread patterns. Our social network and coordinated travel, though simplistic, demonstrate that social networks have some influence on disease spread patterns. It can affect how fast the disease spreads, and where disease is spread.

Future research should focus on implementing more comprehensive joint travel logic based on social network connections in the travel demand model, as the scenarios with coordinated travel showed a marked effect on epidemic spread compared to those without. Another angle for refinement is the social network. Currently the social network is based on shared home, neighborhood, work or education locations in the synthetic population. The next iteration can include social connections outside of household, neighborhood, education and work. These connections can reflect general

friendship and social ties. In addition to homophily and reciprocity, this social network can incorporate the transitivity properties of social connections. Nevertheless, this research presents a novel coupling of synthesized social networks, travel demand modeling and epidemic spread modeling. It demonstrates a way to model human connections in human movement, and how an epidemic travels through the human network.

ACKNOWLEDGEMENTS

The research was funded by the German Research Foundation (DFG) under the research project TENGOS: "Transport and Epidemic Networks: Graphs, Optimization and Simulation" (Project number 458548755).

REFERENCES

- Bóta, A., Gardner, L. M., & Khani, A. (2017). Identifying Critical Components of a Public Transit System for Outbreak Control. *Networks and Spatial Economics*, *17*(4), 1137–1159. doi: 10.1007/s11067-017-9361-2
- Dubernet, T., & Axhausen, K. W. (2015, 9). Implementing a household joint activity-travel multi-agent simulation tool: first results. *Transportation*, *42*, 753-769. doi: 10.1007/s11116-015-9645-7
- Hajdu, L., Bóta, A., Krész, M., Khani, A., & Gardner, L. M. (2020). Discovering the Hidden Community Structure of Public Transportation Networks. *Networks and Spatial Economics*, *20*(1), 209–231. doi: 10.1007/s11067-019-09476-3
- Horni, A., Nagel, K., & Axhausen, K. W. (Eds.). (2016). *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press. Retrieved from <http://www.ubiquitypress.com/site/books/10.5334/baw/> doi: 10.5334/baw
- Kim, J., Rasouli, S., & Timmermans, H. J. (2018, 7). Social networks, social influence and activity-travel behaviour: a review of models and empirical evidence. *Transport Reviews*, *38*, 499-523. doi: 10.1080/01441647.2017.1351500
- Ma, H., Ronald, N., Arentze, T. A., & Timmermans, H. J. (2011, 12). New credit mechanism for semicooperative agent-mediated joint activity-travel scheduling: Negotiating with incomplete information. *Transportation Research Record*, 104-110. doi: 10.3141/2230-12
- Mo, B., Feng, K., Shen, Y., Tam, C., Li, D., Yin, Y., & Zhao, J. (2021). Modeling epidemic spreading through public transit using time-varying encounter network. *Transportation Research Part C: Emerging Technologies*, *122*(January). doi: 10.1016/j.trc.2020.102893
- Moeckel, R., Kuehnel, N., Llorca, C., Moreno, A. T., & Rayaprolu, H. (2020). Agent-based simulation to improve policy sensitivity of trip-based models. *Journal of Advanced Transportation*. doi: 10.1155/2020/1902162
- Moore, J., Carrasco, J. A., & Tudela, A. (2013, 7). Exploring the links between personal networks, time use, and the spatial distribution of social contacts. *Transportation*, *40*, 773-788. doi: 10.1007/s11116-013-9467-4
- Moreno, A., & Moeckel, R. (2018, May). Population Synthesis Handling Three Geographical Resolutions. *ISPRS International Journal of Geo-Information*, *7*(5), 174. Retrieved from <http://www.mdpi.com/2220-9964/7/5/174> doi: 10.3390/ijgi7050174
- Müller, S. A., Balmer, M., Charlton, B., Ewert, R., Neumann, A., Rakow, C., ... Nagel, K. (2020). Using mobile phone data for epidemiological simulations of lockdowns: government interventions, behavioral changes, and resulting changes of reinfections. *medRxiv*, *2020.07.22*. Retrieved from <http://medrxiv.org/content/early/2020/07/24/2020.07.22.20160093.abstract> doi: <https://doi.org/10.1101/2020.07.22.20160093>
- Ronald, N., Arentze, T., & Timmermans, H. (2012). Modeling social interactions between individuals for joint activity scheduling. *Transportation Research Part B: Methodological*, *46*, 276-290. doi: 10.1016/j.trb.2011.10.003

- Sun, L., Axhausen, K. W., Lee, D. H., & Cebrian, M. (2014). Efficient detection of contagious outbreaks in massive metropolitan encounter networks. *Scientific Reports*, *4*, 1–6. doi: 10.1038/srep05099
- Swiss Federal Railways. (2020). *Matsim-extensions by sbb*. Retrieved 21.07.2021, from <https://github.com/SchweizerischeBundesbahnen/matsim-sbb-extensions>
- van den Berg, P., Arentze, T., & Timmermans, H. (2010). Location-type choice for face-to-face social activities and its effect on travel behavior. *Environment and Planning B: Planning and Design*, *37*, 1057-1075. doi: 10.1068/b36019

Whose preferences matter more? Handling unbalanced panel data for choice modelling

Laurent Cazor^{*,†}, Mirosława Łukawska[†], Mads Paulsen[†], Thomas Kjær Rasmussen[†], and Otto Anker Nielsen[†]

** Corresponding author (lauca@dtu.dk)*

† Technical University of Denmark, Department of Technology, Management and Economics, Bygningstorvet 116b, 2800 Kgs. Lyngby

February 28, 2023

SHORT SUMMARY

The emergence of GPS-enabled smartphones and crowdsourcing tools are unique opportunities for understanding transport behaviour. However, the datasets they generate are often unbalanced, as individuals may use the service collecting data at different frequencies and periods. This raises important questions: are typical discrete choice models robust to this unbalance? Are model estimates biased towards over-represented individuals?

This paper tackles the issue of handling unbalanced panel datasets for route choice modelling. It first develops a simulation experiment to study to which degree Mixed Logit Models with panel effects reproduce the population preferences using unbalanced data. It then investigates bias reduction strategies, using subsampling and likelihood weighting. These strategies are compared to give guidelines that fit the model purpose. We show that weighting and subsampling techniques can reduce the bias when interpreting the model output for tastes. Combining these techniques helps to find an optimal trade-off between bias and variance of the estimates.

Keywords: Unbalanced panel, panel mixed logit model, subsampling, likelihood weighting, bias-efficiency trade-off

1 INTRODUCTION

An increasing number of large crowd-sourced datasets are available for choice modelling. For instance, smartphone GPS datasets bring researchers new opportunities when analyzing bicycle traffic. They overcome some limitations related to stated preference data or small sample sizes (Nelson et al., 2021; Lee & Sener, 2021). However, due to their crowd-sourced opt-in nature, such datasets may suffer from having a large proportion of the data collected by only a few active users. This may be problematic if the models estimated on these datasets are used for policy implications or forecasting purposes, for instance if the preferences of these active users differ from the population mean.

The issue of repeated observations per individual in the panel setup for the mixed logit model has been addressed in the literature in the context of stated preference (SP) data. Bliemer & Rose (2010) recognized the advantages of panel setup and studied the construction of an optimal experiment design (in terms of the statistical properties of the model) for stated choice surveys with panel information. Rose et al. (2009) found that adding repeated choice observations per individual improves the model accuracy only until a certain point. Including multiple repeated observations from an individual, which are identical in terms of both the set of attributes and the choice outcome, can be used to account for the effect of e.g. habit (Cherchi & Cirillo, 2014) or correlation patterns (Cherchi et al., 2017).

Yáñez et al. (2011) found that the most significant improvement to the model in terms of fit can be attributed to the introduction of panel correlation. Furthermore, including multiple identical observations should not influence the efficiency of the estimated parameters and does not contribute to the improved capability to retrieve the true parameters. The latter can, however, be improved

by introducing weighting since it increases the influence of the fixed part of the utility over the random part.

Two recent studies (van Cranenburgh & Bliemer, 2019; Ortelli et al., 2022) recognize the challenges of estimating models based on rapidly emerging massive data sources. They propose strategies for a dataset size reduction by optimizing multiple criteria, such as model efficiency, estimation bias, out-of-sample performance, computational time, and value of time for relevant parameters. They propose optimizing the mixed logit model setup based on a simpler multinomial model (MNL).

However, these papers do not answer the question: are estimates biased toward individuals contributing more to a crowdsourced dataset? How to deal with a bias-efficiency trade-off? This paper aims to fill this gap, first testing whether discrete choice models estimated with unbalanced data represent the average tastes of the individuals of the sample population and then finding solutions to eliminate the potential sources of bias.

2 METHODS

Estimated model: the Panel Mixed-Logit model (PMXL)

The Mixed Logit model with panel effects, also denoted Panel Mixed-Logit (PMXL), builds on the traditional Multinomial Logit (MNL) model (McFadden et al., 1973; McFadden & Train, 2000; Train, 2009).

The Panel Mixed-Logit model can be derived as follows: a decision maker $n \in \{1, \dots, N\}$ has T_n choice situations $t \in \{1, \dots, T_n\}$. For each situation t , the decision maker n can choose an alternative i from choice set C_{nt} , whose utility U_{nti} can be written as:

$$U_{nti} = V(\beta_n, X_{nti}) + \epsilon_{nti}$$

β_n is a parameter representing the tastes of decision maker n , X_{nti} is a vector of attributes and ϵ_{nti} is a stochastic error term. Under the logit assumption, the ϵ_{nti} 's are independently and identically distributed (iid) according to Gumbel(0,1). The choice probability of alternative i by decision maker n in choice situation t (i.e., the event ($y_{nt} = i$)) is given by:

$$\mathbb{P}(y_{nt} = i | \beta_n, X_{nti}) = \frac{e^{V(\beta_n, X_{nti})}}{\sum_{j \in C_{nt}} e^{V(\beta_n, X_{ntj})}} \quad (1)$$

We assume the utility is linear in parameters, i.e. that $V(\beta_n, X_{nti}) = \beta_n^\top X_{nti}$. The PMXL model assumes that tastes vary across individuals and that these tastes $\beta_n \sim f(\beta | \theta)$ are iid across decision-makers. The PMXL probability of the sequence of choices $\mathbf{i}_n = i_1, \dots, i_{T_n}$ for decision maker n is then given by:

$$P_{\mathbf{i}_n}(\beta) = \int \prod_{t=1}^{T_n} \mathbb{P}(y_{nt} = i_t | \beta, X_{nt}) f(\beta | \theta) d\theta \quad (2)$$

This probability is approximated through simulation, using:

$$\hat{P}_{\mathbf{i}_n}(\beta) = \frac{1}{R} \sum_{r=1}^R \prod_{t=1}^{T_n} \mathbb{P}(y_{nt} = i_t | \beta_r, X_{nt}) \quad (3)$$

Where the β_r are drawn from the distribution $f(\beta | \theta)$. The parameter β is estimated through Maximum Simulated Likelihood Estimation (MLSE). We define the simulated log-likelihood of the observations (\mathbf{y}, \mathbf{X}) as:

$$\text{LL}(\beta | \mathbf{y}, \mathbf{X}) = \sum_{n=1}^N \ln \hat{P}_{\mathbf{i}_n}(\beta) \quad (4)$$

Path-Size correction

Route choice modelling often involves choice sets with overlapping routes, leading to correlated alternatives and violating the independence assumption of irrelevant alternatives (IIA) in the MNL. To account for this, each alternative receives an additional attribute measuring their overlap with other routes of the choice set. For an alternative i of choice situation t or an individual n , the Path-Size correction (Ben-Akiva & Ramming, 1998) is defined as:

$$\text{PS}_{nti} = \sum_{a \in \Gamma_i} \frac{l_a}{L_i} \frac{1}{\sum_{j \in C_{nt}} \delta_{aj}}, \quad (5)$$

where Γ_i is the set of links for alternative i , C_{nt} is the set of alternatives for choice situation t , l_a is the length of link a , L_i is the length of alternative i , and δ_{aj} equals 1 if j includes link a and 0 otherwise. The utility of an alternative U_{nti} can be written as:

$$U_{nti} = V(\beta_n, X_{nti}) + \beta_{PS} \ln \text{PS}_{nti} + \epsilon_{nti} \quad (6)$$

Where β_{PS} is the Path-Size coefficient to estimate.

Model evaluation

To compare model outputs, we need metrics evaluating bias and precision of the estimations. We denote the Maximum Likelihood estimator of a true parameter vector $\beta = (\beta_1 \dots \beta_K)^\top$ of size $(1 \times K)$ as $\hat{\beta} = (\hat{\beta}_1 \dots \hat{\beta}_K)^\top$.

Bias The bias of an estimator is the difference between this estimator's expected value and the true value of the estimated parameter. It is given by $\text{Bias}(\hat{\beta}) = \mathbb{E}(\hat{\beta}) - \beta$. In the case of a multidimensional estimator, we calculate $\|\text{Bias}(\hat{\beta})\|^2 = \sum_{k=1}^K \text{Bias}(\hat{\beta}_k)^2$

D-error The D-error is an efficiency metric commonly used in experimental designs (see Kessels et al. (2006) for an overview). It is defined as the determinant of the AVC matrix, exponentially scaled w.r.t. to the number of parameters. We calculate the AVC matrix $\Omega = \mathbf{H}^{-1}$ as the inverse of the log-likelihood Hessian matrix at the estimates. For $i, j \in \{1, \dots, K\}$ the Hessian matrix coefficients H_{ij} are given by,

$$H_{ij}(\hat{\beta}) = \mathbb{E} \left[\frac{\partial^2 LL(\beta)}{\partial \beta_i \partial \beta_j} \right]_{\beta=\hat{\beta}}$$

Then, the D-error is given by:

$$\text{D-error} = \det \Omega^{1/K} \quad (7)$$

Minimizing the D-error is minimizing variances and covariances of the estimates. Lower D-error values indicate higher efficiency of the estimated parameter results regarding standard errors. However, it does not provide information on the bias of the estimated parameters.

3 SIMULATION EXPERIMENT

In order to test if the Panel Mixed Logit model reproduces the population parameters, we need to conduct a simulation experiment with known true parameters and sample composition. The following simulation mimics a route choice modeling framework. For one of the population parameters (linked to Elevation gain), we will assume that the number of observations in our dataset is correlated to the parameter individual value.

Step 1 - Network and attributes: We design a small network, with $p \in \{1, \dots, P\}$ Origin-Destination (OD) pairs and choice sets \mathcal{C}_p of routes linking them. These routes have k attributes. For an alternative l of a pair p , we can store these attributes in a vector $X_{l,p} \in \mathbb{R}^k$. We call \mathbf{X}_p the matrix storing the attributes of all the alternatives of \mathcal{C}_p .

The network consists of three OD pairs, A-B, B-C, and A-C (see Figure 1), each linked by 9 routes, composed of all the combinations of links going closer to the destination. Each link has

four attributes: Length (associated to parameter β_L), Elevation Gain (β_E), Bicycle Infrastructure (β_I) and Surface type (β_S). A Path-Size correction term (β_{PS}) (see Equation 5) handles the correlation between routes.

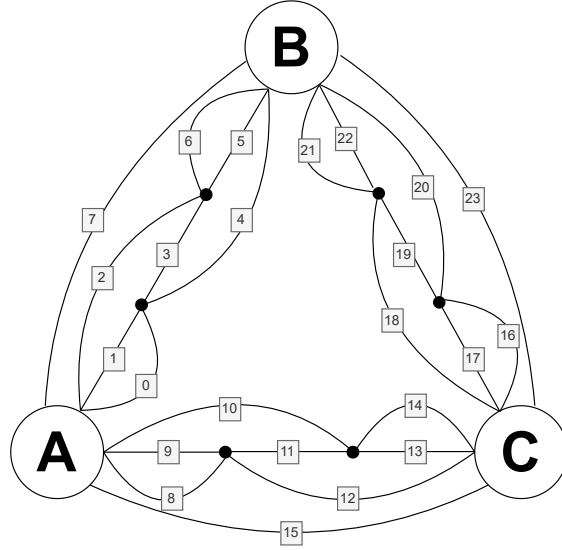


Fig. 1: Network, composed of three ODs and 24 links

Step 2 - Draw population: Assume that the general population follows a multivariate normal distribution for their parameters, i.e., the parameter vector $\beta \sim \mathcal{N}(\beta|\mu, \Sigma)$. $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_k^2)$.

In our simulation, we assume the true parameters values given in Table 1, and that $\sigma_S = \sigma_L = \sigma_{PS} = 0$.

We draw samples of $N = 100$ individuals. The distributions and histograms for the random parameters are plotted below (see Figure 2). We write $\beta_n = (\beta_{L,n} \ \beta_{E,n} \ \beta_{I,n} \ \beta_{S,n} \ \beta_{PS,n})^\top$ the individual parameter drawn for individual n .

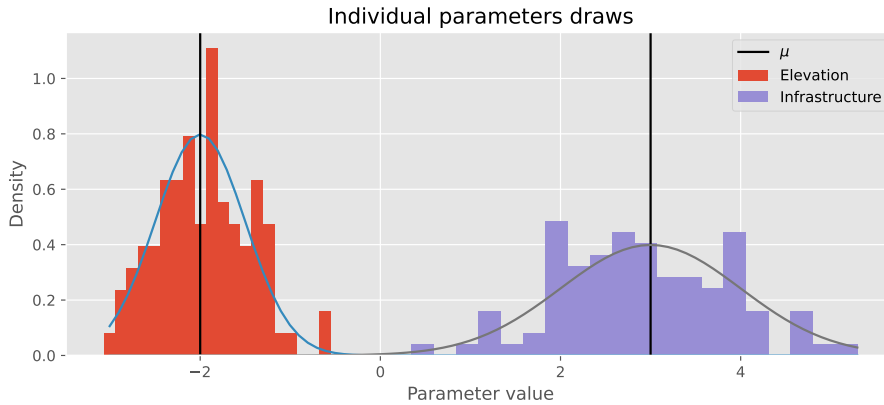


Fig. 2: Histogram of 100 draws (a population sample) for the random parameters β_E and β_I

Step 3a - Draw OD pairs: For each individual n , draw uniformly a permutation π_n of the trip purposes fulfilled on each OD. For each drawn individual, the points A, B, C can either be their "home", "work/study place" or "leisure" place. These are allocated randomly with an equal probability of $\frac{1}{3}$.

Step 3b - Draw number of observations: The number of drawn observations depends on the β_E values. This means individuals with more observations in the dataset are the least sensitive to

elevation gain. For $n \in \{1, \dots, N\}$:

$$T_n = \phi(\beta_n) = \lceil a \exp(b * \lambda(\beta_{E,n})) \rceil \quad (8)$$

$\lambda(\beta_{E,n})$ is the index of $\beta_{E,n}$ in the sequence of $\beta_{E,n}$ in increasing order. The maximum number of draws per individual has been set to $n^* = 200$, so we chose $a = \frac{n^*}{\exp(bN)}$ so that $n_1 = 1$ and $n_N = n^*$. $b = 0.075$ is a scaling constant. We draw a total of 2819 observations.

Step 4 - Draw of observations For $n \in \{1, \dots, N\}$ and $t \in \{1, \dots, T_n\}$ do:

1. Draw the used OD pair $p_{n,t} \sim \text{Categorical}(q_1, \dots, q_P | \pi_n)$ using the probability distribution of bicycle trip purposes given by the Danish National Travel Survey (Christiansen & Baescu, 2022). $q_1 = \mathbb{P}(\text{Home-Work}) = 0.46$, $q_2 = \mathbb{P}(\text{Home-Leisure}) = 0.24$, $q_3 = \mathbb{P}(\text{Work-Leisure}) = 0.3$.
2. Draw the chosen alternative $y_{n,t} \sim \text{Categorical}(L(\beta_n, \mathbf{X}_{p_{n,t}}))$, where

$$L(\beta_n, \mathbf{X}_{p_{n,t}}) = \frac{\exp(\beta_n^\top \mathbf{X}_{p_{n,t}})}{\sum_{l \in \mathcal{C}_{p_{n,t}}} \exp(\beta_n^\top \mathbf{X}_{l,p_{n,t}})} \quad (9)$$

Step 5: Based on these observations re-estimate a discrete choice model.

Step 4 is repeated $N_{exp} = 100$ times, to account for the randomness of the dataset creation process. The flow-chart (Figure 3) illustrates the described steps.

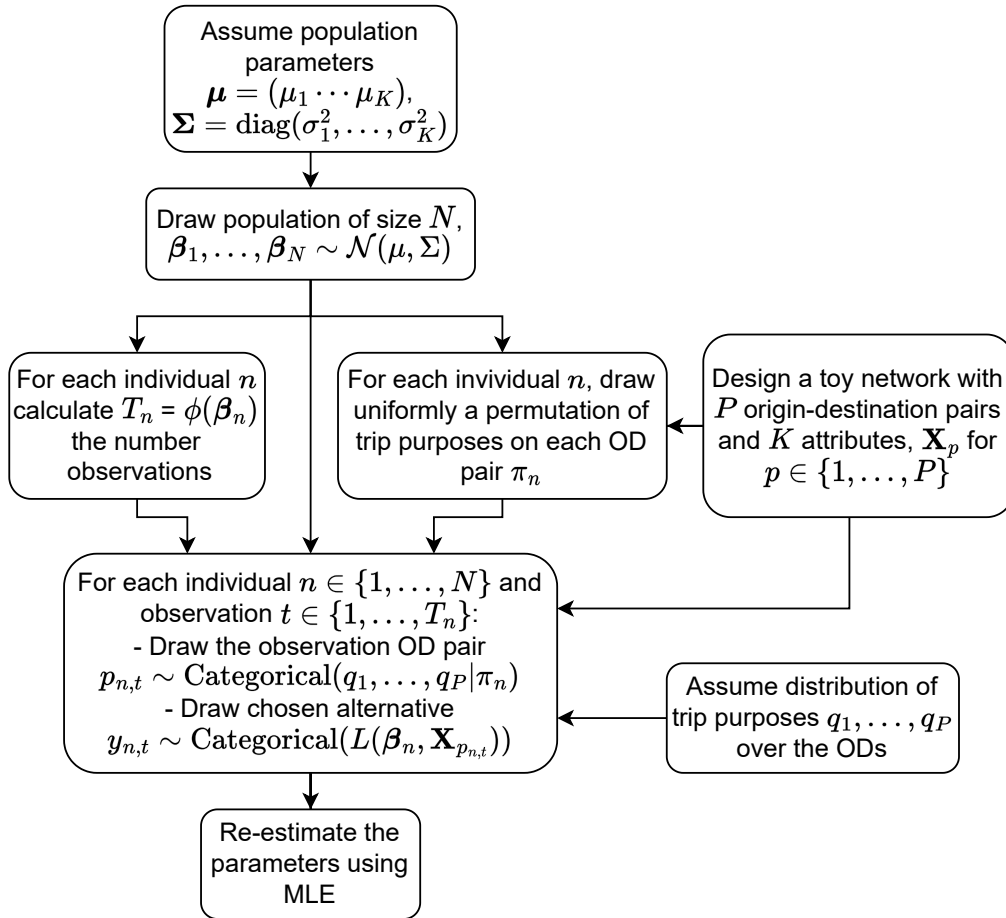


Fig. 3: Flowchart of the simulation experiments

All models are estimated using the Python library *xlogit* (Arteaga et al., 2022).

4 RESULTS: BASE MODELS

Three base models have been estimated: a Multinomial Logit model (MNL), a Mixed Logit model (MXL) and a Panel Mixed Logit Model (PMXL). The estimated parameters are summarized on Table 1. The estimated distributions are also plotted on Figure 4. We calculate the Marginal rates of substitution (onwards referred to as *tastes*) $Taste(x) = \frac{\mu_x}{\mu_L}$ as the ratio of any coefficient and the Length coefficient. This allows separating the issue of the estimation of the model scale and the derivation of people’s preferences. While the model scale defines one’s sensitivity to an attribute change on choice probabilities, tastes allow understanding the relative *value of distance* of each attribute.

The bias of tastes is defined as:

$$\text{Bias of tastes} = \left(\frac{\mu_E}{\mu_L} - \frac{\hat{\mu}_E}{\hat{\mu}_L} \right)^2 + \left(\frac{\mu_I}{\mu_L} - \frac{\hat{\mu}_I}{\hat{\mu}_L} \right)^2 + \left(\frac{\mu_S}{\mu_L} - \frac{\hat{\mu}_S}{\hat{\mu}_L} \right)^2$$

Tab. 1: Estimates, tastes, bias of tastes and D-error for the base models

	μ_L	μ_E	μ_I	μ_S	μ_{PS}	σ_E	σ_I	$\frac{\mu_E}{\mu_L}$	$\frac{\hat{\mu}_E}{\hat{\mu}_L}$	$\frac{\hat{\mu}_S}{\hat{\mu}_L}$	Bias of tastes	D-error
True value	-10	-2	3	1	1.5	0.5	1	0.2	-0.3	-0.1	-	-
MNL	-5.783	-0.775	1.729	0.486	1.327	-	-	0.133	-0.299	-0.0849	0.0584	0.0011
MXL	-9.938	-1.334	2.991	0.974	1.518	0.345	0.945	0.134	-0.302	-0.0982	0.0657	0.0028
PMXL	-9.964	-1.694	2.986	0.993	1.491	0.484	0.863	0.170	-0.300	-0.0997	0.0296	0.0013

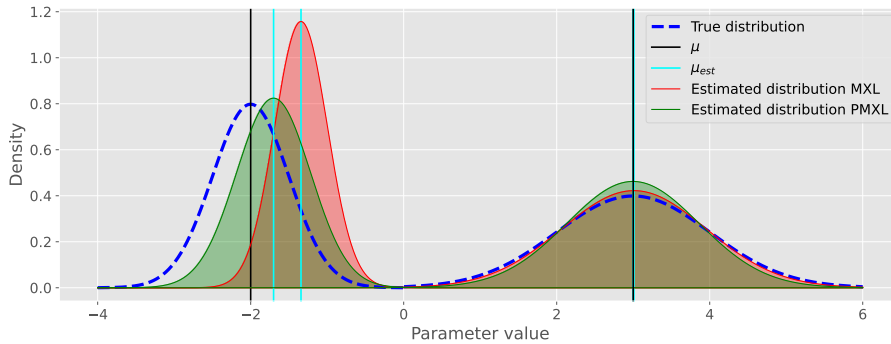


Fig. 4: Base models estimated parameters for elevation (left distribution) and infrastructure (right distribution) or distribution against true distributions

For the MNL model, the model marginal substitution rate is -13.3% for elevation gain, while the actual value is -20%. For bicycle infrastructure, the model outputs a taste of +29.9%, while the actual value is +30%. The MXL shows similar bias. Moreover, the MXL estimates a standard deviation for elevation gain that is way lower than the actual value. The PMXL shows less bias than the other base models, and estimates closer standard deviations to the true values. However it does not represent the average tastes of the individuals in the dataset. The taste for elevation gain is still shifted towards over-represented individuals (see also Figure 4). For the other parameters, however, all models show almost no bias. This is because the individual number of observations is uncorrelated with the other parameter values.

These estimations allow us to search for ways to decrease the taste bias. The developed strategies are presented in the following section.

5 SAMPLING AND WEIGHTING STRATEGIES

We implement a number of strategies to correct the bias in tastes of the **Panel Mixed Logit model** estimated on unbalanced data. These methods use subsampling, weighting techniques, or a combination of both.

Sampling strategies

To correct the dataset unbalance, several sampling strategies have been tested out. Their purpose is to reduce the bias in the estimated parameters compared to the actual parameters of the population.

Naive subsampling: We reduce the dataset size by randomly drawing a subset of the observation. This method does not aim to reduce the bias, but is used as a benchmark for bias and efficiency.

Pruning: We remove any individual with less than k_0 observations from the dataset.

Uniform random subsampling: To keep the same number of choice experiments for each individual, we choose k as the minimum number of observations for an agent in the dataset and select k observations for each individual randomly. This method is naive, but some extensions could be added, e.g. keeping dissimilar observations.

Uniform random truncation: For many experiments, some individuals will only have one observation in the dataset ($k = 1$). Thus, another method would be to randomly choose $n > 1$ and select n observations per individual. If an individual has less than n observations in the dataset, all their observations would be selected.

Subsampling of repeated observations: Another possibility for subsampling that would keep more variability in the dataset is to have a subsampling method that keeps, for each individual, one observation per choice scenario. The unbalance in the resulting dataset can again be handled by weighting the likelihood function.

A Maximum Weighted Likelihood Estimation (MWLE)

Sampling strategies reduce bias by reducing the potential over-representation of some individuals in the dataset. However, it also leads to lower efficiency of the estimates. Another way to deal with this bias would be to modify the likelihood function so that the estimator accounts for the dataset unbalance by penalizing over-represented individuals.

We implemented a new method to weigh the likelihood function. Let $\beta = (\mu, \Sigma)$ be the model parameters. The goal of the weighting algorithm we describe below is that each individual contributes equally (has the same weight) in the likelihood function. For individuals $n \in \{1, \dots, N\}$, we note $\mathbf{w} = (w_1 \dots w_N)$ the vector of individual weights. We note $\mathbf{LL}(\beta) = (\ln \hat{P}_{\mathbf{i}_1}(\beta) \dots \ln \hat{P}_{\mathbf{i}_N}(\beta))$ the vector of individual contributions to the log-likelihood. The weighted likelihood function is thus given by:

$$\text{LL}_{\mathbf{w}}(\beta) = \mathbf{w}^\top \mathbf{LL}(\beta) = \sum_{n=1}^N w_n \ln \hat{P}_{\mathbf{i}_n}(\beta)$$

Our goal is to find the vector of weights \mathbf{w}^* for which each individual gives the same contribution to the weighted likelihood function evaluated at the estimated parameters. Each iteration calculates weights that are inversely proportional to the weighted likelihood contribution of an individual, using the weights of the previous iteration. This is equivalent to solving the following fixed-point problem: $\mathbf{w}^* = F(\mathbf{w}^*)$, where, for an element w_j of \mathbf{w} , we have:

$$F(w_j) = \frac{(w_j \ln \hat{P}_{\mathbf{i}_j}(\hat{\beta}))^{-1}}{\frac{1}{N} \sum_{n=1}^N (w_n \ln \hat{P}_{\mathbf{i}_n}(\hat{\beta}))^{-1}}; \hat{\beta} = \arg \max_{\beta} \text{LL}_{\mathbf{w}}(\beta) \quad (10)$$

$\phi = \left(\frac{1}{N} \sum_{n=1}^N (w_n \ln \hat{P}_{\mathbf{i}_n}(\hat{\beta}))^{-1} \right)^{-1}$ is a normalizing constant ensuring that $\sum_{n=1}^N w_n = N$; which will be useful to compute the AVC matrix of the estimates.

To solve this fixed-point problem, the solution algorithm builds a sequence of weight vectors $\mathbf{w}^{(k)} = (w_1^{(k)} \dots w_N^{(k)})$ which can be described by the pseudo-code below.

Algorithm 1: Algorithm to determine optimal weights; $\mathbf{w}^* = F(\mathbf{w}^*)$

Input: $\mathbf{X}, \mathbf{y}, n_0, \varepsilon$
Result: \mathbf{w}^* , the vector of optimal weights
Initialization:
 $\mathbf{w}^{(1)} \leftarrow (1 \dots 1)$
 $k \leftarrow 1$
while $\|F(\mathbf{w}^{(k)}) - \mathbf{w}^{(k)}\| > \varepsilon$; // F also depends on \mathbf{X}, \mathbf{y}
do
 $\hat{\mathbf{w}}^{(k+1)} \leftarrow F(\mathbf{w}^{(k)})$;
 if $k > n_0$ **then**
 $\mathbf{w}^{(k+1)} \leftarrow \lambda_k \hat{\mathbf{w}}^{(k+1)} + (1 - \lambda_k) \mathbf{w}^{(k)}$; // Method of successive averages
 else
 $\mathbf{w}^{(k+1)} \leftarrow \hat{\mathbf{w}}^{(k+1)}$;
 end
 $k \leftarrow k + 1$
end

The Method of Successive Averages (MSA) (Robbins & Monro, 1951), ensures the convergence of the sequence. We use $\lambda_k = \frac{1}{k-n_0}$, so that the k^{th} calculated weight vector is the arithmetic mean of the previously computed weights, i.e. $\mathbf{w}^{(k)} = \frac{1}{k-n_0} \sum_{i=n_0}^k \hat{\mathbf{w}}^{(i)}$. This method begins to be applied after n_0 iterations, so the first weights are not included in the average.

The simulation experiment described in section 3 is then carried out for the following setups:

1. Whole dataset, weighted
2. Random naive subsampling at 500 observations
3. Pruning individuals with less than 5 observations
4. Randomly truncated at 2, 5, 10, 20, 50 observations, unweighted and weighted
5. Randomly subsampled at the minimum number of observations per individual (equivalent to a truncation at 1 observation with this dataset)
6. Random subsampling of unique observations

Setups 2 to 5 use random subsamples of the generated datasets. The subsampling algorithms are used 100 times for the same dataset, and the results are averaged (the standard deviation is also calculated). With 100 different generated populations and observations, these setups are repeated 10000 times.

6 RESULTS

This section compares two metrics for the different strategies: Bias of tastes and D-error, stored in Table 2. Some insights given by these tables are:

- For the models using the whole datasets, the bias in tastes has been significantly reduced by the weighting algorithm while not affecting the model efficiency (see Figure 5 for a plot of the estimated mixing distributions).
- As expected, the naive and pruning strategies, used as benchmarks, give worse results than the base model in efficiency and bias.
- The unweighted truncation give the best results in terms of *Bias of tastes* for low truncation thresholds (see Figure 6). The more the dataset is truncated, the more it is balanced, and the more the estimates are close to the population's mean. Conversely, lower truncation thresholds also lower the estimated parameters' efficiency, and increase variability between random subsamples. The bias-variance trade-off is highlighted by the green curve on Figure 8.

- The weighted truncation, while slightly increasing the D -error for the same truncation threshold, shows a decrease in bias of tastes when the threshold increases at 50 (see Figure 7). The red curve on Figure 8 shows that the weighting algorithm breaks the bias-variance trade-off when a certain truncation threshold is exceeded.
- Removing repeated observations behave similarly to truncation.

Moreover, random truncation to lower thresholds gives more variability to the model output; it is important to repeat random truncation several times and average the results to lower bias.

Tab. 2: Estimates, tastes, bias of tastes and D-error for the base models

	β_L	β_E	β_I	β_S	β_{PS}	σ_E	σ_I	$\frac{\beta_E}{\beta_L}$	$\frac{\beta_I}{\beta_L}$	$\frac{\beta_S}{\beta_L}$	Bias of tastes	D-error
<i>True value</i>	-10	-2	3	1	1.5	0.5	1	0.2	-0.3	-0.1	-	-
MNL	-5.783	-0.775	1.729	0.486	1.327	-	-	0.133	-0.299	-0.0849	0.0584	0.0011
MXL	-9.938	-1.334	2.991	0.974	1.518	0.345	0.945	0.134	-0.302	-0.0982	0.0657	0.0028
PMXL	-9.964	-1.694	2.986	0.993	1.491	0.484	0.863	0.170	-0.300	-0.0997	0.0296	0.0013
PMXL, w	-10.307	-2.136	3.162	1.021	1.529	0.617	0.960	0.207	-0.307	-0.0991	0.0107	0.0011
Naive	-10.188	-1.538	3.154	1.002	1.530	0.365	0.998	0.150	-0.310	-0.0987	0.0395	0.0093
Pruned	-9.979	-1.509	2.954	0.993	1.495	0.357	0.838	0.151	-0.296	-0.0996	0.0490	0.0013
Trunc 2	-10.333	-1.999	3.109	1.002	1.574	0.423	1.018	0.193	-0.301	-0.0975	0.0071	0.0421
Trunc 5	-10.038	-1.849	3.067	0.982	1.522	0.379	1.003	0.184	-0.306	-0.0981	0.017	0.0143
Trunc 10	-9.989	-1.786	3.082	0.984	1.497	0.369	1.005	0.179	-0.309	-0.0987	0.023	0.0074
Trunc 20	-9.988	-1.749	3.104	0.989	1.491	0.373	1.006	0.175	-0.311	-0.0992	0.0271	0.0042
Trunc 50	-10.001	-1.727	3.094	0.992	1.496	0.402	0.975	0.173	-0.309	-0.0993	0.0288	0.0023
Trunc 2, w	-12.908	-2.600	3.990	1.257	2.073	0.473	1.199	0.201	-0.309	-0.0980	0.0106	0.0703
Trunc 5, w	-11.421	-2.252	3.612	1.113	1.743	0.451	1.126	0.197	-0.316	-0.0978	0.0175	0.0194
Trunc 10, w	-10.927	-2.165	3.501	1.072	1.636	0.465	1.097	0.198	-0.320	-0.0983	0.0214	0.0091
Trunc 20, w	-10.687	-2.161	3.438	1.051	1.589	0.495	1.073	0.202	-0.322	-0.0985	0.0228	0.0048
Trunc 50, w	-10.460	-2.163	3.311	1.033	1.554	0.545	1.015	0.207	-0.317	-0.0988	0.0189	0.0023
Obs	-9.977	-1.781	3.076	0.993	1.524	0.362	1.001	0.178	-0.309	-0.0997	0.0231	0.0075
Obs, w	-11.886	-2.338	3.735	1.169	1.778	0.464	1.148	0.196	-0.314	-0.0989	0.0156	0.0397

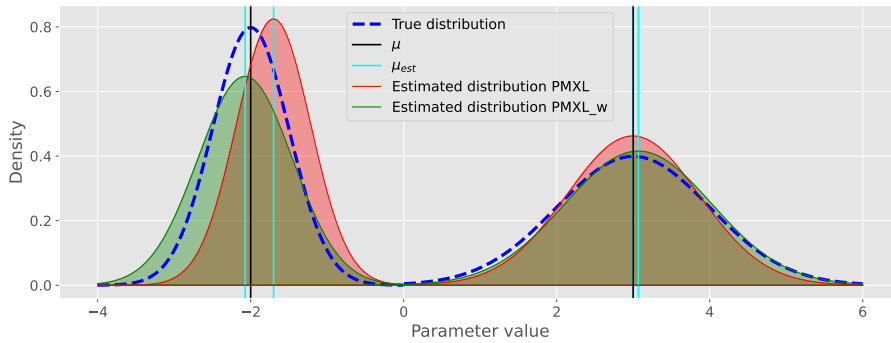


Fig. 5: Estimated distributions for elevation gain (left) and infrastructure (right), PMXL and PMXL weighted

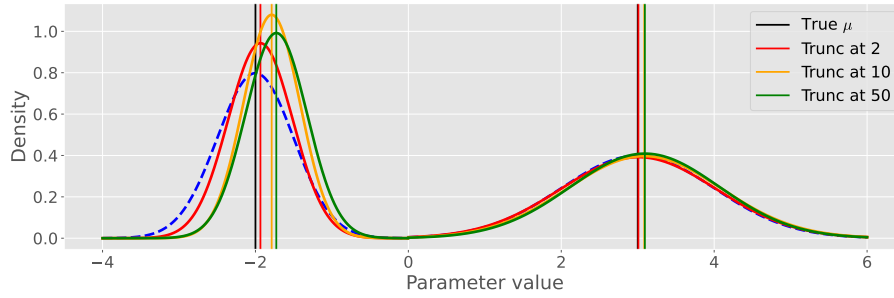


Fig. 6: Estimated distributions for elevation gain (left) and infrastructure (right), Truncated at 2, 10 and 50

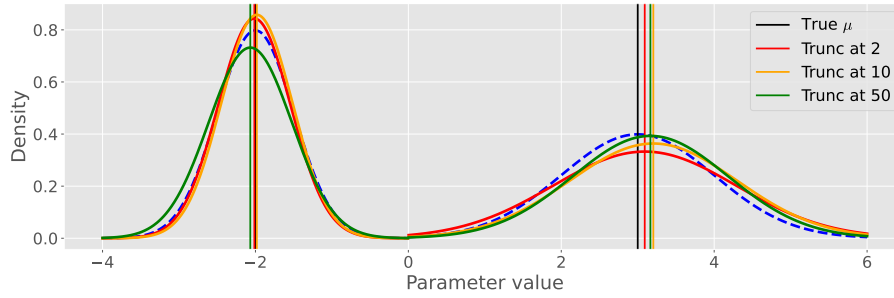


Fig. 7: Estimated distributions for elevation gain (left) and infrastructure (right), Truncated at 2, 10 and 50, weighted

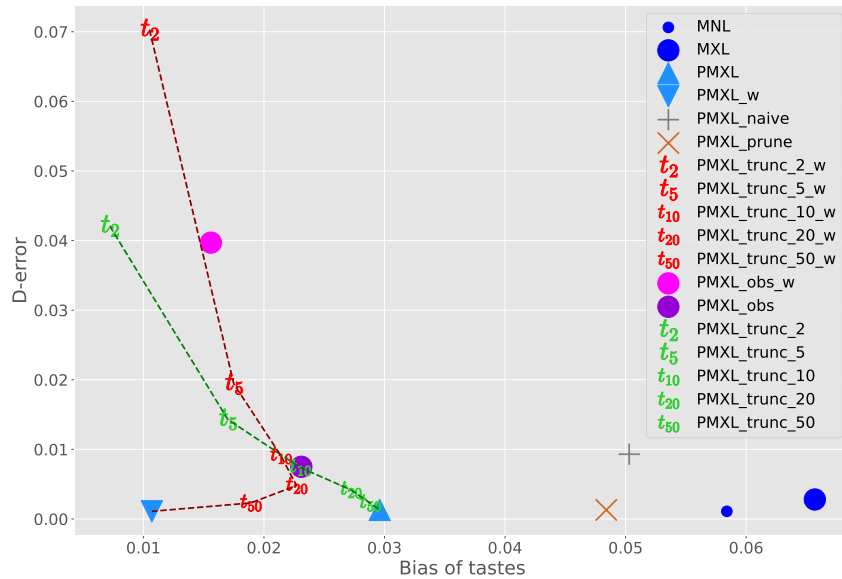


Fig. 8: Bias of tastes vs. D-error for all the different strategies

7 CONCLUSION AND FUTURE WORK

The simulation study has shown that it is possible to remove bias by applying weighting and subsampling methods. However, it also shows the bias-variance trade-off a modeller may face when choosing an optimal strategy. The newly-developed weighting algorithm breaks this trade-off by evening the contribution of each individual in the likelihood function. This allows maximum potential efficiency, while keeping a reliable explanatory model that is not biased towards over-represented individuals.

This simulation included one mixed parameter that was correlated (β_E) to the number of observation and one that was not (β_I). The results show that the models estimates are mostly unbiased for the parameter that did not correlate. As individual parameters are randomly distributed in

the population, we deduce that if the over-representation of some tastes is randomly distributed, the change in estimation may be negligible. The next step is to test these different strategies on different datasets to see how the preferences may change, i.e. how the number of observations correlates with the individual taste. Future work could also encompass a more thorough analysis on large-scale datasets, showcasing and tackling further challenges, such as bias in scale or computational burden.

ACKNOWLEDGEMENTS

We acknowledge Prof. Anders Fjendbo Jensen for his help in the early steps of the project.

REFERENCES

- Arteaga, C., Park, J., Beeramoole, P. B., & Paz, A. (2022). xlogit: An open-source python package for gpu-accelerated estimation of mixed logit models. *Journal of Choice Modelling*, 42, 100339.
- Ben-Akiva, M., & Ramming, S. (1998). Lecture notes: Discrete choice models of traveler behavior in networks. *Prepared for Advanced Methods for Planning and Management of Transportation Networks. Capri, Italy*, 25.
- Bliemer, M. C., & Rose, J. M. (2010). Construction of experimental designs for mixed logit models allowing for correlation across choice observations. *Transportation Research Part B: Methodological*, 44(6), 720–734.
- Cherchi, E., & Cirillo, C. (2014). Understanding variability, habit and the effect of long period activity plan in modal choices: a day to day, week to week analysis on panel data. *Transportation*, 41(6), 1245–1262.
- Cherchi, E., Cirillo, C., & de Dios Ortúzar, J. (2017). Modelling correlation patterns in mode choice models estimated on multiday travel data. *Transportation Research Part A: Policy and Practice*, 96, 146–153.
- Christiansen, H., & Baescu, O. (2022). The danish national travel survey: Annual statistical report for denmark for 2021.
doi: 10.11581/dtu:00000034
- Kessels, R., Goos, P., & Vandebroek, M. (2006). A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, 43(3), 409–419.
- Lee, K., & Sener, I. N. (2021). Strava metro data for bicycle monitoring: a literature review. *Transport reviews*, 41(1), 27–47.
- McFadden, D., et al. (1973). Conditional logit analysis of qualitative choice behavior.
- McFadden, D., & Train, K. (2000). Mixed mnl models for discrete response. *Journal of applied Econometrics*, 15(5), 447–470.
- Nelson, T., Ferster, C., Laberee, K., Fuller, D., & Winters, M. (2021). Crowdsourced data for bicycling research and practice. *Transport reviews*, 41(1), 97–114.
- Ortelli, N., de Lapparent, M., & Bierlaire, M. (2022). Faster estimation of discrete choice models via dataset reduction.
- Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The annals of mathematical statistics*, 400–407.
- Rose, J. M., Hess, S., Bliemer, M. C. J., & Daly, A. (2009). The impact of varying the number of repeated choice observations on the mixed multinomial logit model. *Transportation Research Record*(September), 1–15.
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

- van Cranenburgh, S., & Bliemer, M. C. (2019). Information theoretic-based sampling of observations. *Journal of choice modelling*, *31*, 181–197.
- Yáñez, M. F., Cherchi, E., Heydecker, B. G., & de Dios Ortúzar, J. (2011). On the Treatment of Repeated Observations in Panel Data: Efficiency of Mixed Logit Parameter Estimates. *Networks and Spatial Economics*, *11*(3), 393–418. doi: 10.1007/s11067-010-9143-6

Resilience-Oriented Design for Public Transport Networks

Christina Iliopoulou^{1*}, Ph.D., Michalis Makridis¹, Ph.D, Anastasios Kouvelas¹, Ph.D.

¹Institute for Traffic Planning and Transport Systems, ETH Zurich, Stefano Franscini Square 5,
8093 Zurich, Switzerland, *e-mail:ciliopoulou@ethz.ch

SHORT SUMMARY

Public transport systems are typically designed based on estimated passenger demand and supply patterns, yet may often be called to operate under vastly different operational settings. To systematically design resilient transit systems, it is necessary to “weave” resilience-oriented thinking into the established public transport network design process, moving from an abstract concept to an implementable methodology. This study aims to effectively and efficiently design resilient public transport networks through the integration of Reinforcement Learning (RL), Local Search operators and Particle Swarm Optimization. We present a redundancy indicator and integrate it within a hybrid RL-enhanced metaheuristic solution framework to design more resilient route structures. We apply the proposed Memetic algorithm to an established benchmark from the literature and validate the proposed approach under a series of random and targeted attacks, simulating link disruptions. Results demonstrate that resilience can be enhanced through redundancy without adversely impacting average travel times.

Keywords: Transit Route Network Design, Resilience, Vulnerability, Redundancy, Memetic algorithms, Reinforcement learning.

1. INTRODUCTION

Most existing systems are over-optimized to predefined design inputs. However, as a great deal of uncertainty persists, these systems become fragile to ever-changing external conditions. This is especially true when it comes to transportation systems, and particularly public transport networks (Mattsson and Jenelius, 2015). Indeed, these systems were once designed for predicted passenger demand and supply patterns, yet are now called to operate under vastly different operational settings. Still, medium and long-term disruptions (e.g. road closures, maintenance works) can induce significant changes in supply and trigger shifts in passenger demand, rendering public transport systems vulnerable and often unviable. In such cases, associated costs incurred by passengers and operators are unaccounted for in the design process, albeit significant. Due to rigid constraints and limitations of the underlying network structure, strategic interventions are limited in such cases, with disruptions associated with wide and sustained implications (Jenelius and Cats, 2015).

Under this scope, an important consideration refers to the ability of transport networks to withstand perturbations and maintain their serviceability. Towards this goal, several concepts have been introduced in the literature to evaluate the performance of transport networks under disruption, notably robustness, vulnerability, and resilience (Ge et al., 2022). In transportation-related literature, vulnerability and resilience are viewed as core properties of public transport systems,

both considering the decrease in network performance under perturbations (Ge et al., 2022; Mattsson and Jenelius, 2015).

Under this scope, a large stream of studies sought to identify critical network links/segments, a posteriori investigating the link between network design and vulnerability (Cats, 2016; Mattsson and Jenelius, 2015; Rodríguez-Núñez and García-Palomares, 2014; von Ferber et al. 2012). Despite these observations, in terms of a priori designing resilient transit networks, there is a large gap in the respective literature. The problem of optimally designing surface public transportation systems, referred to as Transit Route Network Design Problem (TRNDP), has attracted the interest of the research community for over five decades (Iliopoulou et al. 2019). Still, despite the vast literature on the general TRNDP, methods for designing resilient transit networks are lacking.

Motivated by the gap in the respective literature, this study presents a methodological framework for enhancing resilience within the TRNDP, without negatively impacting performance. To achieve this, we incorporate Reinforcement Learning (RL) within a metaheuristic solution framework to reinforce the resilience of transit networks, under the planning paradigm for enhanced connectivity. In particular, the availability of trip alternatives, i.e. the amount of redundancy offered, allows for the impacts of link-based incidents to be absorbed (Rodríguez-Núñez and García-Palomares, 2014). To that end, a particle swarm optimization (PSO) algorithm integrated with neighborhood operators which manipulate the degree of connectivity, referred to as Memetic PSO, is developed and enhanced through RL to reinforce path redundancy. We adopt the view that resilience is related to the network's performance decrease, as computed based on demand coverage and transport efficiency, under a shift in operating conditions and develop a set of bidirectional link-based disruption scenarios to be investigated, representing full link closures (Cats, 2016).

2. METHODOLOGY

This section presents a general formulation for the TRNDP, the proposed redundancy indicator and outlines the components of the developed algorithmic framework.

The TRNDP

In general, there is no commonly accepted mathematical programming formulation for the TRNDP due to its inherent complexity and discrete nature (Iliopoulou et al., 2019). A high-level mathematical formulation for the TRNDP can be given as follows. Let:

ATT_{RS} :	Average Travel Time for route set RS
\mathbf{C} :	the vector of path costs on the transit network
d_{0RS} :	Percentage of passenger demand satisfied without transfers for route set RS
d_{1RS} :	Percentage of passenger demand satisfied with one transfer for route set RS
d_{2RS} :	Percentage of passenger demand satisfied with two transfers for route set RS
d_{unRS} :	Percentage of unsatisfied demand for route set RS
R :	Route $\in RS$
$\hat{\mathbf{RS}}$:	Vector of optimal routes
RM :	Maximum number of routes R in RS
RS :	Set of routes $\{R\}$
\mathbf{Q} :	the vector of segment flows on the transit network
p_{RS} :	Path redundancy score for route set RS
s_R :	Number of stops per route R

s_{min} :	Minimum number of stops
s_{max} :	Maximum number of stops
U :	the user route choice model function
ω :	Weighting factor

$$(\hat{\mathbf{RS}}) = \arg \min Z (\mathbf{RS}, \mathbf{Q}) \quad (1)$$

$$Z = ATT_{RS} - \omega p_{RS} \quad (2)$$

s.t.

$$(\mathbf{Q}, ATT_{RS}, d_{0RS}, d_{1RS}, d_{2RS}, d_{unRS}, p_{RS}) = U(C(\mathbf{RS})) \quad (3)$$

$$s_{min} \leq s_R \leq s_{max}, \quad \forall R \in RS \quad (4)$$

$$R \neq K \quad \forall R \in RS \quad (5)$$

$$|RS| \leq RM \quad (6)$$

$$d_{unRS} = 0 \quad (7)$$

The problem seeks to determine the route set that minimizes the objective function (Eq. 1). The latter represents the user cost associated with the route set, defined in this case as a score resulting from the weighted difference of average travel time (ATT) and the redundancy indicator (Eq. 2). In this case, ATT is the most important metric, as it also reflects direct ridership due to transfer penalization (Fan and Mumford, 2010). The values of ATT , the redundancy indicator and other route evaluation criteria, are derived from the transit assignment process, which is represented by Equation (3). Equation (4) specifies the minimum and the maximum number of stops for routes. Equation (5) states that two individual routes cannot coincide, while Equation (6) specifies the maximum number of lines. Last, Equation (7) states that the percentage of unsatisfied passengers must be zero.

Redundancy

The challenge in statically capturing resilience through a performance indicator is that one should account for the behaviour of the network under several disruption scenarios, as a full network scan is not computationally feasible for each candidate solution during the optimization. To that end, the number of alternative paths offers useful insights and is linked to better performance under disruptions, as route redundancy allows for flexibility to passengers (Cats, 2016). Based on this observation, we propose the use of a modified global efficiency indicator, which we name path redundancy, defined as follows:

$$p = \frac{\sum_{w \in W} d_w m_w}{\sum_{w \in W} d_w} \quad (8)$$

Where w denotes an OD pair, d_w the corresponding demand and m_w denotes the number of distinct transit paths offered based on the physical network. This distinction is important, as, different route combinations could use the same physical path. We filter the number of available paths through the physical road network to determine the number of distinct physical paths exploited by the route network. However, physical path overlap must still be taken into account, as it may lead to overestimating the number of distinct alternatives. To account for the reduction in redundancy due to overlapping segments we scale the number of physical paths by considering an overlap coefficient. More specifically, given a set of feasible paths L^w serving a specific OD pair w , the number of distinct physical paths is computed by considering the equivalent path index o_l for each path l , defined as follows:

$$o_l = \frac{\sum_{u \in U_l} \frac{1}{k_u^w}}{|U_l|} \quad l \in L^w \quad (9)$$

Where u is an edge of path l , U^l the set of edges comprising path l , k_u^w the number of paths for the specific OD pair w traversing segment u . Finally, the total number of distinct paths for a pair w is given by summing the corresponding values for all paths l in L^w :

$$m_w = \sum_{l \in L^w} o_l \quad (10)$$

So that m_w is the number of distinct physical paths for OD pair w , accounting for similarities.

The proposed Memetic RL-enhanced PSO algorithm

Motivated by the performance of emerging RL-enhanced PSO algorithms and Memetic PSO variants, we propose a discrete-space Memetic PSO where Q-learning is employed to select the search actions of each particle, referred to as MQLPSO. The proposed algorithm flexibly incorporates the discrete PSO operators for the TRNDP to effectively perform exploration for promising solutions within the entire region, a local-search procedure as the refinement step and a Q-learning framework as the operator selection mechanism. We employ four local search operators in total, which specifically target connectivity and thus, path redundancy, allowing the algorithm to perform exploitation for solution improvement in subregions. Each action contains a set of movements; three global search operators are used to move particles towards the global and personal best and four local search operators are used to refine individuals. The framework of the proposed algorithm is shown in Figure 1.

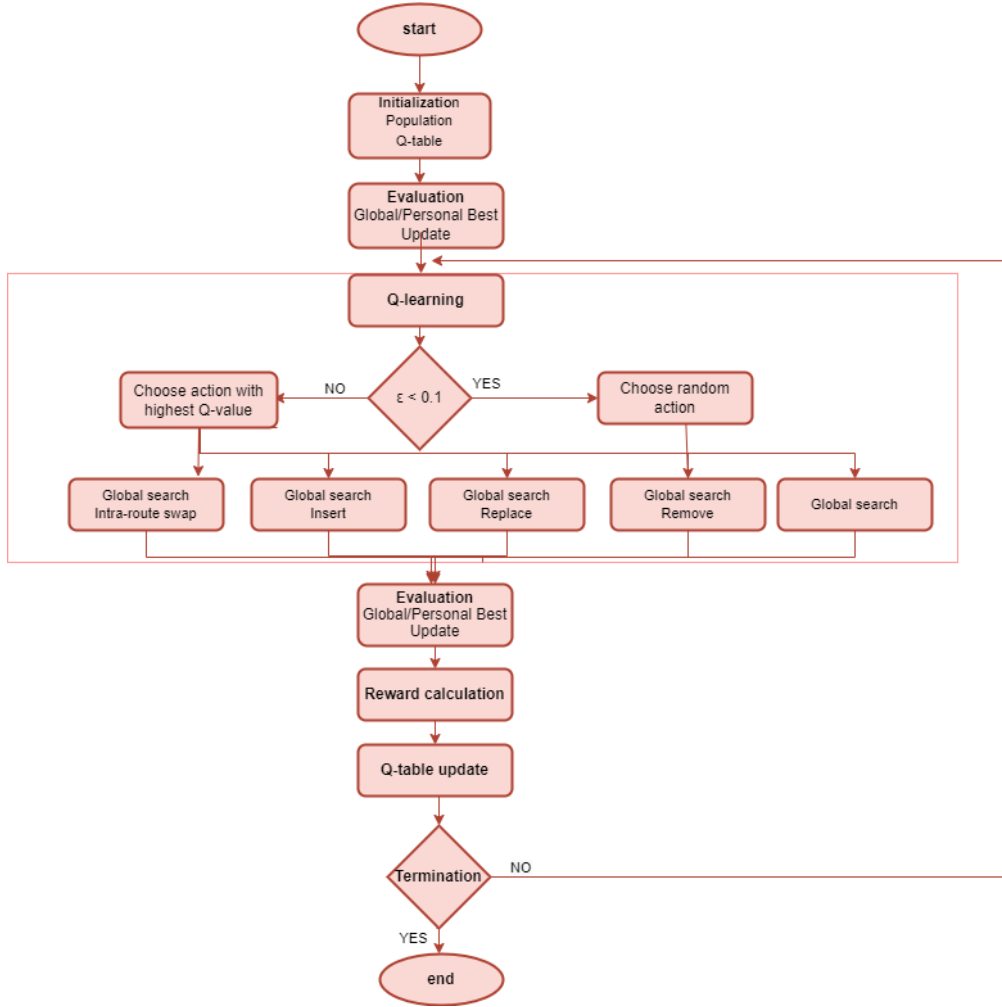


Figure 1 MQLPSO framework

In this specific instance, we are interested in designing a network that offers low average travel times with as much redundancy as possible. Therefore, the following three-piece reward function is defined, after experimentation:

$$r = \begin{cases} 1 & , \text{ if } ATT' < ATT \text{ and } p' > p \\ 0.5 & , \text{ if } ATT' < ATT \text{ or } p' > p \\ -1 & , \text{ otherwise} \end{cases} \quad (11)$$

Where ATT and ATT' denote the value of ATT for the previous solution and the current, respectively; p and p' are the redundancy scores of the previous and current solution, respectively.

3. RESULTS AND DISCUSSION

The road network used as input by the proposed MQLPSO algorithm is based on a real Swiss road network (Mandl, 1980), comprised of 15 nodes and 21 links., and is the widely accepted benchmark for the TRNDP. The demand matrix is symmetric, and the routes run in both directions. The configuration is shown in Figure 2.

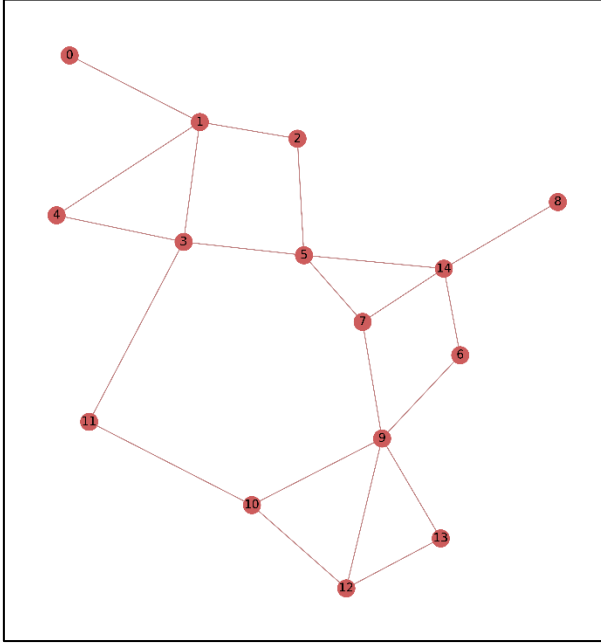


Figure 2. Mandl's Network configuration

In this case, we simulate link-based attacks, by removing links both at random and based on their criticality. For random attacks, we simulate the removal of 1- 10 random links and run 100 experiments per case to assess the impact on network performance to evaluate a representative set of scenarios (Matisziw et al., 2009). For targeted attacks, we consider scenarios without the most critical link and the most critical sequence of up to 10 link disruptions. This process enables the identification of links absorbing trips that have been diverted because of disruption in different scenarios. We identify critical links based on the value of the passenger betweenness centrality indicator (Cats, 2016). The measure is defined as follows:

$$PBC_u = \frac{1}{\sum_{w \in W} d_w} \sum_{w \in W} g_w(u) d_w \quad (18)$$

Where $g_w(u)$ denotes the fraction of shortest paths for OD pair w traversing link u . We recalculate betweenness centrality after each edge removal.

For evaluation, we will compare the following algorithmic setups using the same random seed, conducting 20 experiments in all cases:

- i. MQLPSO with coefficient $\omega=0.1$ in the objective function, so that *ATT* and redundancy are incorporated both at the objective and reward functions.
- ii. MQLPSO with coefficient $\omega=0$ in the objective function, so that *ATT* minimization is the optimization objective and redundancy is enforced only through the reward function.
- iii. The PSO global search step combined with random selection of local search operators, referred to as Memetic PSO (MPSO) aiming to minimize only *ATT*. We use this as a

benchmark reflecting the typical TRNDP design process while capturing the exploitation capabilities of the neighborhood operators to some extent.

Table 1 shows the results for transit network configurations with 4 routes.

Table 1. Comparison of solutions generated among methods for 4-Route Case

Performance Criteria	Algorithm	MLPSO best	MLPSO best	MPSO best	MLPSO avg	MLPSO avg	MPSO avg
	ω	0.1	0	0	0.1	0	0
ATT (min)		10.6	10.51	10.54	10.7	10.68	10.71
p		3.42	3.42	3.03	3.17	3	2.97
d_0 (%)		89.21	90.88	91.59	89.71	89.94	89.86
d_1 (%)		10.79	9.12	7.71	10.03	9.86	9.85
d_2 (%)		0	0	0.71	0.26	0.2	0.29
d_{un} (%)		0	0	0	0	0	0
Run time (s)					72	60	78

As seen in Table 1, MQLPSO produces similar quality results in both cases. The best solutions produced in this case feature the same redundancy value (3.42 vs 3.03 of MPSO, i.e. a 13% improvement), yet the solution under $\omega=0$ features a lower value for ATT and improved direct demand coverage, which is reasonable. In the average case, higher redundancy values are generated if both criteria are enforced through the objective function (3.17 on average), with a slight improvement of passenger-related performance criteria observed under $\omega=0$. Both cases of MQLPSO produce superior solutions to MPSO with zero two-transfer shares, demonstrating that the proposed RL scheme can improve both ATT and redundancy at the same time. To illustrate the value of reinforcing redundancy we compare the route configuration generated by MPSO with the route configuration under $\omega=0.1$, as on average it yields improved values for redundancy. Figure 3 shows the performance decrease of the network with $p=3.42$ vs $p=3.03$ under a series of random attacks, with boxplots summarizing 100 random runs.

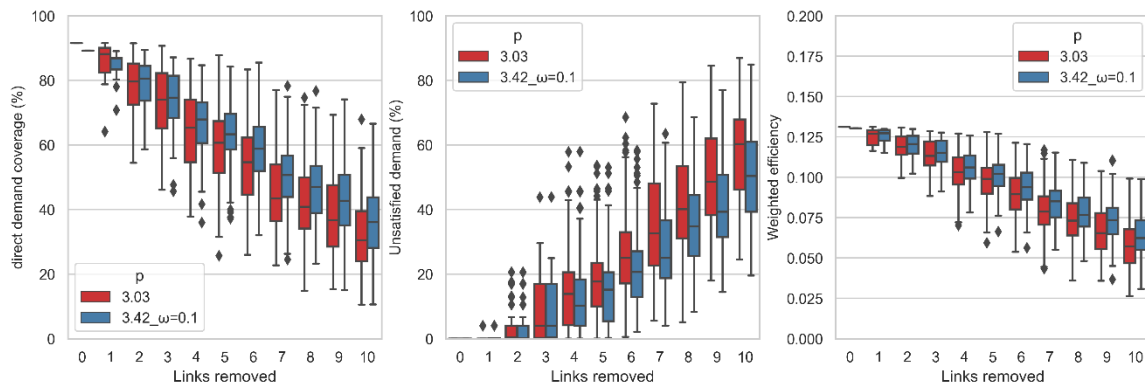


Figure 3. Random attack impacts for the 4-route case.

Figure 3 clearly shows the difference in performance decrease between the two network configurations, across all relevant indicators. As a general observation, a clear trend may be discerned where the route network with the smaller redundancy exhibits larger variability in terms of demand coverage under random attacks, with proportionally more scenarios resulting in worse outcomes. In fact, for the route network with the lowest path redundancy 25% of scenarios feature direct coverage between 45% and 64%, after the removal of 3 links versus 56% - 69% with

$p=3.42$, besides a couple of outliers. After 4 link removals, the median value for unsatisfied demand is consistently higher in the unprotected network, with discrepancies becoming larger with the extent of disruption. The removal of a 4th link seems to be a turning point for performance loss, as a notable rise in unsatisfied demand and an abrupt decline in the efficiency of both networks is observed with consistently inferior values for $p=3.03$. For 7-link disruptions, the differences between the two networks become more pronounced, with the more redundant network retaining a larger portion of its serviceability with under 25% of unsatisfied passengers in the median case, compared to 33% for the less redundant network. Even in the case of an extended 10-link disruption, the former maintains a 10% advantage in terms of demand satisfaction over the latter.

Figure 4 shows the performance decrease under targeted attacks up to 10-edge disruptions.

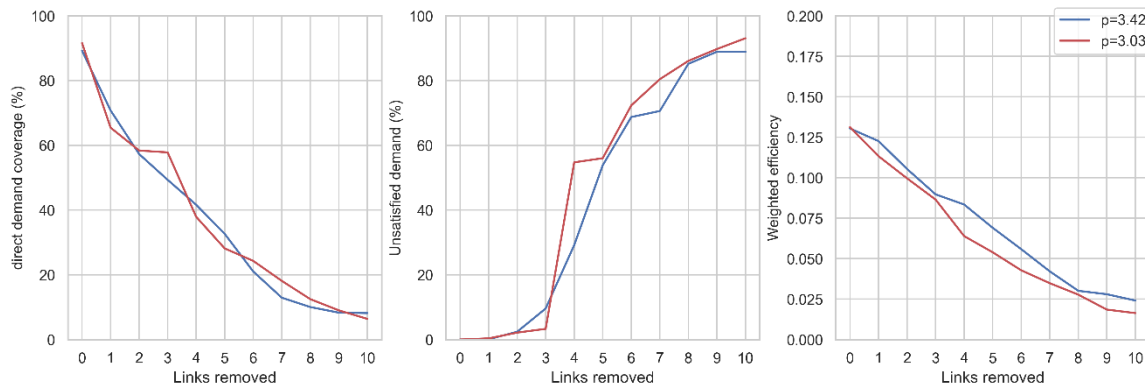


Figure 4. Targeted attack impacts for the 4-route case.

As seen in Figure 4, the value of redundancy becomes apparent after 3 consecutive removals, with unsatisfied demand higher across all cases after that point for the network with the lower redundancy measure. The efficiency and thus the passenger carrying capacity drops more abruptly in this case, with 4-link disruptions being the critical point. Indeed, for a sequence of 4 critical links removed, there is more than 25% difference in unsatisfied demand between the two networks and 20% in efficiency. Even though the relative gap becomes smaller with successive removals, the results for the unprotected network are consistently superior to those for the reinforced network. Even at 10 successive removals of the most critical link, the more redundant network serves 6% more passengers.

4. CONCLUSIONS

This study showcased a two-pronged resilience-oriented design framework: on the one hand, a redundancy indicator was developed and incorporated in the TRNDP solution process within an intelligent optimization framework and on the other hand, multiple simulation runs were performed to assess generated solutions and evaluate the design process. Results demonstrated that redundancy can reduce the impacts of link disruptions, reducing associated repercussions on demand coverage, with negligible costs in terms of average travel times. The more redundant network incurred a lower share of unsatisfied demand for multiple-link disruptions, either random or targeted, and a smoother decline in efficiency, retaining a larger portion of its serviceability. Respectively, the networks with lower redundancy exhibited more unpredictable behavior under attacks, with higher variability and more damaging worst-case scenarios. For targeted attacks, the value of redundancy is still evident, yet relatively lower compared to the case of random attacks.

This may perhaps be expected, as targeted attacks are based on the maximum weighted betweenness centrality measure, which may still be high even in redundant networks.

ACKNOWLEDGEMENTS

This work was supported by the “Swiss Government Excellence Programme” through a Postdoctoral Fellowship (recipient: Christina Iliopoulou) under the project RESPOND.

REFERENCES

- Cats, O. 2016. The robustness value of public transport development plans. *Journal of Transport Geography*, Vol. 51, pp. 236-246.
- Fan, L., Mumford, C. L. 2010. A metaheuristic approach to the urban transit routing problem. *Journal of Heuristics*, Vol. 16, No. 3, pp. 353–372.
- Ge, L., Voß, S., Xie, L. 2022. Robustness and disturbances in public transport. *Public Transport*, pp. 1-71.
- Iliopoulou, C., Kepaptsoglou, K., Vlahogianni, E. 2019. Metaheuristics for the transit route network design problem: a review and comparative analysis. *Public Transport*, Vol. 11, No. 3, pp. 487-521.
- Jenelius, E., Cats, O. 2015. The value of new public transport links for network robustness and redundancy. *Transportmetrica A: Transport Science*, Vol. 11, No. 9, pp. 819-835.
- Mandl, C. E. 1980. Evaluation and optimization of urban public transportation networks. *European Journal of Operational Research*, Vol. 5, No. 6, pp. 396–404.
- Matisziw, T. C., Murray, A. T., Grubestic, T. H. 2009. Exploring the vulnerability of network infrastructure to disruption. *The Annals of Regional Science*, Vol. 43, No. 2, pp. 307-321.
- Mattsson, L. G., Jenelius, E. 2015. Vulnerability and resilience of transport systems—A discussion of recent research. *Transportation Research Part A: Policy and Practice*, Vol. 81, pp. 16-34.
- Rodríguez-Núñez, E., & García-Palomares, J. C. 2014. Measuring the vulnerability of public transport networks. *Journal of transport geography*, Vol 35, pp. 50-63.
- von Ferber, C., Berche, B., Holovatch, T., Holovatch, Y. 2012. A tale of two cities. *Journal of Transportation Security*, Vol. 5, No. 3, pp. 199-216.

Acceptance of car-reducing measures: observed factors and latent attitudes

Filippos Adamidis*¹, Sara Moghavam Ghaffari¹, Constantinos Antoniou²

¹M.Sc., Technical University of Munich, Germany

² Prof. Dr., Technical University of Munich, Germany

SHORT SUMMARY

Major cities are increasingly willing to reclaim public space from cars. This paper analyses the acceptance of car-reducing measures by different segments of the population. The respondents of a stated preference survey in Munich, Germany, were asked whether they accept one or more measures designed to decrease the ownership and use of private cars, and to state their opinion on theoretical statements regarding private cars and the environment. Factor analysis and binomial regression were employed to model the relationship between the established travel behaviour, socio-demographics and latent attitudinal constructs on the one side, with acceptance on the other. The results showed that age, education, occupation and income, as well as environmentally friendly travel behaviour and attitudes play a major role in acceptance, thus providing valuable policy recommendations.

Keywords: attitudes, factor analysis, socio-demographics, travel behaviour, transport policy

1. INTRODUCTION

Measures against private cars have the goal to reduce traffic congestion, to mitigate the external costs of transportation and to reclaim urban space for social, commercial and recreational activities. With this aim, many researchers proposed relevant measures such as road pricing, parking restrictions, improved infrastructure for active modes, incentives for public transport, within-neighbourhood mobility and concentration of land uses to reduce distances (Gärling, Gärling, & Johansson, 2000). However, the public acceptance of those measures varies and turns out to be a key issue for their success (Banister, 2008).

Such measures can be distinguished based on their coerciveness. According to Loukopoulos et al. (2005), coercive measures, also named hard or structural measures, are less acceptable by the public than noncoercive measures – an example of the former is the prohibition of car traffic in city centres. By contrast, noncoercive measures (e.g. reduced fares in public transport), also called soft or psychological measures, may be politically and socially more feasible (Friman, Larhult, & Gärling, 2013). Similarly, push measures are perceived as ineffective, unfair and not acceptable, whereas pull measures are perceived to be effective, fair and acceptable (Eriksson, Garvill, & Nordlund, 2008).

Another important factor is the effectiveness of those measures. Romero et al. (2019) investigated the influence of driving restrictions in Madrid and found the modal shift towards public transport to be modest, explained likely by the large number of cars driving into the city from the outskirts.

*Corresponding author: filippos.adamidis@tum.de

Yan, Levine, & Marans (2019) investigated the responses of travellers to different parking attributes, e.g. search time and parking cost, and found that the primary response was changing parking location rather than shifting to another mode. In Gonzalez, Gomez, & Vassallo (2022) parking restrictions and low emission zones were found to encourage greener mobility, although owners of cleaner vehicles were unwilling towards shifting to public transport due to the perceived benefits stemming from their vehicles. Becker, Ciari, & Axhausen (2018) stated that free-floating car-sharing could reduce car ownership, however, there are differences compared to station-based schemes.

The objective of this study is to provide insight into the factors associated with the acceptability of measures against private cars. These factors can be related to the established travel behaviour, the socio-demographics and the underlying attitudes of individuals when concerned with questions about car ownership and the environment. More information about the dataset, the methodology, the results and the main findings are presented in the next sections.

2. METHODOLOGY

Data collection

The dataset was collected through an online panel (Schlesinger Group) as part of a stated preference mode choice survey in Munich, Germany. The survey includes four bundles of measures that aim to reduce the ownership and use of private cars and to encourage the use of active and environmentally friendly alternatives:

1. Extending the existing mobility hub network.
2. Removing on-street parking spots in favour of multipurpose garages.
3. Promoting neighbourhood mobility by creating attractive public spaces.
4. Facilitating active mobility and restricting motorised private transport.

The respondents could accept one or more bundles or to reject the measures altogether (*None of the above - everything should remain as it is*). Furthermore, the survey included twelve attitudinal statements (Table 1) about car ownership and the environmental concerns of the individuals. The participants were asked to express their opinion about the statements in a five-level rating (Likert) scale from *Strongly disagree* to *Strongly agree*. Other collected data include information about their usual travel behaviour and socio-demographic questions.

Data analysis

Two modelling techniques were used in this study, namely factor analysis and binomial regression. Exploratory factor analysis (EFA) is a statistical technique that aims to explore the underlying correlations between measurable variables. The result of EFA is a latent construct (*factor*) and the association strength between the latent construct and the measured variables (*loadings*). Confirmatory factor analysis (CFA) is a structural equation modelling technique that investigates relations between latent constructs and observed variables in an a priori specified theoretical model. Therefore, to apply CFA in this study, we assume that underlying attitudes that influence the acceptance of the bundles of measures exist. Further, we assume that the acceptance is influenced by the established travel behaviour and to the socio-demographics of the sample. The null hypothesis in factor analysis is that the correlation matrix of the assumed model does not differ from the one implied by the data.

To construct the models, we assume that pro-environmental attitudes are associated with the acceptance of measures, while attitudes in favour of cars contribute to the rejection of the measures. The structural models are estimated using the *R*-package *lavaan* (Rosseel, 2012). Between EFA and CFA we employ binomial regression in order to perform a first selection of the observed variables.

Table 1: Statements regarding car ownership and the environment

Statement	Variable
Car ownership	
A car is a symbol of social status for me.	Symbol
Having access to a car invokes to me a feeling of independence and freedom.	Independence
The brand/manufacturer is important to me when choosing to buy a car.	Brand
I deserve to own a good car because I have been successful in life.	Success
I feel accomplished and fulfilled after buying a car.	Accomplishment
A car is essential to my everyday mobility needs.	Essential
Environment	
The use of individual motorised transport threatens the environment	Threat
It is my obligation to protect the environment through my transportation mode choice.	Protection
The government should increase the price of fuel in order to invest in public transport.*	Fuel Price
A way to reduce congestion is to ban cars from city centres.	Carfree
I am concerned about the future of our planet.	Future
I have already moved towards a more environmentally friendly lifestyle.	Lifestyle Change

*Source: Schmid, Schmutz, & Axhausen (2016)

3. RESULTS AND DISCUSSION

Established travel behaviour and socio-demographics

Some observations were filtered out of the initial sample, either because the respondents speeded through the questionnaire (completion time less than one third of the estimated time of 15 minutes), or because their socio-demographic category was underrepresented, e.g. gender *Diverse* or *I prefer not to say*. The resulting sample size was $N=1497$. At least one bundle was selected by 1230 respondents, whereas 267 respondents declined all measures. A comparison of the sample with the latest published census (Federal Statistical Office, 2011) can be seen in Table 2.

Table 2: Key characteristics of the respondents

Variable	Answer	Sample (%) <i>N=1497</i>	Census (%) <i>2011</i>
Gender	Female	46.8	51.7
	Male	53.2	48.3
Age	≤17	0	14.6
	18-29	18.0	17.2
	30-39	21.6	16.7
	40-49	18.7	16.2
	50-59	21.0	11.8
	60-69	14.8	10.7
	70-79	4.9	8.4
	≥80	0.8	4.4
Occupation	Full-time work	60.5	56.5
	Part-time work	16.2	
	Pupil, student or apprentice	6.5	4.5
	Retired	11.4	18.3
	Housewife/Househusband	1.9	2.9
	Other	2.4	17.8
	No answer	1.1	
Size of household (no. of people)	1	30.0	50.3
	2	37.3	28.8
	3	13.4	10.6
	≥4	18.9	10.3
Driving license	Yes	90.0	88.9*
	No	10.0	11.1*
Car ownership	0	29.8	44.0*
	1	52.0	49.0*
	≥2	18.2	7.0*

**Mobilität in Deutschland (infas, DLR, IVT & infas 360, 2018)*

Attitudes about car ownership and the environment

The answers to the attitudinal statements are summarised in Figure 1.

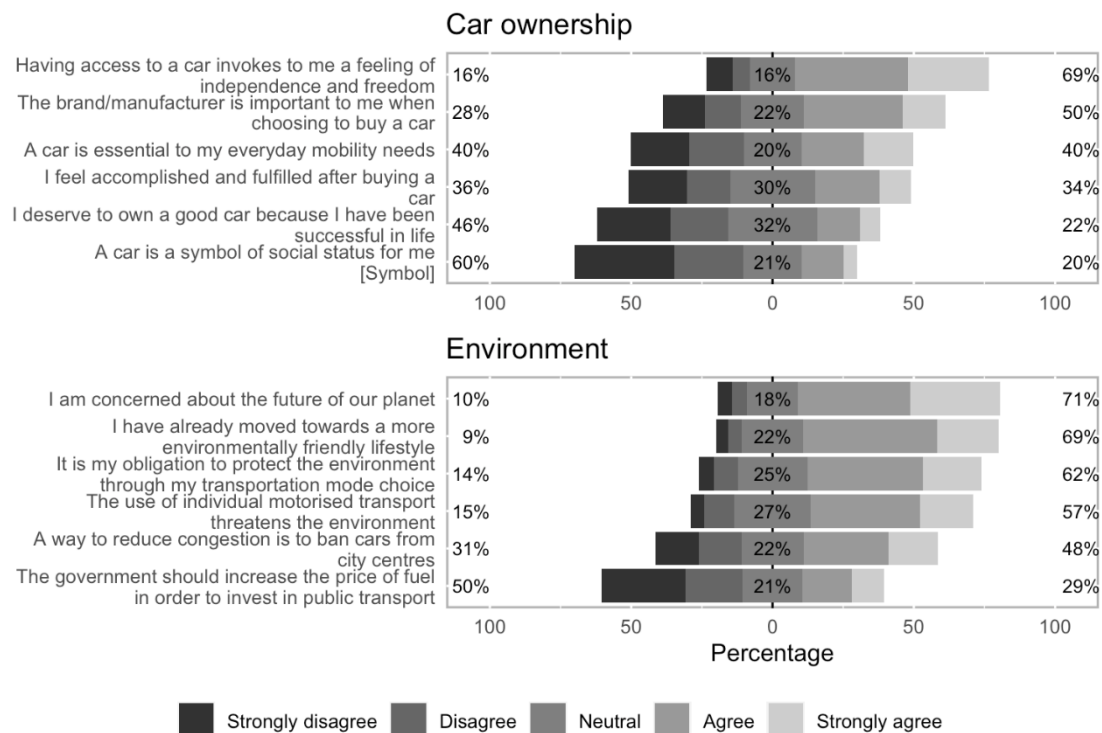


Figure 1: Responses to attitudinal questions

To check the conformity of the dataset with the assumptions of EFA, two inspection metrics are calculated. The χ^2 -statistic with 66 df, calculated by Bartlett's test of sphericity, is close to 500, which indicates that the data are not an identity matrix at 95% confidence level. Furthermore, the Kaiser-Meyer-Olkin measure of sampling adequacy is 0.88, which suggests that the data are suitable for EFA (Howard, 2016). The eigenvalues of factors suggest empirically that between two and four factors should be retained in the model. Comparing the full model (four factors) with the parsimonious model of two factors results in a loss of explained variance of 0.05, while gaining in interpretation. Table 3 shows the factor loadings of the EFA model with two factors after oblique rotation, which results in an interpretable structure when the factors are correlated. It can be seen that:

- 1) All variables about environmentally friendly travel behaviour were associated with Factor 1 and
- 2) All variables related to car use and ownership were associated with Factor 2.

Therefore, Factor 1 is interpreted as "Pro-Environment" and Factor 2 as "Pro-Car" attitudes. We expect pro-environment attitudes to be associated with higher willingness to accept any of the measures, while attitudes in favour of cars are could be more resistant to changes.

Table 3: Factor loadings for attitudes towards car ownership and the environment (loadings ≤ 0.3 not presented, ≥ 0.7 highlighted)

Variable	Factor 1	Factor 2
Symbol		0.718
Independence		0.605
Brand		0.686
Success		0.820
Accomplishment		0.828
Essential		0.531
Threat	0.741	
Protection	0.834	
Fuel Price	0.647	
Carfree	0.713	
Future	0.727	
Lifestyle change	0.673	
Summary statistics		
Proportional variance	0.269	0.251
Cumulative variance		0.520
χ^2 -statistic 644.12		
Factor interpretation	<i>Pro-Environment</i>	<i>Pro-Car</i>

Modelling the acceptance

First, we model the acceptance by binomial regression using as explanatory variables the measurable travel behaviour and the socio-demographics of the respondents. We check for separation effects in terms of the levels of the response variable and remove all instances with less than five occurrences in the sample. Non-significant variables with the appropriate sign are grouped with significant variables to enhance the interpretability of the model ($AIC = 1170.9$, McFadden pseudo- $R^2 = 0.45$). The main findings are summarised below:

- (1) Age: people 30-59 years old seem to disregard the measures, a possible explanation being that individuals in working age have less time to investigate alternatives and habitually select one mode of transport.
- (2) Education: basic education (finished high school) and other types of education, e.g. professional training in Germany, are also connected with lower willingness to accept. Those education types are probably associated with out-of-office activities that require freedom of mobility.
- (3) Occupation: students seem to accept the measures, while housewives and househusbands seem to disregard them. Students often use public transport, on the contrary, housewives and househusbands travel often for shopping trips, where avoiding the car is not always an alternative.
- (4) Household income: medium to higher income households (4000-7000 € per month) tend to accept the measures; due to their financial flexibility, they are likely willing to try alternatives.
- (5) Duration of residence: living in Munich for 1-3 years is associated with the adoption of measures, while other durations did not result in any significant relation.
- (6) Subscription for public transport: regular users indicate their agreement with the measures, which also aim to shift much of the demand to public transport.

- (7) Driving license: the positive association may relate to the fact that the vast majority of respondents hold a license. In contrast, people without a license are probably driven around by others or belong to a population segment that was not captured well by this survey.
- (8) Home office: 2-5 days per week seems to contribute to the acceptance, because people do not have to change their commute patterns. By contrast, no conclusion can be drawn for those who work remotely one day per week or less.
- (9) Modes of transport: active transportation, such as walking and bike, are associated positively with the measures. As expected, using a car for leisure trips relates to a negative impact. Furthermore, commuters to work with travel companions are also more likely to accept, possibly because they have a lower sensitivity to changes of travel time.
- (10) Ownership of vehicles: bike owners are willing to accept the measures, which largely could improve the conditions for them, while no pattern could be identified for owners of other vehicles, including private cars, e-scooters and cargo bikes.
- (11) Use of micromobility: respondents that indicated to use on-demand micromobility regularly stated that they will accept, as the measures aim to improve conditions for micromobility too.

No meaningful relation was found for *disabled* people, which were expected to vastly disregard the measures. Additionally, households *owning* or *planning to buy* a private vehicle were not related with the acceptance of the measures. The *household size* did not impact the acceptance of the models, despite our expectation that having children could impact the acceptance negatively.

The detailed results of the binomial model are not given. Instead, we add the latent factors Pro-Environment and Pro-Car and create a structural equation model to reveal the correlation between the underlying attitudes and the acceptance of the car-reducing measures. Although the coefficient estimates of the measurable characteristics change slightly in comparison to the binomial model, they remain consistent in sign and magnitude (Table 4).

Table 4: Structural modelling results

Indicator	Estimate	Std. Error	t-stat.
Regressions			
Age ₃₀₋₅₉	-0.27	0.10	-2.85**
Education _{High School}	-0.18	0.09	-1.95 .
Education _{Other}	-0.72	0.27	-2.66**
Occupation _{Student}	0.55	0.29	1.93 .
Occupation _{Housewife/husband}	-0.53	0.25	-2.09**
Household income _{4000-7000 €/month}	0.21	0.10	2.12**
Home office _{2-5 days}	0.24	0.09	2.60**
Resident _{1-3 years}	0.60	0.28	2.13**
Subscription _{Public Transport}	0.28	0.11	2.60**
Driving License	0.35	0.14	2.49**
Car _{to Leisure}	-0.42	0.10	-4.26***
Public transport _{to work}	0.43	0.12	3.76***
Bike _{to Work}	0.47	0.17	2.82**
Bike _{to shopping}	0.27	0.14	1.94 .
Walk _{to Shopping}	0.22	0.10	2.15**
Car with companion _{to Work}	0.94	0.29	3.29**
Own Bike	0.22	0.10	2.13**
Use micromobility	0.70	0.12	5.76***
<i>Pro-Environment</i>	<i>0.71</i>	<i>0.05</i>	<i>13.82***</i>
<i>Pro-Car</i>	<i>0.08</i>	<i>0.06</i>	<i>1.32</i>
Covariances			
<i>Pro-Environment ~ Pro-Car</i>	<i>-0.17</i>	<i>0.01</i>	<i>-12.57***</i>
Summary statistics			
R^2	0.537		
χ^2 -statistic	2838.958 with 279 df		
CFI	0.904		
TLI	0.973		
RMSEA	0.078, 90% CI [0.076, 0.081]		

Significance: 0 '****' 0.001 '***' 0.05 '.' 0.1

The results suggest that pro-environmental attitudes partially explain the willingness to accept car-reducing measures. On the other hand, attitudes in favour of owning and using a private car do not necessarily associate with the rejection of the proposed measures. This is indicated by the *t*-statistic, which is large in the case of Pro-Environment, meaning that the null hypothesis of the coefficient estimate being equal to zero can be rejected, whereas it is low for Pro-Car, meaning that there is not enough evidence to reject the null hypothesis at a confidence level of 90% at least. Overall, the assumed model is valid as indicated by the Comparative Fit (CFI) and the Tucker-Lewis (TLI) indices being over 0.9.

4. CONCLUSIONS

In this short paper, citizens' acceptance of car-reducing measures is assessed using data from a survey. Through explanatory factor analysis, the responses of individuals to questions relating to

the environment and to private cars were clustered into latent attitudinal constructs. Subsequently, binomial regression and structural equation modelling reveal that pro-environmental attitudes are affiliated with a higher willingness to accept the measures, whereas attitudes associated with owning and using a car do not provide sufficient evidence against them. Other key factors include medium to high household income, possibility to work from home, public transport subscription, driving license and the habitual use of active modes of transport and micromobility.

A methodological limitation is that the presented models do not account for interactions between variables. Further, some answers in this study might be biased; the respondents were not aware of the goal of this project, which was to achieve modal shift towards alternative modes of transport but may have been able to infer this from the context and the formulations of the four bundles of measures. In the future, further effort should go into understanding which are the factors that influence each measure separately.

Overall, this work can be seen as a tool to target the relevant audiences when local authorities take decisions to curb motorised traffic. They can either target the mentioned population groups and maximise their acceptance of the measures or nudge the groups that would otherwise not accept the measures and shift their grounds.

ACKNOWLEDGEMENTS

This study has been conducted within the project “MCube aqt” with funding from the Federal Ministry of Education and Research of Germany.

REFERENCES

- Banister, D. (2008). The sustainable mobility paradigm. *Transport Policy*, 15(2), 73-80.
- Becker, H., Ciari, F., & Axhausen, K. W. (2018). Measuring the car ownership impact of free-floating car-sharing – A case study in Basel, Switzerland. *Transportation Research Part D*, 65, 51-62.
- Eriksson, L., Garvill, J., & Nordlund, A. M. (2008). Acceptability of single and combined transport policy measures: The importance of environmental and policy specific beliefs. *Transportation Research Part A*, 42(8), 1117-1128.
- Federal Statistical Office. (2011). *Zensus 2011*. Bonn, Germany.
- Friman, M., Larhult, L., & Gärling, T. (2013). An analysis of soft transport policy measures implemented in Sweden to reduce private car use. *Transportation*, 40, 109-129.
- Gärling, T., Gärling, A., & Johansson, A. (2000). Household choices of car-use reduction measures. *Transportation Research Part A*, 34, 309-320.
- Gonzalez, J. N., Gomez, J., & Vassallo, J. M. (2022). Do urban parking restrictions and Low Emission Zones encourage a greener mobility? *Transportation Research Part D*, 107, 103319.
- Howard, M. C. (2016). A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? *International Journal of Human-Computer Interaction*, 32(1), 51-62.
- infas, DLR, IVT & infas 360. (2018). *Mobilität in Deutschland*. Federal Ministry of Digital and Transport, Bonn, Germany.
- Loukopoulos, P., Jakobsson, C., Gärling, T., Schneider, C. M., & Fujiji, S. (2005). Public attitudes towards policy measures for reducing private car use: evidence from a study in Sweden. *Environmental Science & Policy*, 8(1), 57-66.
- Romero, F., Gomez, J., Rangel, T., & Vassallo, J. M. (2019). Impact of restrictions to tackle high pollution episodes in Madrid: Modal share change in commuting corridors. *Transportation Research Part D*, 77, 77-91.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36.
- Schmid, B., Schmutz, S., & Axhausen, K. W. (2016). Explaining mode choice, taste heterogeneity, and cost sensitivity in a post-car world. *Transportation Research Board 95th Annual Meeting*.
- Yan, X., Levine, J., & Marans, R. (2019). The effectiveness of parking policies to reduce parking demand pressure and car use. *Transport Policy*, 73, 41-50.

Word count

Abstract: 138

Short paper: 2994

An activity-based latent class modelling approach to assess the impact of hybrid working on travel demand in the Netherlands after COVID-19

H. Zhou^{*1}, Y. Araghi¹, B. Ashari¹, and M. Snelder^{1, 2}

¹Scientific researcher, Sustainable Urban Mobility and Safety, Dutch Applied-Science Organization (TNO), 2595 DA The Hague, The Netherlands

²Associate Professor, Delft University of Technology, 2628 CN Delft, The Netherlands

SHORT SUMMARY

After COVID some employees can continue to work from home or at their work location. This hybrid way of working can impact transport demand and traffic conditions. Current models can not fully capture mobility patterns caused by hybrid working. We developed a dedicated latent class hybrid working model to predict which individuals will choose to WFH and how frequently they will WFH and integrated it into an activity-based model. We illustrate the potential of the model by simulating travel demand in a metropolitan region in the Netherlands. The results show that under some scenarios hybrid working can reduce mobility demand but under other scenarios these gains in work-home travel is lost by additional activities.

Keywords: activity-based, travel demand, hybrid working, latent class

1 INTRODUCTION

After the COVID19-pandemic, employees were allowed to continue to partially work from home (WFH) and partially at office, thus providing a hybrid way of working. However, the level of impact of hybrid working on the mobility patterns remains to be fully investigated.

Caldarola & Sorrell (2022) studied hybrid working in England and indicated that it leads to fewer commutes but not necessarily reducing the distance travelled by employees. In the United States, the traffic worsened because of cuts in the transit network (resulting in less public transport) during the pandemic (Mack et al. (2021)) and more solo driving. The vehicle kilometres travelled in July 2020 were restored to 104% of pre-COVID levels in NYC (Wang et al. (2021)). A survey conducted in Melbourne, Australia also reported increased car usage post pandemic (Currie et al. (2021)). In the Netherlands, 27% of hybrid workers expected to WFH more often in the future, according to a study using the Netherlands Mobility Panel 2020 (de Haas et al. (2020)). A 2020 survey (*MenE*-team (2020)) by the Dutch Ministry of Infrastructure and Water Management shows a similar pattern. Both surveys have shown that people prefer to use cars, bikes and walking than pre-pandemic.

Since hybrid working potentially impacts transport demand and traffic conditions (Beck & Hensher (2022)), it is important to understand its role in mobility patterns. However, current traffic and transport models do not capture the extra activities that employees may do while working from home, which leads to inaccurate mobility assessments and traffic management, which may cause errors in decisions in congestion management and for large infrastructural investments.

Hybrid working may depend on many factors e.g. the type of work, socio-demographic attributes, living/work locations, employer's willingness to allow WFH. Since the decision to WFH is largely person-specific, it fits well with the domain of activity-based modelling (ABM), where detailed personal and household data is used to predict daily activity and travel schedules. The schedule includes the individuals' mobility patterns, where and when they are carried out, and the travel modes used.

The study by Cruz (2021) analysed the impact of COVID-19 on travel behaviours and in-home activities using ABM. However, the method used in that study has not yet fully integrated within an ABM and uses aggregated values for activity choices. These may underestimate the impact of hybrid working on people's destination choices, travel patterns and joint activities among household members. A case study by Wang et al. (2021) in New York City, using MATSim (Horni et al. (2016)), captured the preferences of WFH by updating the mode choice utility functions for the synthetic population and the travel schedules are modified to have the suitable WFH ratio based

on Dingel & Neiman (2020) and GTFS (General Transit Feed Specification) data to reflect its effect.

To the best of our knowledge, a dedicated ABM model determining the individual’s choice of hybrid working is missing in the literature. In this paper we fill this gap by initially applying latent class models and segment employees regarding their level of hybrid working, using empirical data from the Netherlands Working Conditions Survey (NWCS Hooftman et al. (2020)). The model aims to capture the heterogeneity of individuals and take gender, size of the company, work sector, household income, urbanization degree and age into account when creating latent segments of employees based on their decision to WFH.

Next, we use the latent class hybrid working model outputs to integrate them within an existing ABM framework. The improved ABM model has the capability of evaluating the effect of hybrid working-related mobility patterns. To demonstrate the potential of our hybrid working decision model within ABM, we simulate the potential impact of hybrid working in an illustrative study in the Metropolitan Region Rotterdam The Hague (MRDH) in The Netherlands.

The remainder of this paper is organized as follows: Section 2 explains the construction of the hybrid working model using survey data, section 3 presents the estimation results, and the results of an illustrative example using ABM model with the integrated hybrid working model. Finally, Section 4 presents the conclusions, discussion, and recommendations for future research.

2 METHODOLOGY

The latent class hybrid working model has been developed as a component of an agent-based model (ABM). To explain how this model interacts with an ABM framework, we use a specific framework called ActivitySim (Gali et al. (2008)). Using a population synthesizer (Snelder et al. (2021)), the ABM determines individuals’ work or school locations, their level of hybrid working, and their daily activity patterns (DAP). Based on this information, the model predicts the number of tours that an individual will undertake in a given day, as well as the number of stops in each tour. This includes information about the start time, duration, destinations, and modes of each tour. The trip mode chosen at this stage is considered the main mode. Next, our tour-based mode chain choice model determines the access and egress modes to generate a feasible trip mode combination for each tour (Zhou et al. (2023)).

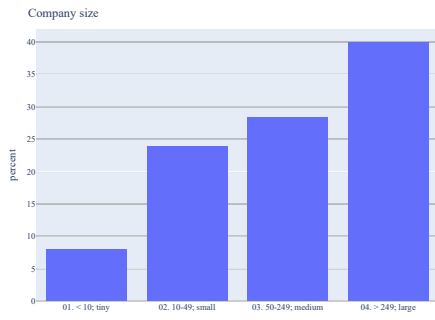
Survey data

Data used to develop the models in this study are taken from NWCS (Hooftman et al. (2020)), a periodic survey carried out jointly by TNO and CBS and focusing on the labour situation among Dutch employees since 2003. It provides information on the working conditions, employability and health of a representative sample of the working population (age range between 15 and 75) in The Netherlands. Since the COVID-19, NWCS surveys added questions, amongst others, about employees’ expectations of working from home. The survey from November 2021 has been adopted to reflect better people’s opinions on the number of hours WFH at the time of the survey and in the future.

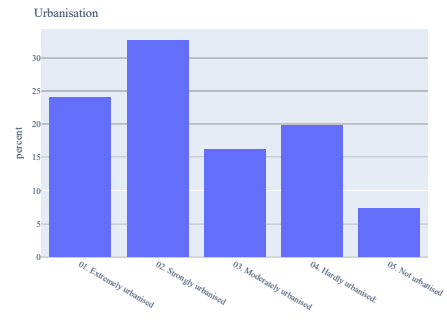
We have filtered out those respondents who did not complete their desired days of WFH in the future (post-pandemic), which resulted in 6359 respondents being used for latent class estimation. The sample distributions of several relevant socio-demographic and work attributes are explored (Figure 1).

Latent Class Cluster Analysis (LCCA)

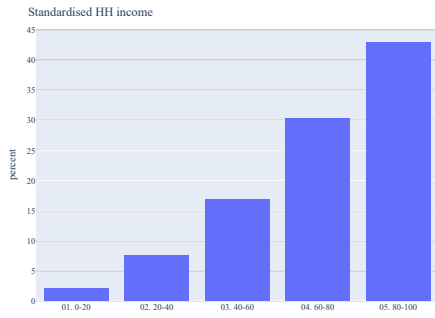
We used LCCA (Vermunt & Magidson (2004)) to group individuals into different latent classes based on their responses to observed indicators (Molin et al. (2016)), which we call manifest indicators. The goal is to create latent segments based on the available data to maximize the homogeneity within the latent classes and the heterogeneity between clusters. Using the LCCA method, one can predict a probability of a respondent belonging to a particular class. We used Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) criterion, Chi-Squared and Log-likelihood ratio test as indicators to determine the best model fit (Oberski et al. (2013)). Furthermore, the LCCA models can incorporate covariates, which in this case are the socio-demographic characteristics of individuals. These covariates are used as additional predictors of class membership. This is based on the probability of observing a particular sequence of responses,



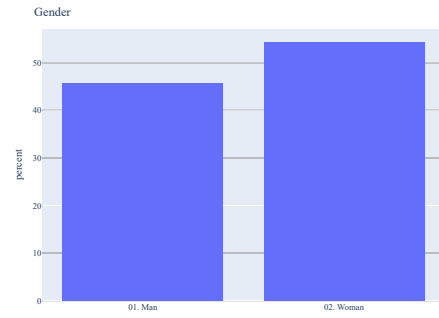
(a)



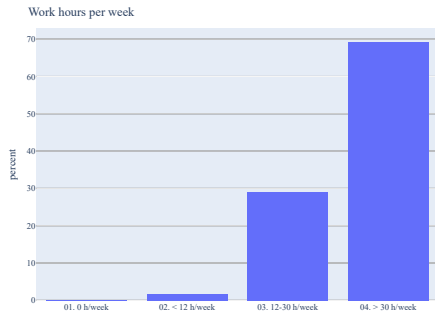
(b)



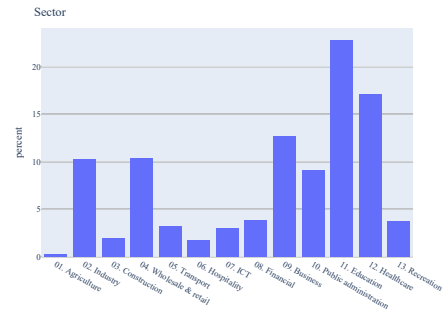
(c)



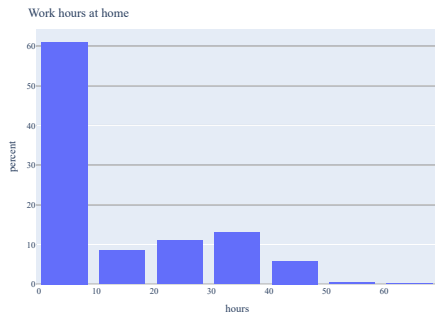
(d)



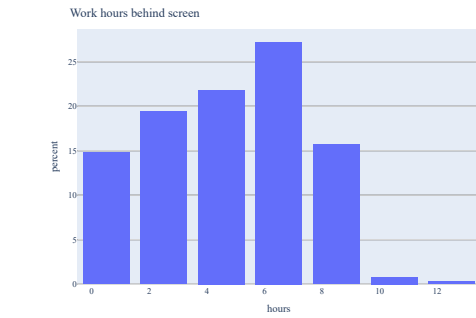
(e)



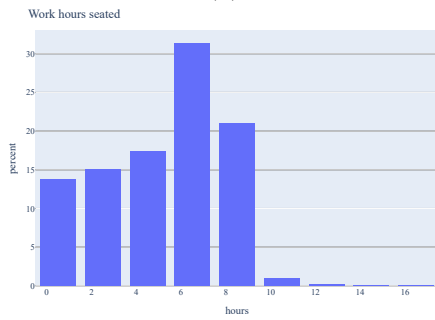
(f)



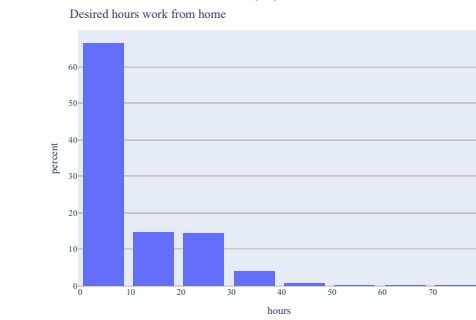
(g)



(h)



(i)



(j)

Figure 1: Distribution of several attributes

i.e. the response pattern, on the questions asked from the respondents.

This research is posited on the assumption that various groups exist in the population that have different approaches towards hybrid working, which are determined by working situations, e.g. less active work (such as sitting behind a desk) or being active at work or having the possibility to work at home (or at a distance from the employers' location) and socio-demographic covariates, e.g. age, gender, work sector, household income, urbanization degree and company size, number of contract hours per week or per month etc.

Hybrid working model within ABM

We have integrated a hybrid working component into our ABM framework using the outputs of the LCCA model described in Section 2. The the membership likelihood function is used to predict which cluster each employee belongs. The component also incorporates the probabilities of each hybrid working alternative from the LCCA model outputs and selects an alternative based on these probabilities. Once the hybrid working alternative is determined, the ABM predicts individuals' daily activity pattern (DAP), which includes mandatory activities such as work or school, non-mandatory activities, and home activities, using a multinomial logit model (MNL).

3 RESULTS

First we show the LCCA results and then present the hybrid working outcomes that is integrated in the ABM model. We define different levels of hybrid working by the number of hours/week an employee could WFH in four ordinal categories: category 1 is the ones that can not WFH, which we call "No-hybrid", category 2 are called "light-hybrid" for employees that WFH for less than 16 hours/week, category 3 was called "moderate hybrid" referring to those who WFH for 16 to 24 hours/week and category 4 is called "heavy hybrid" referring to those that WFH more than 24 hours/week.

LCCA model to determine hybrid working latent classes

The LCCA model was estimated from 1 to 6 classes, and based on the statistical criteria of Log-likelihood, BIC and AIC shown in table 1, we conclude that the 4 latent class model gives the best-fit.

Number of Classes	Log-Likelihood	BIC	AIC	Chi-square goodness of fit
1	-29081.89	58268.98	58187.89	32335.57
2	-23157.33	46638.7	46388.67	9335.125
3	-22139.75	44822.48	44403.51	5811.823
4	-21293.41	43348.72	42760.81	3113.202
5	-22576.24	46133.33	45376.47	5726.725
6	-27892.79	56985.37	56059.58	17234.99

Table 1: LCCA model fit statistics, 4 class model is selected

To estimate these LCCA models, we used 4 manifest variables and 6 demographic variables. The manifest variables are 1) the number of hours/week the employees worked at home (at the time of the execution of the survey Nov. 2021), 2) the number of hours/week the employees wished (i.e. desired) to work at home if things went back to normal and if they were able to choose, 3) the number of hours/day the employees worked behind a desk, 4) the number of hours/day the employees worked with a computer, tablet or laptop which had a screen. The socio-demographic variables used as covariates in the model were the following: 1) the size of the company the employee worked for, 2) the urbanisation level of the place of living of the employee, 3) total household income, 4) the number of work hours per week, 5) the sector to which the employee worked for, 6) the gender of the employee.

Figure 2 shows the percentage of employees among different manifest variables per latent class cluster. And Figure 3 shows percentages among different covariates per latent class.

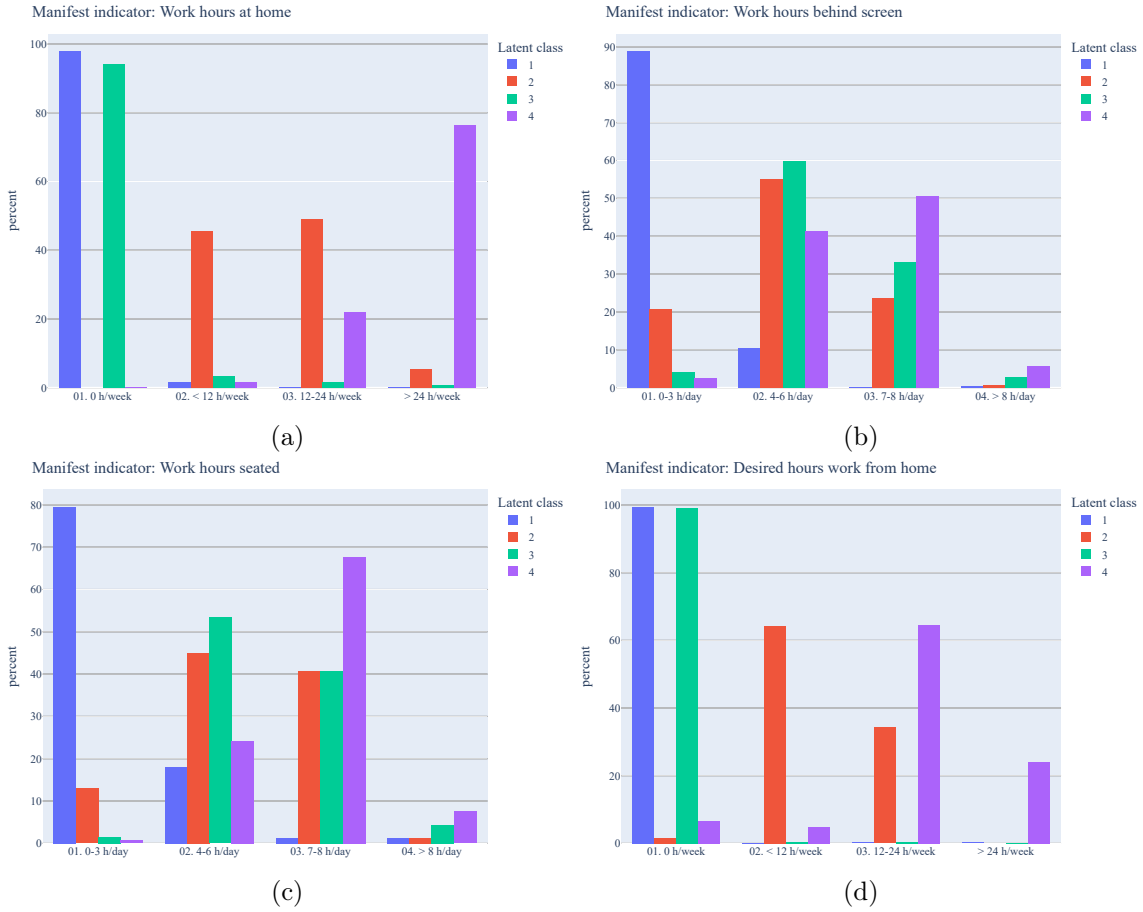


Figure 2: Distribution of manifest indicators of each latent class.

The description of each latent class is derived from the distribution of the manifest variables from figure 2. We see that latent classes 1 and 3 are mainly include employees that have not reported WFH and additionally have no intention to WFH. This is mainly due to their work types. However, these 2 classes differ from each other when it comes to work hours behind the screen and seated. Workers of classes 2 and 4 are working +2 days per week from home and intent to keep WFH but slightly less than what they were already doing at the time of survey (Nov 2021). Table 2 presents more features of each of the 4 classes and descriptions, together with the probabilities of each hybrid-working alternative per latent class.

Illustration example

In this section, we explore whether the hybrid working model leads to different travel behaviour. To do this, we run the entire ABM model for the MRDH region in the year 2022 with the integrated hybrid working component. We chose this year because the hybrid working survey was conducted at the end of 2021, which gives a reasonable prediction for 2022.

Input data

The MRDH region, located in the Netherlands, has an area of about 1130 km². The synthesized population of the region of MRDH is generated through a population generator based on data from the Dutch State Statistics (CBS) (Centraal Bureau voor de Statistiek (2020)). This synthesized population consists of 2,387,032 individuals spread over a total of 1,322,202 households. It includes characteristics such as age, vehicle ownership, education level, work sector, company size etc. The second type of data is the land use from the V-MRDH 2.6 model (Schoorlemmer (2020)). It concerns the land use of 7,011 pre-specified traffic analysis zones (TAZ) in the Netherlands. Of these, 5924 TAZs are within the MRDH; see Figure 4. Each TAZ contains information such as the number of employment places (offices, shops, etc.), the number of education places (i.e., schools), the actual area of the TAZ and its urbanisation level (i.e., the population density), the number of paid and non-paid parking spots and the average hourly parking costs.

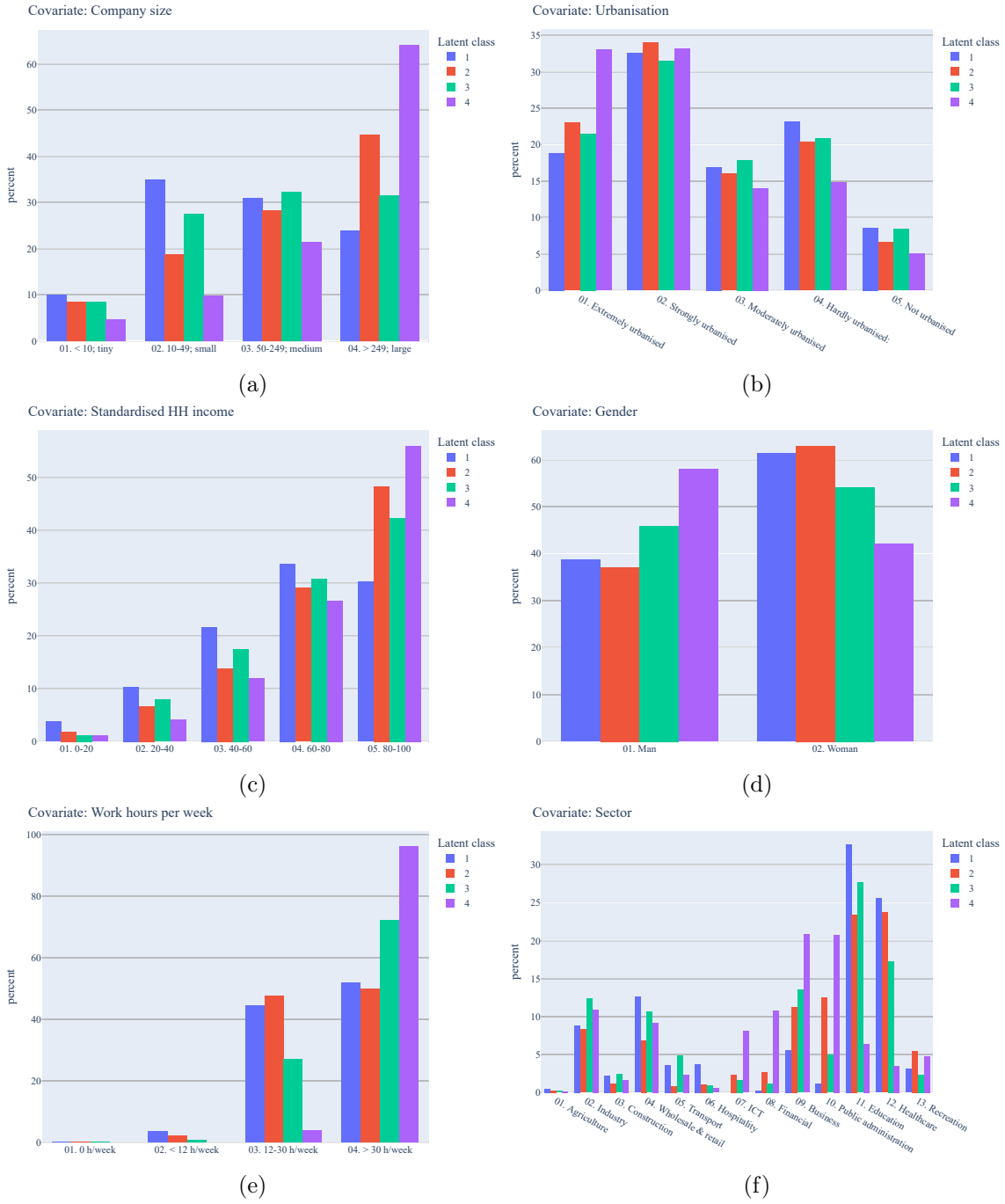


Figure 3: Distribution of covariates of each latent class.

The third type of input data concerns level-of-service data for each pair of TAZ's (origin-destination pair). For each possible pair and each of the seven unimodal travel modes (i.e. walk, bike, ebike, car, car-passenger, demand response transport and public transport (i.e., bus, metro, tram and train) that we consider, we generate travel time, cost, and distance for three different periods over the day (morning peak, evening peak, and off-peak).

Scenario description

The following three scenarios are considered::

1. The first scenario is the "Reference" which assumes that the transport system is not affected by the hybrid working since 2020.
2. The second scenario, "Hybrid working fix", which allows people to choose to work sometimes at home and sometimes in the office. in this scenario we assume that employees stay at home as much as possible and are not further mobile while working.

Class	Description	No-WFH	Light-WFH	Moderate-WFH	Heavy-WFH
1	very limited WFH, mainly working on-site limited screen & sit-work more active work types Female-dominated (61%) working in small to medium companies	99.39%	0.16%	0.27%	0.18%
2	light hybrid workers intention to WFH 2 or 3 days/week; female dominated relatively better-paid jobs in larger companies	4.22%	59.83%	35.95%	0.00%
3	very limited hybrid work mainly screen & sit-work (e.g. administrative) Working in relatively bigger companies	99.25%	0.31%	0.31%	0.12%
4	moderate to heavy WFH high behind screen & sit-work work mainly in large companies; intention to continue WFH ≥ 3 days/week male-dominated (58%)	6.02%	5.53%	64.01%	24.44%

Table 2: Description of latent classes and the corresponding probabilities of WFH for each latent class

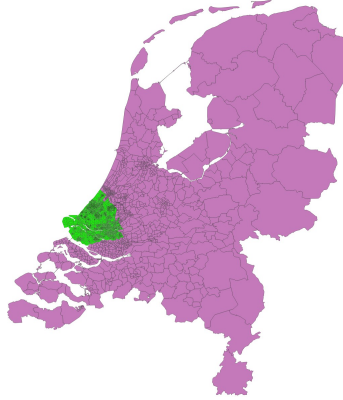


Figure 4: The Netherlands consists of 7011 TAZs, 5924 of which represent the MRDH region.

3. The third scenario, "Hybrid working flex", assumes that people have the flexibility to arrange their working time while WFH. Employees may engage in other activities, such as shopping, picking up children, or walking/cycling, in addition to working from home.

Simulation results

We use the integrated ABM model described in Section 2 to simulate the three scenarios described in the previous section. Table 3 presents the indicators derived from the simulation results of each scenario. Additionally, Figure 5 shows the number of trips per trip purpose, and Figure 6 displays the departure time distribution of all trips.

In both hybrid working scenarios (Table 3), 73% of employees in the study area cannot WFH. This is not surprising, as not all types of work can be done remotely. However, this percentage is higher than the result reported by the NWCS survey (63.1%). This discrepancy could be explained by the fact that the synthesised population in MRDH area has a lower income, works shorter hours per week, and lives in more densely populated urban areas than the main survey population. These factors may contribute to a higher percentage of people who cannot WFH in the MRDH area. Among those who can WFH, 8.1% choose to work less than 2 days per week, while 14.8% would like to WFH 2 or 3 days. Only 4.1% of employees would like to WFH more than 3 days per week. In scenario 2, we saw on average a 3% decrease in the number of trips per person in a day. This decrease is expected because in this scenario we assumed people who WFH do not further adjust their activity patterns. This reduction in trips leads to a significant decrease of 2.93% in the total

Indicators	Reference	Hybrid Working Fix (% change w.r.t. reference)	Hybrid Working Flex (% change w.r.t. reference)
Hybrid working		no hybrid: 73% light hybrid: 8.1% moderate hybrid: 14.8% heavy hybrid: 4.1%	(The same as the 'Fix' scenario)
Tours p.p	1.132	1.096 (-3.2%)	1.138 (0.5%)
Trips p.p.	2.795	2.709 (-3.1%)	2.808 (0.5%)
Total trips	6,672,069	6,465,695 (-3.1%)	6,703,805 (0.5%)
Car kilometer traveled(million km)	15.962	15.495 (-2.93%)	15.912 (-0.3%)

Table 3: Indicators

car travel distance. However, there is no remarkable difference in the distribution of the departure time of the trips(Figure 6), which is not surprising since employees who do not work at home do not adjust their departure time in the model.

In scenario 3, employees still make fewer work-home trips but more other tours/trips, especially for groceries, visiting doctors (Figure 5) as they are more flexible during working hours. The total car travelled distance is higher than scenario 2 (see Table 3), which is justifiable since people may make non-work tours/trips during their work hours instead of staying home. However, it is still lower than that of the reference scenario, which can be explained by the fact that the non-work activity destinations are closer to their homes. The departure time shifts slightly towards off-peak hours since more non-work trips are made.

Overall, we conclude that the integrated ABM provides insights into the changes in travel behaviour.

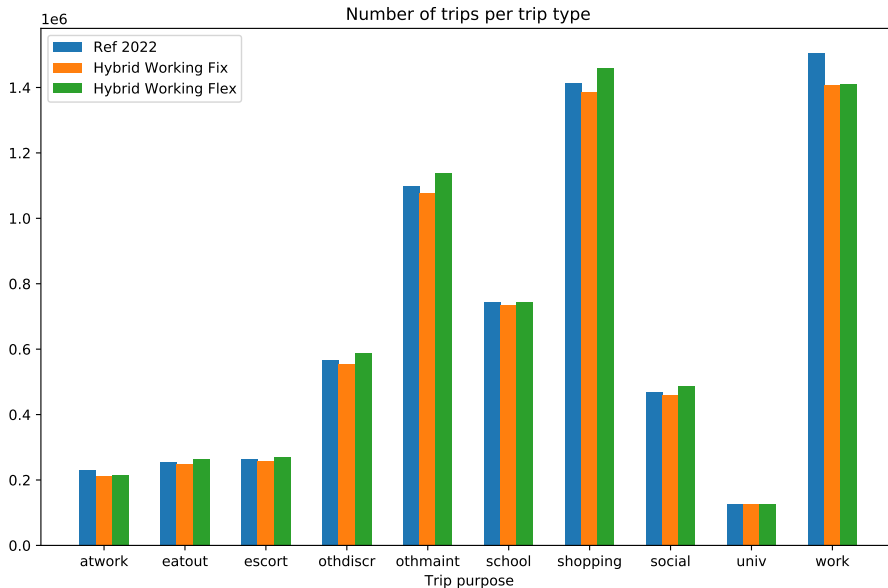


Figure 5: Number of trips per trip purpose

4 CONCLUSIONS AND DISCUSSION

In this study, we used empirical data to develop latent class clusters of employees based on their hybrid working levels. Our analysis showed that several factors, such as company size, urban area type, household income, weekly working hours, work sectors, and gender, play a crucial role in people's choice of hybrid working. The developed hybrid working model was then integrated into an ABM framework and applied to a synthesised population of the MRDH region. In the scenario

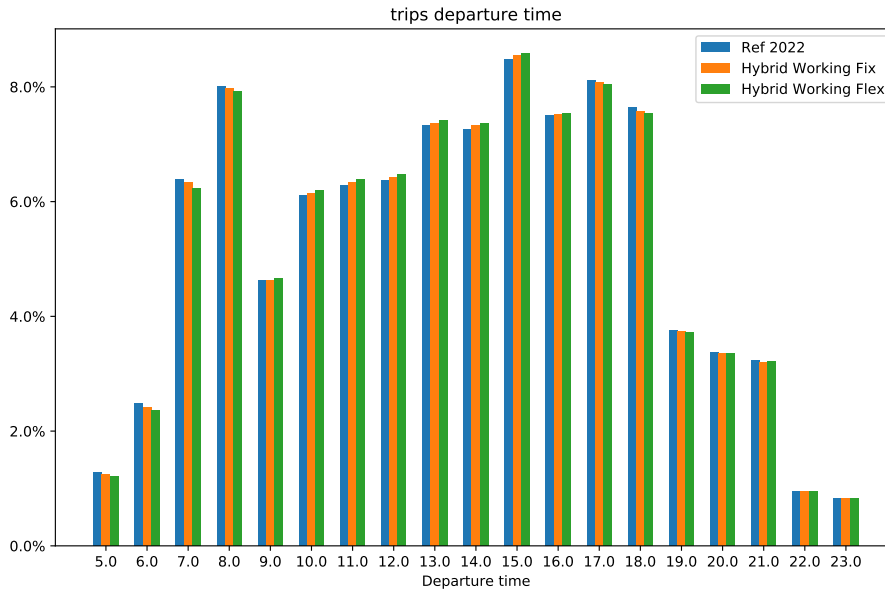


Figure 6: Distribution of trip departure time

where we assumed employees do not further travel during the day while they WFH, we saw a 2.93% reduction in car travel distance. We conclude that hybrid working has the potential to reduce the total travel demand. And policies regarding the departure time, especially during peak hours, could be explored to make further improvements on the travel demand.

In the scenario in which people who WFH are allowed to be flexible in doing other activities, the car travel distance is hardly reduced, because of an increase in shopping and other maintenance trips due to hybrid working. We conclude that this way of hybrid working (i.e. scenario 3) could positively and negatively affect travel demand. On the one hand, fewer work trips could reduce traffic congestion during peak hours. On the other hand, more non-work trips could increase overall travel demand, increasing traffic congestion during off-peak hours.

As the trend towards hybrid working is expected to continue, it is important to develop better models to evaluate its impact on cities and urban areas and inform policy decisions. In this regard, we recommend that future research focuses on calibrating the mode choices while accounting for hybrid working. Such updates could provide a more accurate representation of the impact of hybrid working. In addition, we recommend surveys to add questions about employees' mobility patterns while they work at home. Including this information in the model can increase the accuracy of the whole ABM model.

Notably, there is no significant difference in the distribution of departure times of trips. This could imply that travellers not working from home are not adjusting their departure times, which could be a potential area for further exploration in future studies.

ACKNOWLEDGEMENTS

This research was supported by TNO Dragon's Den program 2022. The authors would like to thank Wendela Hooftman, Reinier Sterkenburg and Paco Hamer for their contributions.

REFERENCES

- Beck, M. J., & Hensher, D. A. (2022, nov). Australia 6 months after COVID-19 restrictions part 2: The impact of working from home. *Transport Policy*, 128, 274–285. doi: 10.1016/j.tranpol.2021.06.005
- Caldarola, B., & Sorrell, S. (2022, may). Do teleworkers travel less? evidence from the english national travel survey. *Transportation Research Part A: Policy and Practice*, 159, 282–303. doi: 10.1016/j.tra.2022.03.026

- Centraal Bureau voor de Statistiek. (2020). *Results based on calculations by tno using non-public microdata from statistics netherlands* (Tech. Rep.). Retrieved from <https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/catalogus-microdata>
- Cruz, V. A. (2021). *An activity-based modeling approach to assess the effects of activity-travel behavior changes and in-home activities on mobility* (Unpublished master's thesis). TU Delft.
- Currie, G., Jain, T., & Aston, L. (2021). Evidence of a post-covid change in travel behaviour—self-reported expectations of commuting in melbourne. *Transportation Research Part A: Policy and Practice*, 153, 218–234. doi: 10.1016/j.tra.2021.09.009
- de Haas, M., Faber, R., & Hamersma, M. (2020). How COVID-19 and the dutch ‘intelligent lockdown’ change activities, work and travel behaviour: Evidence from longitudinal data in the netherlands. *Transportation Research Interdisciplinary Perspectives*, 6, 100150. doi: 10.1016/j.trip.2020.100150
- Dingel, J., & Neiman, B. (2020, apr). How many jobs can be done at home? doi: 10.3386/w26948
- Gali, E., Eidenbenz, S., Mniszewski, S., Cuellar, L., & Teuscher, C. (2008). *ActivitySim: large-scale agent based activity generation for infrastructure simulation*.
- Hooftman, W. E., Mars, G. M. J., Knops, J. C. M., van Dam, L. M. C., de Vroome, E. M. M., Janssen, B. J. M., ... van den Bossche, S. N. J. (2020). *Nationale Enquête Arbeidsomstandigheden 2019. Methodologie en globale resultaten* (Tech. Rep.). Leiden: TNO. Retrieved 2022-10-14, from <https://repository.tno.nl//islandora/object/uuid:1716ea15-af48-4b47-8b0a-db3bbb7debbf>
- Horni, A., Nagel, K., & Axhausen, K. (Eds.). (2016). *Multi-agent transport simulation matsim*. London: Ubiquity Press. doi: 10.5334/baw
- Mack, E. A., Agrawal, S., & Wang, S. (2021). The impacts of the covid-19 pandemic on transportation employment: A comparative analysis. *Transportation Research Interdisciplinary Perspectives*, 12, 100470. doi: <https://doi.org/10.1016/j.trip.2021.100470>
- MenE-team. (2020, December). *Achtergrondrapportage ‘monitoring mobiliteiten vervoer’* (techreport No. 35). Ministerie van Infrastructuur en Waterstaat.
- Molin, E., Mokhtarian, P., & Kroesen, M. (2016). Multimodal travel groups and attitudes: A latent class cluster analysis of dutch travelers. *Transportation Research Part A: Policy and Practice*, 83, 14–29.
- Oberski, D. L., van Kollenburg, G. H., & Vermunt, J. K. (2013). A monte carlo evaluation of three methods to detect local dependence in binary data latent class models. *Advances in Data Analysis and Classification*, 7, 267–279.
- Schoorlemmer, S. (2020, February). *Verkeersmodel V-MRDH 2.6 Een addendum op de technische documentatie van V-MRDH 2.0 en 2.4* (Tech. Rep. No. 005556.20200218.N1.01). Deventer (The Netherlands): Goudappel Coffeng. Retrieved from <https://www.mrdh.nl/project/verkeersmodel>
- Snelder, M., Araghi, Y., Ashari, B., Charoniti, E., Klunder, G., Sterkenburg, R., ... de Romph, E. (2021). *Rapport A: Methode Urban Tools Next II - toelichting op gekozen aanpak voor parkeren, ketens en hubs, nieuwe mobiliteitsconcepten*. Den Haag (The Netherlands). Retrieved from <https://publications.tno.nl/publication/34639979/i0d13z/TNO-2021-R10644.pdf>
- Vermunt, J. K., & Magidson, J. (2004). Latent class analysis. *The sage encyclopedia of social sciences research methods*, 2, 549–553.
- Wang, D., Tayarani, M., He, B. Y., Gao, J., Chow, J. Y., Gao, H. O., & Ozbay, K. (2021). Mobility in post-pandemic economic reopening under social distancing guidelines: Congestion, emissions, and contact exposure in public transit. *Transportation Research Part A: Policy and Practice*, 153, 151–170. doi: 10.1016/j.tra.2021.09.005

Zhou, H., Dorsman, J., Mandjes, M., & Snelder, M. (2023, mar). Sustainable mobility strategies and their impact: a case study using a multimodal activity based model. *Case Studies on Transport Policy*, 11, 100945. doi: 10.1016/j.cstp.2022.100945

The stochastic multiple depot electric vehicle scheduling problem with recourse

Léa Ricard^{*1}, Guy Desaulniers², Andrea Lodi³, and Louis-Martin Rousseau²

¹Department of Computer Science, Université de Montréal, Montréal, Canada

²Department of Mathematics and Industrial Engineering, Polytechnique Montréal, Montréal, Canada

³Jacobs Technion-Cornell Institute, Cornell Tech and Technion - IIT, New York, USA

SHORT SUMMARY

Zero emission policies in urban centers are promoting the conversion of transit agencies fleets to battery electric buses (BEBs). This transition raises questions about battery management and more specifically about the best way to mathematically model this resource in order to respect energy feasibility constraints while being as little conservative as possible. In an attempt to partially answer these questions, this work presents a two-stage stochastic model with recourse for the multiple depot electric vehicle scheduling problem with stochastic travel time and energy consumption (S-MDEVSP). Vehicles are allowed to be partially recharged and a non-linear charging function is considered. Our model takes advantage of the full information on the current state of charge that is available in operation by allowing planned charge time to be extended when energy consumption deviations are observed. We propose a column-generation-based heuristic featuring stochastic pricing problems to solve a real-life instance from the city of Montréal, Canada. An analysis of the relevance of our approach for different commercially available BEBs is also provided.

Keywords: column generation, electric bus, two-stage stochastic program, vehicle scheduling

1 INTRODUCTION

The multiple depot electric vehicle scheduling problem (MDEVSP) is an extension of the multiple depot vehicle scheduling problem with additional limitations, including shorter driving range, longer refueling time, and special charging infrastructure. It aims at finding a set of vehicle routes that covers each timetabled trip exactly once while minimizing the operational costs and respecting energy feasibility and depot capacity constraints. These vehicle routes are subject to operational uncertainties (e.g., traffic jams, extreme weather conditions, or special happenings in the city) that impact travel time and energy consumption. Nevertheless, the MDEVSP is generally solved without taking these uncertainties into account. This strong assumption may compromise schedule adherence and lead to solutions with sub-optimal true costs (including recourse costs). A simple way to guarantee energy feasibility is to adopt a robust optimization approach, i.e., ensuring that energy feasibility is respected for the worst case energy consumption scenarios (see for example the work of Bie et al. (2021)). Some less conservative approaches, that we group into stochastic optimization (Li et al., 2021), robust optimization with cardinality constrained set (Jiang et al., 2021), and dynamic optimization (Tang et al., 2019), have been proposed in the literature to address the MDEVSP with uncertain travel time and/or energy consumption.

This work presents the first stochastic model for the MDEVSP with stochastic travel time and energy consumption (S-MDEVSP). We formulate the S-MDEVSP as a two-stage stochastic program and introduce a recourse policy to recover energy feasibility when the vehicle routes outputted a priori turn out to be infeasible. The main idea of our approach is to take advantage of the fact that charging time can be adjusted from day-to-day to cope with energy consumption deviations. This flexibility in the charging time could allow us to output less conservative vehicle routes than the robust optimization approach while guaranteeing energy feasibility. However, this flexibility may also induce delays. To control the build-up of delays, that can also be caused by travel time deviations, we add a penalty for delays in the objective function as in Ricard et al. (2022). Our objective is to assess the relevance of our two-stage stochastic model for commercially available battery electric buses (BEBs). Precisely, we want to verify if a substantial reduction in the optimal

fleet size can be archived by introducing a recourse policy. We propose a branch-and-price algorithm to solve this challenging optimization problem and test our solution approach on a real-life instance of the city of Montréal.

This paper is organized as follows. Section 2 deals with the problem definition and a two-stage stochastic program is introduced. We devise a method compute the second stage cost analytically. A column generation-based solution approach is presented in Section 3. We present the results of computation in Section 4 and discuss the relevance of our approach for different commercially available BEBs. Our main conclusions are stated in Section 5.

2 MATHEMATICAL MODEL

Let a timetable of trips \mathcal{T} , where trip $i \in \mathcal{T}$ is schedule to start at d_i , a set of depots \mathcal{D} , such that $|\mathcal{D}| \geq 2$, and a set of charging stations \mathcal{Q} . Each charging station $q \in \mathcal{Q}$ is time-expanded as $\tilde{q} = \{qs_1, \dots, qs_k\}$, where s_1, \dots, s_k are k time intervals of ρ minutes. We denote the set of time-expanded charging stations $\tilde{\mathcal{Q}}$. The S-MDEVSP is defined on the acyclic connection-based networks $G^d(V_d, A_d)$, for $d \in \mathcal{D}$, with node set $V_d = \mathcal{T} \cup \{n_0^d, n_1^d\} \cup \tilde{\mathcal{Q}}$, where n_0^d and n_1^d represent depot d at the beginning and the end of the day, respectively, and arc set A_d . Given the probability mass function (PMF) with finite supports of the travel time ($h_i(t)$) and the PMF of the energy consumption ($e_i(\mu)$) of each timetabled trip $i \in \mathcal{T}$ as well as the travel time t_{ij} and the energy consumption e_{ij} between the end location of node i and the start location of node j , for all pairs of nodes $i, j \in V_d$, the first stage problem of the S-MDEVSP consists of finding a priori set of vehicle routes \mathcal{R}^* that covers exactly once each trip $i \in \mathcal{T}$ and respects the capacity b_d of each depot $d \in \mathcal{D}$. A vehicle route is defined as a sequence of timetabled trips and time-expanded charging nodes starting and ending at a depot $d \in \mathcal{D}$. The amount of energy recharged at each time-expanded charging node included in a vehicle route is derived from a piecewise linear function similar to the one used in Montoya et al. (2017). In the second stage, the travel time and energy consumption values are revealed and the a priori plan is modified with respect to a recourse policy to guarantee energy feasibility. A vehicle route is considered feasible if the state of charge (SoC) of the BEB never falls below σ^{min} (e.g., 0%), or if one or several recourse actions can be taken to regain energy feasibility. A recourse action is taken at the second stage if the SoC of a bus is under ω (e.g., 50%) after a charging activity. It consists in extending the charging activity by one or several time intervals in order to reach a SoC of at least ω .

Our model for the S-MDEVSP uses the following notation. Let \mathcal{R} be the set of all feasible vehicle routes, \mathcal{R}^d be the subset of these routes starting and ending at the depot d , y_r be a binary variable equal to 1 if vehicle route r is selected, and a_{ir} be a binary parameter equal to 1 if route r covers trip $i \in \mathcal{T}$. The S-MDEVSP can be formulated as the following integer linear program:

$$\min \quad \sum_{r \in \mathcal{R}} \bar{c}_r y_r \quad (1)$$

$$\text{s.t.} \quad \sum_{r \in \mathcal{R}} a_{ir} y_r = 1, \quad \forall i \in \mathcal{T} \quad (2)$$

$$\sum_{r \in \mathcal{R}^d} y_r \leq b_d, \quad \forall d \in \mathcal{D} \quad (3)$$

$$y_r \in \{0, 1\}, \quad \forall r \in \mathcal{R}, \quad (4)$$

where $\bar{c}_r = c_r + \beta \mathbb{E}[W_r(t, \mu)]$ is the expected cost of vehicle route r , c_r is the operational costs of r , β is a weighting factor, and $\mathbb{E}[W_r(t, \mu)]$ is the expected second-stage cost of r . This latter cost penalizes the delay a passenger is likely to encounter in route r . Specifically, $\mathbb{E}[W_r(t, \mu)] = \sum_{i \in r \cap \mathcal{T}} \alpha_i \mathbb{E}(X_i^r)$, where α_i is the relative passenger flow (or demand volume) on timetabled trip i and X_i^r is the secondary delay of timetabled trip i covered by route r (in minutes). A vehicle route r may be delayed because the travel times of its trips deviate from the planned time, because buffer times before trips are not sufficient, or because recourse actions are required. By adjusting the weighting factor β , one can find solutions with different trade-offs between operational costs and the expected second-stage cost. In general, the larger the β the more reliable the S-MDEVSP solutions. Analytical equations to compute in the first stage $\mathbb{E}[W_r(t, \mu)]$, for all $r \in \mathcal{R}$ generated, are developed in the following two sections.

Probability of using a recourse action.

Consider a vehicle route $r = (1, 2, \dots, i, i+1, \dots, j-1, j, \dots, n)$ with trips i and $j \in \mathcal{T}$ interspersed by a charging activity of $j-i$ time intervals (i.e., $i+1, i+2, \dots, j-1$ are time-expanded charging nodes). Let $m_j^r(z)$ be the PMF with finite supports of the SoC of bus r just before trip j . The probability of not having to extend the charging time is

$$\Pr\{0 \text{ extra charge periods before } j \in r\} = \sum_{z=\omega}^{100} m_j^r(z), \quad (5)$$

and the probability of having to extend the charging time of ϕ charge periods is

$$\begin{aligned} \Pr\{\phi \text{ extra charge periods before } j \in r\} &= \Pr\{z | \Lambda_\omega(z) = \phi\} \\ &= \sum_{z=\sigma^{min}}^{\omega-1} m_j^r(z) [\Lambda_\omega(z) = \phi], \quad \phi = 1, 2, \dots, k, \end{aligned} \quad (6)$$

where $\Lambda_\omega(z)$ is a function outputting the minimum number of additional charge time periods to be performed when the initial SoC of a BEB is equal to z in order to get an updated SoC of at least ω . We use the Iverson bracket (Iverson, 1962) notation (i.e., $[P]$ is equal to 1 if P is true and 0 otherwise).

Delay propagation.

Let $f_i^r(y)$ be the PMF with finite supports of the actual start time of activity i assigned to route r and $g_i^r(x)$ be the PMF with finite supports of X_i^r , i.e., of the secondary delay of trip i , such that $g_i^r(x) = f_i^r(x + d_i)$ when $i \in \mathcal{T}$. For $i \notin \mathcal{T}$, $g_i^r(x)$ is not defined.

Consider a route $r = (1, 2, \dots, n)$ and denote $P_{ir}^0 := \Pr\{0 \text{ extra charge periods before } i \in r\}$ and $P_{ir}^\phi := \Pr\{\phi \text{ extra charge periods before } i \in r\}$. We assume that the first timetabled trip of a vehicle route r is never delayed (i.e., $f_1^r(d_1) = 1$). Consider a trip $j \in (2, 3, \dots, n)$ preceded by a trip i . The distribution of the actual start time of trip j can be recursively computed as

$$f_j^r(y) = \begin{cases} \sum_{t=t_i^{min}}^{t_i^{max}} h_i(t) \sum_{y'=d_i}^{d_j - t - \Upsilon(i,j)} \left[f_i^r(y') P_{jr}^0 + \sum_{\phi=1}^k f_i^r(y' - \rho\phi) P_{jr}^\phi \right], & \text{if } y = d_j; \\ \sum_{t=t_i^{min}}^{t_i^{max}} h_i(t) \left[f_i^r(y - t - \Upsilon(i,j)) P_{jr}^0 + \sum_{\phi=1}^k f_i^r(y - t - \Upsilon(i,j) - \rho\phi) P_{jr}^\phi \right], & \text{if } y > d_j; \\ 0, & \text{otherwise,} \end{cases} \quad (7)$$

where t_i^{min} and t_i^{max} are the minimum and the maximum possible travel time values of timetabled trip $i \in \mathcal{T}$, respectively, and $\Upsilon(i, j)$ is equal to $t_{ij} + \tau$ if there is no charging activity between trips i and j , or $t_{iq} + t_{qj} + \tau + (j-i)\rho$ if there is a charging activity of $j-i$ time intervals at station $q \in \mathcal{Q}$ between trips i and j . Here, τ is the minimum layover time before each timetabled trip.

The expected secondary delay of a trip j assigned to r is expressed as $\mathbb{E}(X_j^r) = \sum_{x=0}^{x_{jr}^{max}} x \times f_j^r(x + d_j)$, where $x_{jr}^{max} = d_i + x_{ir}^{max} + t_{ij} + \tau + t_i^{max} - d_j$ is the maximum possible secondary delay of trip j when covered by vehicle route r . It should be observed that $f_j^r(y)$, $m_j^r(z)$, P_{ir}^0 , and P_{ir}^ϕ are, by definition, route-dependent. Since the routes are not enumerated but rather generated in our algorithm, it is impossible to compute $f_j^r(y)$, $m_j^r(z)$, P_{ir}^0 , and P_{ir}^ϕ for all $i \in \mathcal{T}$ and $r \in \mathcal{R}$ beforehand. Instead, the latter are dynamically generated throughout the solution process.

Every time a trip i is delayed of 1 minute, a penalty of $\beta\alpha_i$ is paid. Depending on the transport agency's level of delay aversion, the weighting factor β can be adjusted to find an appropriate trade-off between the operational costs and reliability. Generally speaking, the larger the β , the more reliable (or delay-tolerant) the S-MDEVSP solutions.

3 HEURISTIC BRANCH-AND-PRICE ALGORITHM FOR THE S-MDEVSP

Since there is generally a very large number of feasible vehicle routes in the S-MDEVSP, we propose a branch-and-price solution approach that generates columns (i.e., vehicle routes) instead

of enumerating them. We use the same heuristic branching strategy as in Ricard et al. (2022) to obtain integer solutions in a reasonable amount of time.

To identify columns that could be potentially useful to add, we solve one pricing problem per depot $d \in \mathcal{D}$ at each iteration. These pricing problems are defined on the networks G^d , for $d \in \mathcal{D}$, with modified arc costs \tilde{c}_{ij}^r defined as

$$\tilde{c}_{ij}^r = \begin{cases} \bar{c}_{ij}^r - \pi_d, & \text{if } i = n_o^d \\ \bar{c}_{ij}^r - u_i, & \text{if } i \in \mathcal{T}, \end{cases} \quad (8)$$

where $(u_i)_{i \in \mathcal{T}}$ and $(\pi_d)_{d \in \mathcal{D}}$ are dual variables associated with constraints (2) and (3), respectively. Since the cost of the arcs is stochastic and path-dependent in the S-MDEVSP, these pricing problems correspond to shortest path problems with stochasticity (Boland et al., 2015; Wellman et al., 2013) that can be solved by a modified version of the labeling algorithm (see Ahuja et al. (1993) for more details on this algorithm). Next, we specify the main characteristics of the dynamic programming algorithm, namely the labels, the extension functions, and the stochastic dominance criteria.

Labels.

Each label stores a representation of the actual start time cumulative distribution function (CDF), a representation of the SoC CDF, and the accumulated reduced cost. Let $F_j^p(y)$ be the CDF of $f_j^p(y)$ at node j defined as

$$F_j^p(y) = \sum_{y'=d_j}^y f_j^p(y'), \quad (9)$$

and let $M_j^p(z)$ be the CDF of $m_j^p(z)$ at node j defined as

$$M_j^p(z) = \sum_{z'=\sigma^{min}}^z m_j^p(z'). \quad (10)$$

The label L_j^p of path p at node j is defined as $L_j^p = (F_j^p(d_j), \dots, F_j^p(y_{jp}^{max}), M_j^p(\sigma^{min}), \dots, M_j^p(100), C_j^p)$, where y_{jp}^{max} is the maximum value of $F_j^p(y)$ and C_j^p is the accumulated reduced cost.

Extension functions.

We want to extend a label $L_i^{p'} = (F_i^{p'}(d_i), \dots, F_i^{p'}(y_{ip'}^{max}), M_i^{p'}(\sigma^{min}), \dots, M_i^{p'}(100), C_i^{p'})$ associated with node i along arc (i, j) to create label L_j^p . The accumulated reduced cost C_j^p at node j is given by

$$C_j^p = C_i^{p'} + \tilde{c}_{ij}^p. \quad (11)$$

In Section 2, we devised a method to analytically compute the propagation of delays in a sequence of timetabled trips. Here, we specify this method in the form of an extension function. The PDF of the actual start time of trip j covered by path p is given by

$$f_j^p(y) = \begin{cases} \sum_{t=t_i^{min}}^{t_i^{max}} h_i(t) \sum_{y'=d_i}^{d_j-t-t_{ij}-\tau} f_i^{p'}(y'), & \text{if } i, j \in \mathcal{T} \text{ and } y = d_j \\ \sum_{t=t_i^{min}}^{t_i^{max}} h_i(t) f_i^{p'}(y-t-t_{ij}-\tau), & \text{if } i, j \in \mathcal{T} \text{ and } y > d_j \text{ or } i \in \mathcal{T} \text{ and } j \in \tilde{\mathcal{Q}} \\ f_i^{p'}(y-\rho), & \text{if } i, j \in \tilde{\mathcal{Q}} \\ \sum_{y'=d_i}^{d_j-t_{ij}-\tau} \left[f_i^{p'}(y') P_{jp}^0 + \sum_{\phi=1}^k f_i^{p'}(y'-\rho\phi) P_{jp}^\phi \right], & \text{if } i \in \tilde{\mathcal{Q}}, j \in \mathcal{T}, \text{ and } y = d_j \\ f_i^{p'}(y-t_{ij}-\tau) P_{jp}^0 + \sum_{\phi=1}^k f_i^{p'}(y-t_{ij}-\tau-\rho\phi) P_{jp}^\phi, & \text{if } i \in \tilde{\mathcal{Q}}, j \in \mathcal{T}, \text{ and } y > d_j \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

The components $M_j^p(\cdot)$ are computed as

$$m_j^p(z) = \begin{cases} \sum_{\mu=\sigma^{min}}^{100} e_i(\mu) m_i^{p'}(z + \mu + e_{ij}), & \text{if } i \in \mathcal{T} \\ \sum_{z'=\sigma^{min}}^{100} m_i^{p'}(z') [\lambda(z', \rho \times \Lambda_\omega(z')) = z], & \text{if } i, j \in \tilde{\mathcal{Q}} \\ m_i^{p'}(z + e_{ij}), & \text{if } i \in \tilde{\mathcal{Q}}, j \in \mathcal{T}, P_{jp}^0 = 1 \\ m_i^{p'}(z + e_{ij}) + \sum_{z'=0}^{\omega-1} m_i^{p'}(z') [\lambda(z', \rho \times \Lambda_\omega(z')) = z], & \text{if } i \in \tilde{\mathcal{Q}}, j \in \mathcal{T}, \\ & z \geq \omega, 1 - P_{jp}^0 > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (13)$$

where $\lambda(z, t)$ is a piecewise linear function giving the final SoC of a battery after a charge of t minutes that started with an initial SoC of z . We assume all BEBs start the day fully charged.

Stochastic dominance criteria.

Consider two paths p^1 and p^2 with resident node i . Path p^1 dominates path p^2 when the following conditions hold:

1. $C_i^{p^1} \leq C_i^{p^2}$
2. $F_i^{p^1}(y) \geq F_i^{p^2}(y)$, for all $y \in \{d_i, d_i + 1, \dots, d_i + \max\{y_{ip^1}^{max}, y_{ip^2}^{max}\}\}$
3. $M_i^{p^1}(z) \leq M_i^{p^2}(z)$, for all $z \in \{\sigma^{min}, \sigma^{min} + 1, \dots, 100\}$

All dominated paths can be safely discarded because they are not part of the Pareto-optimal set of paths or will not be extended into Pareto-optimal paths.

4 COMPUTATIONAL RESULTS

We tested our model on a real-life instance from the city of Montréal of 273 timetabled trips, 2 depots, and 2 charging stations. To minimize battery degradation, σ^{min} is set to $\sigma^{min} = 25\%$. We compared our approach for two different types of commercially available BEBs; an electric shuttle with battery capacity (C) of $C = 80$ kWh, charger power (W) of $W = 60$ kW, and an average consumption rate of 0.76 kWh/km (Gao et al., 2017) and a 35-foot transit bus with $C = 492$ kWh, $W = 221$ kW, and an average consumption rate of 1.57 kWh/km (Proterra, 2022). We assume the energy consumption distributions follow a normal distribution.

The heuristic performance of our solution approach and the quality of the solutions are reported in Table 1 and 2, for the first and second type of BEB, respectively, for β values ranging from 0 to η , where η is the cost per bus used, and ω values ranging from σ^{min} to 75% of the battery capacity. When $\omega = \sigma^{min}$, our approach is equivalent to a robust optimization approach (i.e., no corrective actions). The columns display the relative difference in percentage between the upper bound and the lower bound (Gap), the number of branching nodes explored (Nodes), the computing times (CPU time), including the total time in seconds (Total), the portion of the total time dedicated to solve the root node (Root) and the pricing problems (Pricing), the operational costs (Op. costs), the fleet size (bus) and the total penalty for delays ($\sum_{r \in \mathcal{R}^*} \mathbb{E}[W_r(t, z)]$).

For both vehicle types, all problems are solved in less than 2 hours with almost all the computing time spent on solving the pricing problems. Also, the solutions obtained with our approach are at most 0.16 % more expensive than their corresponding lower bound, suggesting that our heuristic can find near-optimal solutions. Generally speaking, when β increases, the operational costs increase and the reliability improves.

For the first type of BEB, namely the shuttle with $C = 80$ kWh and $W = 60$ kW, the introduction of the recourse policy provides significant cost savings. Indeed, the fleet size can be reduced from 30 BEBs to 29 BEBs by introducing a recourse policy with $\omega \geq 50$, which could be considered as a substantial reduction since the number of vehicles used constitutes the major part of the operational costs. Furthermore, the deterioration in reliability that the charging policy introduces can be compensated for by a higher weighting factor β .

For the second type of BEB, namely the 35-foot transit bus with $C = 492$ kWh and $W = 221$ kW, introducing a recourse policy does not improve the cost of the solutions found nor does it reduce the size of the fleet. Thus, for this second type of vehicle with larger battery capacity and higher charging power, our approach is not useful and a simple robust optimization approach should be used to find vehicle routes such that the vehicles never run out of energy. Indeed, for this type of

Table 1: S-MDEVSP heuristic performance and quality of the solutions, with $C = 80$ kWh and $W = 60$ kW

		Heuristic performance					Quality of the solutions			
β	ω	Gap (%)	Nodes	CPU time			Op. costs	Bus	$\sum_{r \in \mathcal{R}^*} \mathbb{E}[W_r(t, z)]$	
				Total (s)	Root (%)	Pricing(%)				
0	σ^{min}	0.05	28	1,361	39.0	99.4	32,301.8	30	0.57	
	35	0.05	28	1,362	39.4	99.5	32,301.8	30	0.57	
	50	0.05	27	1,567	42.1	99.5	31,374.0	29	0.70	
	75	0.09	27	2,054	38.5	99.6	31,438.4	29	0.67	
	avg.	0.06	28	1,586	39.8	99.5	31,854.0	30	0.63	
$\eta/2$	σ^{min}	0.13	28	3,264	34.7	99.8	32,364.0	30	0.24	
	35	0.13	28	3,152	34.2	99.8	32,364.0	30	0.24	
	50	0.14	28	5,170	24.9	99.9	31,451.6	29	0.26	
	75	0.16	29	6,341	28.3	99.9	31,430.8	29	0.33	
	avg.	0.14	28	4,482	30.5	99.8	31,902.6	30	0.27	
η	σ^{min}	0.09	27	3,236	38.2	99.8	32,431.2	30	0.12	
	35	0.09	27	3,226	37.8	99.8	32,431.2	30	0.12	
	50	0.14	30	5,681	25.8	99.9	31,526.0	29	0.15	
	75	0.08	25	6,032	36.3	99.9	31,471.0	29	0.22	
	avg.	0.10	27	4,544	34.5	99.9	31,964.8	30	0.15	

Table 2: S-MDEVSP heuristic performance and quality of the solutions, with $C = 492$ kWh and $W = 221$ kW

		Heuristic performance					Quality of the solutions			
β	ω	Gap (%)	Nodes	CPU time			Op. costs	Bus	$\sum_{r \in \mathcal{R}^*} \mathbb{E}[W_r(t, z)]$	
				Total (s)	Root (%)	Pricing(%)				
0	σ^{min}	0.02	21	1,059	49.7	99.6	27,290.4	26	0.49	
	35	0.02	21	1,434	53.8	99.6	27,290.4	26	0.49	
	50	0.02	21	1,145	49.4	99.6	27,290.4	26	0.49	
	75	0.01	19	1,106	53.3	99.6	27,286.4	26	0.41	
	avg.	0.02	21	1,186	51.5	99.6	27,289.4	26	0.47	
$\eta/2$	σ^{min}	0.01	19	2,025	49.2	100.0	27,325.2	26	0.16	
	35	0.01	19	2,855	53.4	100.0	27,325.2	26	0.16	
	50	0.01	19	2,051	49.1	100.0	27,325.2	26	0.16	
	75	0.01	21	2,567	49.1	99.9	27,325.2	26	0.16	
	avg.	0.01	20	2,374	50.2	100.0	27,325.2	26	0.16	
η	σ^{min}	0.01	21	2,091	46.7	100.0	27,362.6	26	0.11	
	35	0.01	21	3,002	50.1	100.0	27,362.6	26	0.11	
	50	0.01	21	2,294	47.0	99.9	27,362.6	26	0.11	
	75	0.01	22	2,051	47.5	99.9	27,364.0	26	0.11	
	avg.	0.01	21	2,360	47.8	99.9	27,363.0	26	0.11	

BEB, the battery capacity is large enough that vehicles only need to charge once or twice a day. Because timetables of trips typically include off-peak periods with fewer trips to make, charging activities can easily be scheduled during these periods and batteries are often charged to their maximal capacity. In this context, the recourse policy we introduced is never activated.

5 CONCLUSIONS

In this work, we introduced a stochastic model for the MDEVSP that we formulated as a two-stage stochastic program with a recourse action. We proposed an efficient branch-and-price algorithm to solve this challenging problem. Our results indicated that the use of recourse actions is beneficial for shuttle BEBs with relatively small battery capacity and charging power, but not for 35-foot transit BEBs with larger battery capacity and charging power. Medium- to large-scale transit agencies are typically equipped with up-to-date BEBs that resemble the second type of vehicle tested, so our approach is probably not relevant for them. However, our two-stage stochastic model may be relevant for smaller transit agencies or those with access to fewer resources. Future work includes translating our approach to other routing problems with smaller electric vehicles, for example the electric dial-a-ride problem.

ACKNOWLEDGEMENTS

This project was partially supported by GIRO Inc. and the Natural Sciences and Engineering Research Council of Canada under the grants RDCPJ 520349-17 and BESC D3-558645-2021.

REFERENCES

- Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (1993). *Network flows: Theory, algorithms, and applications*. Prentice Hall.
- Bie, Y., Ji, J., Wang, X., & Qu, X. (2021). Optimization of electric bus scheduling considering stochastic volatilities in trip travel time and energy consumption. *Computer-Aided Civil and Infrastructure Engineering*, 36(12), 1530-1548. doi: <https://doi.org/10.1111/mice.12684>
- Boland, N., Dickson, S., Savelsbergh, M., & Smilowitz, K. (2015). *Dominance in pricing problems with stochasticity*. (accessed February 22, 2022)
- Gao, Z., Lin, Z., LaClair, T. J., Liu, C., Li, J.-M., Birky, A. K., & Ward, J. (2017). Battery capacity and recharging needs for electric buses in city transit service. *Energy*, 122, 588-600. doi: <https://doi.org/10.1016/j.energy.2017.01.101>
- Iverson, K. E. (1962). *A programming language*. USA: John Wiley amp; Sons, Inc.
- Jiang, M., Zhang, Y., & Zhang, Y. (2021). Optimal electric bus scheduling under travel time uncertainty: A robust model and solution method. *Journal of Advanced Transportation*, 2021, 19. doi: <https://doi.org/10.1155/2021/1191443>
- Li, L., Lo, H. K., Huang, W., & Xiao, F. (2021). Mixed bus fleet location-routing-scheduling under range uncertainty. *Transportation Research Part B: Methodological*, 146, 155-179. doi: <https://doi.org/10.1016/j.trb.2021.02.005>
- Montoya, A., Guéret, C., Mendoza, J. E., & Villegas, J. G. (2017). The electric vehicle routing problem with nonlinear charging function. *Transportation Research Part B: Methodological*, 103, 87-110. doi: <https://doi.org/10.1016/j.trb.2017.02.004>
- Proterra. (2022). *ZX5 35-foot battery-electric transit bus - platform specifications*. Retrieved 2023-03-01, from https://www.proterra.com/wp-content/uploads/2022/09/SPEC_35_001_METRIC_Q4_2022_V1_09_09_22-1.pdf
- Ricard, L., Desaulniers, G., Lodi, A., & Rousseau, L.-M. (2022). Predicting the probability distribution of bus travel time to measure the reliability of public transport services. *Transportation Research Part C: Emerging Technologies*, 138, 103619. doi: <https://doi.org/10.1016/j.trc.2022.103619>

Tang, X., Lin, X., & He, F. (2019). Robust scheduling strategies of electric buses under stochastic traffic conditions. *Transportation Research Part C: Emerging Technologies*, 105, 163-182. doi: <https://doi.org/10.1016/j.trc.2019.05.032>

Wellman, M. P., Ford, M., & Larson, K. (2013). *Path planning under time-dependent uncertainty*.

A heuristic approach to improve the robustness of a railway timetable in a bottleneck area

Inneke Van Hoeck*¹ and Pieter Vansteenwegen¹

¹KU Leuven Institute for Mobility - CIB, KU Leuven,
Celestijnenlaan 300, Box 2422, 3001 Leuven, Belgium

SHORT SUMMARY

In this paper, a heuristic is presented to improve the robustness of a given railway timetable in a bottleneck area. The timetable can be adapted by adjusting the timing and routing of trains, but cannot deviate too much from the current timetable. Robustness is measured with an objective function that considers the buffer times between train pairs. The developed algorithm updates the routing and the timing in separate steps. It is applied to a bottleneck area of the Belgian network. The results show that the objective can be improved by about 10% when alternative routing options are considered. Additional experiments with smaller instances indicate that the heuristic is capable of finding near-optimal solutions.

Keywords: Operations research, Public transport, Railway transport, Timetable robustness

1 INTRODUCTION

Railway networks are often heavily used. In nearly saturated parts of the network, limited capacity is one of the main reasons for delay propagation (Burggraeve & Vansteenwegen, 2017). Often, delay propagation is closely related to timetable robustness (Dewilde et al., 2014). Robustness is a concept that has many definitions and interpretations, an extensive survey on this topic can be found in Lusby et al. (2018). The aim of this work is to adjust a given timetable to improve its robustness. Our objective function is based on the work of Dewilde et al. (2014) and looks for a good spreading in time of the trains by considering the buffer times between trains. Both the timing and the routing of the trains can be adjusted. However, the new timetable cannot deviate too much from the current timetable. This is required by the Belgian railway companies to ensure the practical applicability. Additional constraints for the timing of the trains are included to guarantee this. Several methods to improve the robustness of a timetable by making relatively small changes have been presented in literature. Some examples can be found in Andersson et al. (2015), Jovanović et al. (2017) and Högdahl et al. (2019). They all present methods to optimize the allocation of time supplements in a given timetable. These methods do not allow alternative routing options, as will be considered in this work.

Thus, in this paper it is examined how the robustness of a given timetable can be improved by changing both the timing and routing of trains, while not deviating too much from the original timetable. The timetable is considered on a microscopic level. A heuristic to solve this problem is presented and applied to a case study for a part of the Belgian railway network.

2 PROBLEM DESCRIPTION

Denote the set of all considered trains as T . A path p is associated with each train to indicate the route that it takes through the network. We refer to a *trainroute* (t, p) when talking about a train t that uses path p . A path is a sequence of infrastructure resources that are used by the train. The set of resources that is used by a trainroute (t, p) is denoted by $R_{(t,p)}$. Each resource r in a trainroute (t, p) will be blocked by the train during a certain period of time, meaning that no other train is allowed to use the resource. The blocking times are determined according to the blocking time theory. Thus, for each $r \in R_{(t,p)}$ the *reserve* and *release* time indicate the time when the blocking time of the resource starts and stops, respectively denoted by $res_{(t,p)}^r$ and $rel_{(t,p)}^r$. These

times are determined relative to the start time of the train t , denoted by S_t . This is the time when a train enters the considered zone. In this work, alternative routing options for trains are considered. Thus, for each train t there are multiple possible paths $p \in P_t$ that can be selected. The point where a train enters and leaves the considered zone is assumed to be fixed for all paths. The reserve and release times for the resources in an alternative path are estimated based on the timing for the path that is used in the current timetable.

The absolute reserve and release times for a resource r in a trainroute (t, p) can be written as:

$$RES_{(t,p)}^r = S_t + res_{(t,p)}^r, \quad (1)$$

$$REL_{(t,p)}^r = S_t + rel_{(t,p)}^r. \quad (2)$$

The *buffer time*, between a pair of trainroutes (t, p) and (t', p') is defined as

$$B_{(t,p),(t',p')} = \min_{r \in R_{(t,p)} \cap R_{(t',p')}} b_{(t,p),(t',p')}^r \quad (3)$$

where $b_{(t,p),(t',p')}^r$ is the minimum time span between the two trainroutes on resource r . It is defined as:

$$b_{(t,p),(t',p')}^r = \begin{cases} \min(RES' - REL, RES + H - REL') & \text{if } t \prec t', \\ \min(RES - REL', RES' + H - REL) & \text{if } t' \prec t. \end{cases} \quad (4)$$

Here, the notation of the absolute reserve and release times is simplified by omitting the explicit reference to the trainroute and resource. Additionally, $t \prec t'$ means that trainroute (t, p) uses resource r before trainroute (t', p') does. The parameter H is the time period that is considered, 1 hour in this case. Note that this definition takes into account that the timetable is cyclic.

Similar to the work of Dewilde et al. (2014), a cost $c_{(t,p),(t',p')}$ is associated with each buffer time as follows:

$$c_{(t,p),(t',p')} = \begin{cases} 100 & \text{if } B \leq 0, \\ \frac{-1}{10}B + 10 & \text{if } 0 < B \leq 60, \\ \frac{-1}{30}B + 6 & \text{if } 60 < B \leq 120, \\ \frac{-1}{390}B + \frac{90}{39} & \text{if } 120 < B \leq 900, \\ 0 & \text{else,} \end{cases} \quad (5)$$

where $B_{(t,p),(t',p')}$ is simply denoted by B . This is a piecewise linear, monotone decreasing function. A small buffer time will thus correspond to a large cost. The objective is to minimize the sum of all costs:

$$\text{minimize } \sum c_{(t,p),(t',p')}. \quad (6)$$

With this objective in mind, let us look more closely at the definition of the cost. A cost of 100 is given when the buffer time is smaller than or equal to 0, i.e. when there is a conflict between the two trains. When the buffer time is larger than 900 seconds, the cost is set to 0 based on the idea that these trains are far enough apart such that they do not influence each other (Dewilde et al., 2014). The slope of the functions that are used for the different intervals for B becomes less steep when B becomes larger. This is done to include the following idea: improving a buffer time from 30 seconds to 1 minute is much more valuable than increasing a buffer time of 10 minutes with 30 seconds. Note that the same principle can be obtained by defining the cost as $1/B$. However, we opt for a piecewise linear function because this allows a MILP formulation of the problem.

In conclusion, the aim of this work is to obtain a new feasible timetable such that the value of the objective function (6) is minimal. A timetable is completely defined when the start time S_t and the selected path $p \in P_t$ is known for each train t . As stated in section 1, the new timetable cannot deviate too much from the current timetable to ensure the practical applicability. Additional timing constraints are imposed to achieve this, but they are not discussed further in this short paper.

3 HEURISTIC APPROACH

Solving the problem described in section 2 with an exact model is not possible within a reasonable amount of time for the instances considered in the case study (see section 4). Therefore, a heuristic

approach is developed. Updating the timing and the routing is done in separate steps. In the remainder of this work, a *route selection* refers to the set of trainroutes (t, p) that are used in the timetable. In a single iteration of the algorithm, a new solution is determined by first calling the `routing update` to find a new route selection, while the timing is fixed, followed by the `timing update` function to change the start times. These functions are now discussed in more detail.

Routing update

The `routing update` is a function that takes the current start times as input. The aim of this function is to find a feasible route selection such that the timetable defined by this route selection and the current start times has the lowest possible cost. Note that the complete impact of the route selection on the cost of the timetable is only known after the `timing update` is applied. However, observations show that timetables with a low cost after only applying the `routing update` lead to timetables with a low cost after the `timing update` is also applied.

The crucial concept in the `routing update` is the *cost graph*. Each trainroute (t, p) is a node in this graph. An edge is added between two nodes if the corresponding trainroutes can be simultaneously included in the route selection. Two conditions must hold for this. First, there cannot be a conflict between the trainroutes. Second, the additional timing constraints must hold, if there are such constraints defined between the trainroutes. The cost $c_{(t,p),(t',p')}$ is associated with the edge between two trainroutes. Denote the set of edges in the cost graph as E . The trainroute combinations that are not connected by an edge are included in the complement set E^C .

A feasible route selection corresponds to a clique of the cost graph where the number of nodes in the clique is equal to the number of trains, since a route must be selected for each train. As stated before, we aim to find the route selection that leads to the timetable with minimal cost. This is done by solving the following ILP model. The decision variables are:

$$\begin{aligned} x_{(t,p)} &= 1 \text{ if the node } (t, p) \text{ is included in the clique, } 0 \text{ otherwise,} \\ y_{(t,p),(t',p')} &= 1 \text{ if the edge } ((t, p), (t', p')) \text{ is included in the subgraph induced by the clique,} \\ &0 \text{ otherwise.} \end{aligned}$$

The objective function is given by:

$$\text{minimize} \quad \sum_{((t,p),(t',p')) \in E} c_{(t,p),(t',p')} \cdot y_{(t,p),(t',p')}. \quad (7)$$

For each train, exactly one route must be included in the clique, thus:

$$\sum_{p \in P_t} x_{(t,p)} = 1 \quad \forall t \in T. \quad (8)$$

If two nodes are not connected in the cost graph, then at most one of them can be included in the clique:

$$x_{(t,p)} + x_{(t',p')} \leq 1 \quad \forall ((t,p), (t',p')) \in E^C. \quad (9)$$

An edge is included in the subgraph induced by the clique if both its incident nodes are included in the clique. The following constraints hold $\forall ((t,p), (t',p')) \in E$:

$$y_{(t,p),(t',p')} \leq x_{(t,p)}, \quad (10)$$

$$y_{(t,p),(t',p')} \leq x_{(t',p')}, \quad (11)$$

$$y_{(t,p),(t',p')} \geq x_{(t,p)} + x_{(t',p')} - 1. \quad (12)$$

The trainroutes that are included in the clique give the new route selection.

Timing update

In the `timing update` function, the current route selection and start times are given as input. The aim is to adjust the timing of the trains to decrease the cost of the timetable. The function keeps iterating until no decrease larger than a certain threshold can be found or when a maximum computation time, for example 1 second, is reached. In a single iteration, first the `possible improvements` function is applied to create a list of possible timings that can be changed. The `select improvements` function is then used to determine which of these possible changes will

actually be applied. These functions are now discussed in more detail.

Because the route selection is fixed in the entire `timing update` function, we simplify the notation by omitting the reference to the path p and refer to train t instead of trainroute (t, p) . For each train t the total cost related to this train in the current timetable is defined as follows:

$$C_t = \sum_{t' \in T \setminus t} c_{t,t'}. \quad (13)$$

The current start times are used to calculate this cost. Now, the `possible improvements` function runs over all trains. For each train t , the aim is to find the optimal start time \hat{s}_t , while the start times of the other trains remain fixed. The definition of the cost related to a train can be extended to be dependent on the start time of that train. Thus C_t^s is calculated as in expression 13, but the start time of train t is now equal to s , instead of its current start time. We look for the value s that gives the minimal related cost C_t^s . Discrete times, with a time step of 6 seconds (0.1 minutes), are considered for s . For a possible start time s it is first verified if this start time leads to a feasible timetable, i.e. there are no conflicts, and if the additional timing constraints are satisfied. If this is the case, the cost C_t^s is determined. The time that leads to the minimum cost is denoted by \hat{s}_t . The improvement that can be made to the objective by changing the start time of train t is defined as $C_t - C_t^{\hat{s}_t}$. If this value is greater than a certain threshold, for example 0.1, then t is added to the list of possible improvements.

Not all of these possible changes can be applied at the same time. When two trains have an impact on each other, meaning that they have common resources or have a timing constraint defined between them, the calculated improvements are no longer correct if their timing is updated at the same time. Therefore, the `select improvements` function determines a subset of the possible improvements such that none of the selected trains have an impact on each other. The total improvement in the objective is then given by the sum of the improvements of the separate trains. Selecting such a subset is done by creating an *improvement graph*. Each train that is included in the list of possible improvements is a node in this graph. An edge between two nodes indicates that they have an impact on each other. Thus, finding trains that can be updated at the same time corresponds to finding an independent set in the improvement graph. This is done in a greedy way as follows. Select the node in the improvement graph with the largest improvement. Remove this node and its adjacent nodes from the graph. Repeat this process until there are no nodes left in the graph. The timing of the selected trains can then be updated.

Overview of the algorithm

Algorithm 1 now presents an overview of the heuristic algorithm. The `update routing` and `update timing` functions that were discussed in the previous sections are the core of the algorithm.

Algorithm 1: Overview of the heuristic algorithm

```

Data: original_solution
Result: optimal_solution
1 solution ← find initial solution(original_solution);
2 while comp_time ≤ max time do
3   new_route_selection ← routing update(start_times);
4   new_start_times ← timing update(start_times, new_route_selection);
5   if new_objective < objective then
6     solution ← new_solution
7     if new_objective < optimal_objective then
8       optimal_solution ← new_solution
9     end
10  else
11    solution ← generate random solution(new_solution)
12  end
13 end

```

The algorithm starts from the original timetable. The function `find initial solution` adapts the timing of the trains to obtain a better spreading. The routes are not adjusted yet. In this function, a MILP model is used to minimize the overall maximum cost that is found in the timetable,

while satisfying the feasibility constraints. Then the `timing update` is applied to create a better spreading in time. As stated before, an iteration starts by calling the `routing update` to generate a new route selection. This route selection and the current start times are then used as input for the `timing update`. Together, the new start times and the new route selection completely define a new timetable. The objective value related to this timetable is referred to as new objective. If this new objective is better than the current objective, the current solution is updated. It is also checked if the new objective is better than the optimal solution found until now, if so, the optimal solution is updated. If the new objective is not better than the current objective, then a new random, feasible solution is generated to continue with in the next iteration. The algorithm continues until a maximum computation time is reached.

4 RESULTS AND DISCUSSION

A part of the Belgian network that is centered around the station of Halle is considered. This is located just outside of Brussels, which is the main bottleneck of the Belgian railway network. Figure 1 shows a macroscopic view of this area. A small part of it is shown in microscopic detail on figure 2 to illustrate the complexity of the network. In this case study, zone 1 and 2 indicated in figure 1 are considered. All the trains that use this part of the network during one hour of the morning peak (between 7:00 and 8:00) are selected. This leads to 35 trains for zone 1 and 40 trains for zone 2. More data and information related to this case study can be found in Uyttendaele et al. (2022). A regular laptop with an 11th Gen Intel Core i7-1185G7 @ 3.00 GHz processor, 16 GB RAM was used to run the experiments.

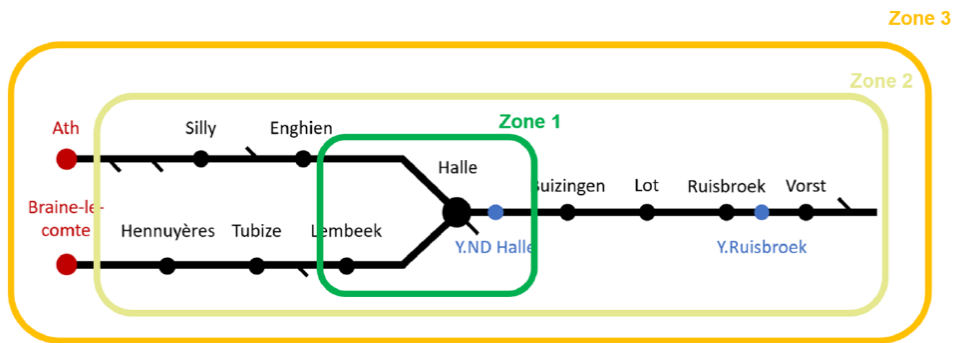


Figure 1: Macroscopic view of the considered part of the network.

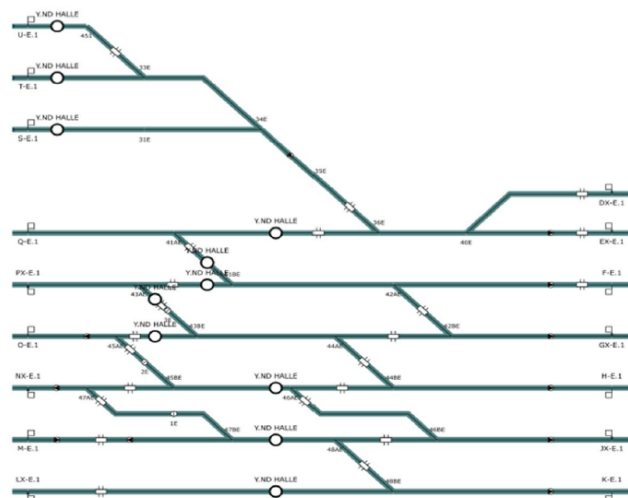


Figure 2: Microscopic view of the switch area Y.ND Halle, which corresponds to the blue dot in zone 1 indicated in figure 1.

The problem is considered for zone 1 and 2 with 5 and 10 possible paths for each train. Only the routes with the shortest duration are considered. It is unlikely that longer routes will lead to good solutions because this typically corresponds to routes with some detours that are not desirable in practice. As stated before, solving this problem with a MILP model is not possible in a reasonable amount of time. After 48 hours of running for zone 1 with 5 alternative routes, the best objective found was 68.08, but there was still an optimality gap of 25%, so no conclusions can be drawn from this. Therefore, the heuristic algorithm was developed. Because the heuristic contains randomness, it is run 10 times for each instance. Each run has a maximum allowed computation time of 30 seconds. The best obtained objective and the average objective are reported in table 1. The improvement compared to the initial objective is given between brackets. This initial objective is the objective after the `find initial solution` is applied.

Table 1: Results of the heuristic algorithm for zone 1 and 2 with 5 and 10 possible routes.

	Initial objective	Heuristic: best (improvement)	Heuristic: average (improvement)
Zone 1 - 5 routes	79.03	69.63 (-11.89%)	71.85 (-9.09%)
Zone 1 - 10 routes	79.03	69.93 (-11.51%)	71.01 (-10.15%)
Zone 2 - 5 routes	205.97	180.11 (-12.56%)	183.03 (-11.14%)
Zone 2 - 10 routes	205.97	184.15 (-10.59%)	184.43 (-10.46%)

The heuristic clearly finds solutions with a lower objective value than the initial objective. For the different instances, the best obtained objective is quite close to the average objective. This indicates that the heuristic performs consistently over the different runs. For both zones, the best solution that is found for 5 possible routes is better than for the case with 10 possible routes, although the difference for zone 1 is quite small. This is an unexpected result because the solution that is found for 5 possible routes is of course also a possible solution when 10 routes are considered. The fact that these solutions are not found by the heuristic is possibly a consequence of the larger search space in the case with 10 routes.

Because the optimal solution is not known for any of the instances, some additional experiments are conducted to gain some insight in the performance of the heuristic. For both zone 1 and 2 with 5 possible routes, 5 instances are generated by randomly selecting 10 trains to consider instead of all the trains that use the zone. For these instances, the optimal solution can be calculated by a MILP model in only a few seconds. It can then be compared with the solution found by the heuristic. Like before, the heuristic is run 10 times, with a computation time of 30 seconds for each run. The results for zone 1 and 2 are presented in table 2 and 3, respectively. The results show that the heuristic is able to find good solutions, i.e. with a small gap. For instances A and D of zone 2, there is a larger gap. These experiments show that the heuristic is indeed capable of finding near-optimal solution. This does however not guarantee that this is also the case for the larger instances of the case study.

Table 2: Results of the smaller instances for zone 1.

	Optimal objective	Heuristic: best (gap)	Heuristic: average (gap)
A	13.40	13.41 (+0.07%)	13.56 (+1.19%)
B	11.19	11.19 (+0.00%)	11.20 (+0.09%)
C	11.01	11.13 (+1.09%)	11.56 (+5.00%)
D	13.62	13.69 (+0.51%)	13.82 (+1.47%)
E	9.16	9.17 (+0.11%)	9.17 (+0.11%)

Table 3: Results of the smaller instances for zone 2.

	Optimal objective	Heuristic: best (gap)	Heuristic: average (gap)
A	5.14	6.21 (+20.82%)	6.21 (+20.82%)
B	9.64	9.66 (+0.21%)	9.66 (+0.21%)
C	15.33	15.39 (+0.39%)	15.88 (+3.59%)
D	13.05	14.09 (+7.97%)	14.75 (+13.03%)
E	11.18	11.50 (+2,86%)	11.58 (+3.58%)

5 CONCLUSIONS

In this paper, a heuristic is presented to improve the robustness of a given timetable by adjusting the timing and routing of the trains. Robustness is measured by considering the buffer times between train pairs. The developed algorithm consists of separate steps to first update the routing and then the timing of the trains. The heuristic is applied to a bottleneck area of the Belgian network. The results for the different instances show that the value of our objective can be improved by about 10% by considering alternative routing options and not only changing the timing. Additional experiments with smaller instances show that the heuristic is capable of finding near-optimal solutions.

ACKNOWLEDGEMENTS

The authors would like to thank Infrabel, the Belgian railway infrastructure manager, and NMBS, the Belgian railway passenger operator, for this collaboration and their support.

REFERENCES

- Andersson, E., Peterson, A., & Törnquist Krasemann, J. (2015). Improved railway timetable robustness for reduced traffic delays - a milp approach. *Proceedings of the 6th International Seminar on Railway Operations Modelling and Analysis (RailTokyo2015)*.
- Burggraeve, S., & Vansteenwegen, P. (2017). Robust routing and timetabling in complex railway stations. *Transportation Research Part B*, 101, 228–244.
- Dewilde, T., Sels, P., Cattrysse, D., & Vansteenwegen, P. (2014). Improving the robustness in railway station areas. *European Journal of Operational Research*, 235, 276–286.
- Högdahl, J., Bohlin, M., & Fröidh, O. (2019). A combined simulation-optimization approach for minimizing travel time and delays in railway timetables. *Transportation Research Part B*, 126, 192–212.
- Jovanović, P., Kecman, P., Bojović, N., & Mandić, D. (2017). Optimal allocation of buffer times to increase train schedule robustness. *European Journal of Operational Research*, 256(1), 44–54.
- Lusby, R., Larsen, J., & Bull, S. (2018). A survey on robustness in railway planning. *European Journal of Operational Research*, 266(1), 1–15.
- Uyttendaele, J., Van Hoeck, I., Besinović, N., & Vansteenwegen, P. (2022). Timetable compression using max-plus automata applied to large railway networks. *TOP*, 1–26.

Post-hoc explanation methods for deep neural networks in choice analysis

Niousha Bagheri Khoulenjani*¹, Milad Ghasri¹, Michael Barlow¹

¹ School of Engineering and IT (SEIT), UNSW Canberra, ACT, 2600, Australia

SHORT SUMMARY

Deep Neural Networks (DNNs) are accurate and powerful tools for modeling travel decisions. Nonetheless, the black-box characteristic of DNNs has decreased their potential implication in discrete choice modeling. In this study, we investigate the potentials of cutting-edge post-hoc interpretation tools in providing behavioral insight into DNN architectures. We evaluate the relationship between the output probabilities and input features using the Shapely Additive explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). Using SwissMetro dataset, we demonstrate that the outputs of SHAP and LIME are consistent with theory when the architecture of DNN is designed based on the Random Utility Maximization (RUM) theory. However, for a fully connected DNN architecture, SHAP and LIME do not provide behaviorally interpretable outputs. Additionally, the prediction accuracy shows the DNN model based on RUM avoids overfitting.

Keywords: Discrete choice modeling, Deep Neural Networks, Explainable AI

1. INTRODUCTION

DNN models have become ubiquitous in Intelligent Transport System (ITS) due to their powerful predictive and efficient learning algorithms and fixable modelling structure. ITS, as an integrated transport management system, refers to the use of data communication, information processing, traffic management technologies and Artificial Intelligence (AI) in transport (Chen, Liu et al. 2020). All applications in ITS that rely on DNNs are categorized into four groups of computer vision, time series prediction, classification, and optimization (Wang, Zhang et al. 2019). Most studies of DNNs in ITS belong to the time series prediction group to model variables such as travel time, traffic flow, and traffic speed prediction. On the other hand, the number of applications in classification using DNNs, particularly modelling travel mode choice, is fairly limited (van Cranenburgh and Alwosheel 2019). Traditionally, discrete choice modelers have mostly used econometric methods, including discrete choice models. These models are based on a theoretical foundation with predefined assumptions and underlying relationships between dependent and explanatory variables (Train 2009). Econometric methods are inferior to DNN methods in terms of prediction accuracy. This is one of the main reasons that DNNs have become pervasive in modelling individuals' behavior (Golshani, Shabanpour et al. 2018).

DNNs encompass a wide range of architectures such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) (Goodfellow, Bengio et al. 2016), but the applications of DNNs in discrete choice modeling are mainly limited to the most basic DNN's architecture, called Multi-Layer Perceptron (MLP). Even this basic MLP model is shown to achieve higher prediction accuracy in comparison with traditional discrete choice models (Assi, Nahiduzzaman et al. 2018). Accurate modelling of travel behavior is essential, and it is equally important that high accuracy is resulted from an interpretable model. Despite all the advantages of DNNs, they are considered as complex black-box (non-interpretable) models because of the numerous parameters in the model. In other words, the structure of DNNs is not directly interpretable, as hundreds of

parameters need to be described (Zhao, Yan et al. 2020). Nonetheless, a wide range of interpretable tools known as Post-hoc explainability techniques are proposed to extract knowledge from complex DNN models (Arrieta, Díaz-Rodríguez et al. 2020). The post-hoc approaches justify how and why a DNN model has arrived at its prediction (Lipton 2018).

Almost all applications of DNN in discrete choice modeling have applied a post-hoc approach to interpret the DNN models. For example, Sifringer, Lurkin et al. (2020) proposed a DNN architecture inspired by RUM, consisting of an interpretable and a fully connected part. The authors calculated the importance of each input variable to the fully connected part using a traditional post-hoc approach named the saliency map. However, in addition to finding the importance of input variables, the main purpose of post-hoc approaches is evaluating the output of complex machine learning models such as DNNs. In other words, the output of the DNN model needs to be evaluated using post-hoc approaches to increase the trust in DNN’s decisions. Although there is a wide variety of post-hoc explanation approaches in the literature, only a few traditional approaches of post-hoc analysis can be found in recent studies e.g. (Wang, Mo et al. 2020, Wang, Wang et al. 2020, Wang, Wang et al. 2021, Wong and Farooq 2021).

This study seeks to evaluate recent DNN models in discrete choice modeling using the state-of-the-art post-hoc approaches. The performance of two novel post-hoc approaches in the literature, Shapely Additive explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) are tested on a fully connected DNN model and a DNN model based on RUM theory. Our study contributes to evaluating the performance and reliability of the recent DNN models in discrete choice modeling. Furthermore, this comparison and evaluation process will help the researcher to select the appropriate DNN architecture and interpretation approach.

2. METHODOLOGY

Deep Neural Networks and Random Utility Maximisation

The architecture of DNNs models used in discrete choice analysis can be divided into two groups. The first group are those that use a fully connected architecture e.g. (Assi, Nahiduzzaman et al. 2018, Zhao, Yan et al. 2020), and the second group are those who use customized architectures to make the model consistent with behavioral theories such as RUM (Wang, Mo et al. 2020, Wong and Farooq 2021). The Fully connected DNN (F-DNN) connects input variables to the output probabilities through several layers with hundreds of parameters (Goodfellow, Bengio et al. 2016). In F-DNN, the utility of each alternative is connected to the attributes of all alternatives. The second group encompasses specific types of DNNs developed based on RUM which computes the utility of each alternative based on its corresponding attributes.

A recent DNN architecture with alternative-specific utility functions (ASU-DNN) proposed by Wang, Mo et al. (2020), could achieve high accuracy levels in modeling discrete choice data. ASU-DNN contains an input layer, two hidden layers and the output layer. Assume input variables to be a vector of \mathbf{x} , and the input variables are divided into a vector of alternative specific variables denoted by x_{ik} and a vector of individual specific variables denoted by x_i where $i \in \{1,2, \dots, n\}$ and $k \in \{1,2, \dots, K\}$. Then, consistent with RUM (Train 2009), the utility of each alternative is defined as a function of individual specific variables and its corresponding alternative specific variables as indicated in equation (1).

$$V_k = V(x_i, x_{ik}) = w_k(g_1 \circ g_2 \dots g_{M_2})((g_1^{x_{ik}} \circ g_1^{x_{ik}} \dots g_{M_1}^{x_{ik}})(x_{ik}), (g_1 \circ g_2 \dots g_{M_1})(x_i)) \quad (1)$$

In this equation M_1 and M_2 are the number of neurons in the first and second hidden layers respectively; $g(t) = \text{Max}(t, 0)$ is the RELU activation function, and w_k is the vector of parameters

to be estimated. In the last layer, the Softmax activation function, shown in equation (2) is applied to the utilities in order to calculate the output probabilities (Goodfellow, Bengio et al. 2016):

$$S(V_{ik}) = \frac{e^{V_{ik}}}{\sum_j e^{V_{ij}}} \quad (2)$$

Interpretation methods

The demands for interpretability in DNNs have increased in recent years (Arrieta, Díaz-Rodríguez et al. 2020). Therefore, many approaches as the post-hoc explainability methods for DNNs, have been developed. The existing post-hoc explainability approaches fall into six categories of text explanation, visual explanation, local explanation, explain by examples, explain by simplification, and feature relevance explanation. For further details about each category, the reader is referred to Arrieta, Díaz-Rodríguez et al. (2020). In this study, we apply two approaches, Shapely Additive explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) from the local explanation category.

SHAP is a game theory interpretation method of machine learning methods that evaluates the negative and positive impact of input variables (Lundberg and Lee 2017). For a given input X and a DNN model $f(X)$, SHAP utilizes an Explanation Model (EM) to evaluate the contribution of each input variable x_i to the model f . EM sets up a relationship between x_i and the model outputs. The parameters of this model are called SHAP value denoted by φ_i . SHAP values are defined as the weighted average of the marginal contributions over all possible coalitions $|F|!$ and are calculated as indicated in equation (3).

$$\varphi_i(f) = \sum_{\{S \subseteq F\} \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(x_{S \cup \{i\}}) - f(x_S)] \quad (3)$$

In this equation, F is the total number of features, and S is a subset of F . $f(x_{S \cup \{i\}})$ is the model output using feature i and features in S , and $f(x_S)$ is the model output using features in S but without feature i . As computing the exact value of φ_i is challenging, several methods have been introduced to approximate SHAP values (Lundberg and Lee 2017). In this study, we apply Kernel SHAP as it is a model-agnostic method that can be used for all types of machine learning models, and it is a reasonable approach when number of input variables are small in the dataset (Lundberg and Lee 2017).

Similar to SHAP, the LIME belongs to the local interpretation category that measures the impact of input variables on the variations of model output (Ribeiro, Singh et al. 2016). LIME generates new datasets around an observation x consisting of the corresponding outputs of the model. Then, an explainable model g is trained on the new dataset that is weighted by the proximity of the sample observations. With the new explainable model g and trained DNN model f , it is possible to provide a rough estimate of the contribution of input variable x to the model f . To accomplish this, the following objective function is minimized:

$$\xi(x) = \operatorname{argmin}_{g \in G} L(f, g, \pi_x) + \Omega(g) \quad (4)$$

Where G denotes the set of all interpretable models, and $\Omega(g)$ is the complexity of model g . π_x is the proximity measure between generated data to sample x . L measures how unfaithful g is in the approximation of f in the locality defined by π_x .

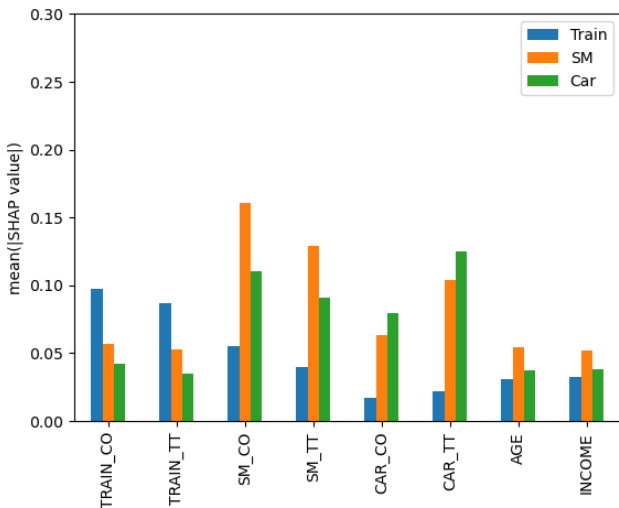
Although it is expected LIME and SHAP yield similar results, they have different structures in interpreting models. While LIME generates a perturbed dataset to fit an explainable model, SHAP requires an entire sample to approximate SHAP values.

3. RESULTS AND DISCUSSION

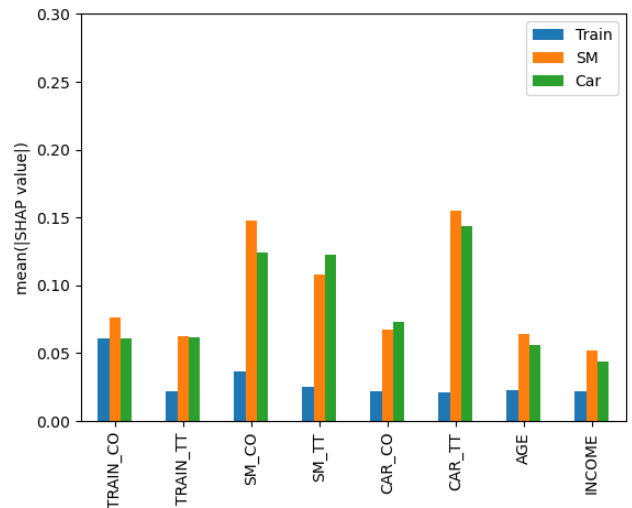
This study uses the Swissmetro dataset for evaluating the interpretation of DNNs. This dataset was compiled in Switzerland in 1998 (Bierlaire, Axhausen et al. 2001). It contains 1192 respondents who were asked to choose their preferred transportation mode among three alternatives of train, Swissmetro and car. This dataset contains 9,036 observations after cleaning. In the current study, Travel Time (TT), Travel Cost (CO), AGE and INCOME are selected among available variables for choice analysis.

The two models of ASU-DNN and F-DNN are developed using the Swissmetro dataset. Then the two methods are SHAP and LIME are used to provide insight into how these models make predictions. SHAP and LIME also show which features are most important in both models separately. In the end, the performance of ASU-DNN and F-DNN is compared through accuracy and log-likelihood. In this experiment, ASU-DNN includes two layers with $M_1 = M_2 = 100$. Similar to ASU-DNN, F-DNN includes 2 layers with 100 neurons in each layer.

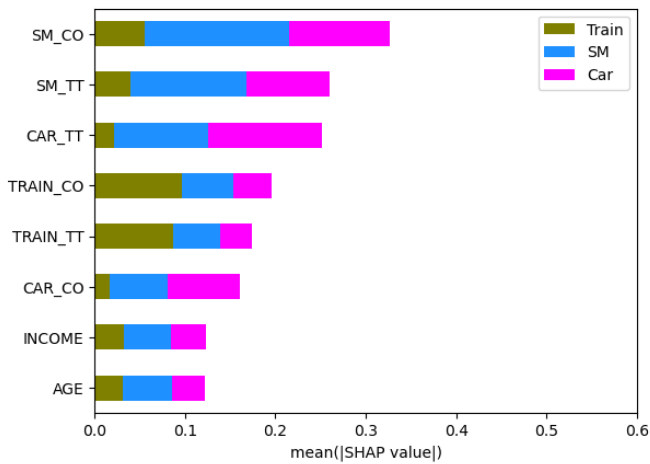
In this study, the impact of an input variable on each class will be calculated using SHAP and LIME. In contrast to an image input with a 2D set of pixels (which is the common application for post-hoc explanation methods), the position of the input variables always has the same meaning. For example, if the first input variable of all observations is age, then the first LIME and SHAP values will always be the impact of a passenger’s age on each class. The average of absolute SHAP values for all features in each class is reported in Figure 1. The first line illustrates the SHAP values in ASU-DNN and F-DNN corresponding to each class, and the second line ranks the summation of SHAP values for each feature. As shown in plot (a), features related to each utility have the most contribution to probability of that utility. For example, Swissmetro Cost (SM_CO) and Swissmetro Travel Time (SM_TT) have the highest impact on Swissmetro. However, from (b), there is no specific connections between input variables and utilities. For example, Train Travel Time (Train-TT) has the highest impact on Swissmetro and Car. Plots (c) and (d) demonstrate the overall importance of input variables on the output probabilities in ASU-DNN and F-DNN. Both models show that SM_CO, SM_TT and Car Travel Time (CAR_TT) have the highest impact on the mode choice decision. Also, INCOME and AGE have the least feature importance in this classification task.



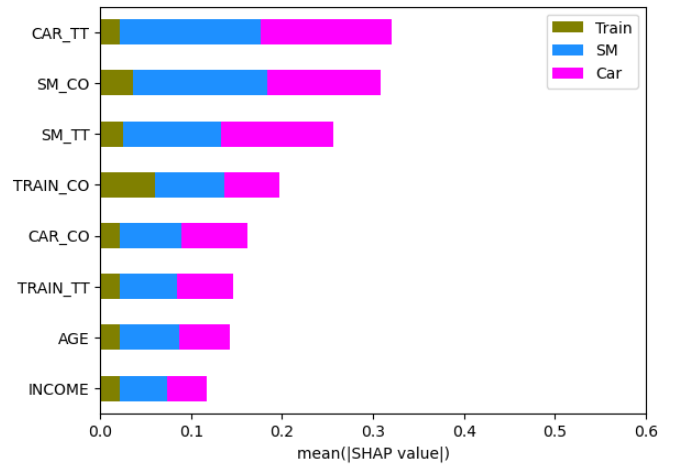
(a) ASU-DNN SHAP values



(b) F-DNN SHAP values



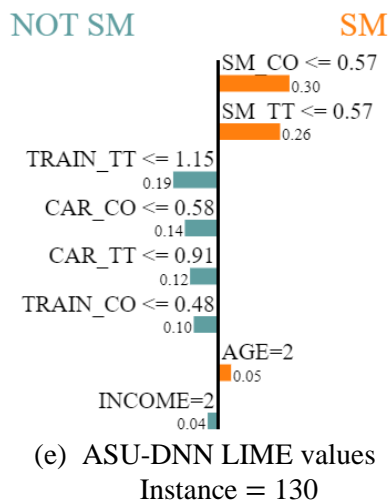
(c) Ranking SHAP values of ASU-DNN



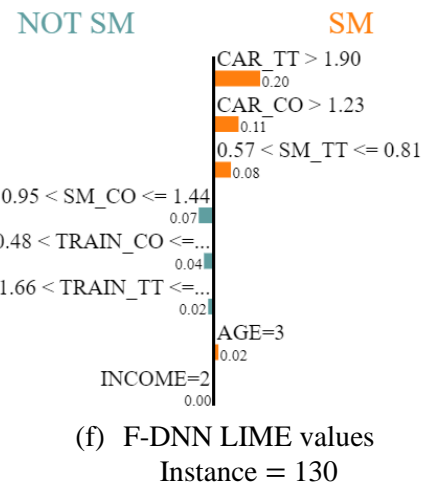
(d) Ranking SHAP values of F-DNN

Figure 1: Interpretation results of ASU-DNN and F-DNN using SHAP

Figure 2 demonstrates the local interpretability analysis with LIME for instance numbers 130 and 2137 (randomly selected). All four bar graphs reflect the contribution of each feature to the classification of respective 130 and 2137 instances. The true classes of both instances are Swissmetro. For ASU-DNN, both (e) and (g) showed that SM_TT and SM_CO have the most contribution in the output probability SM. On the contrary, (f) and (h) shows CAR_TT has the most impact on the probability of Swissmetro in F-DNN model.



(e) ASU-DNN LIME values
Instance = 130



(f) F-DNN LIME values
Instance = 130



(g) ASU-DNN LIME values
Instance = 2187

(h) F-DNN LIME values
Instance = 2187

Figure 2: Interpretation results of ASU-DNN and F-DNN using LIME

For training the DNN models, the dataset is divided into 70% training dataset and 30% test dataset. Table 1 shows the number of parameters, loglikelihood and accuracy of F-DNN and ASU-DNN for the test and train datasets. Although the performance of F-DNN in training is impressive, ASU-DNN outperforms it in terms of accuracy and loglikelihood on the test dataset. This indicates that overfitting is less likely when the RUM theory is implemented in the DNN architecture. In ASU-DNN, the number of parameters is reduced from 21,403 to 1,803 which means in this architecture, connections that are not supported by the theory are removed from the model. Therefore some spurious correlations that potentially could cause overfitting are avoided in this architecture (Yang, Chou et al. 2022).

Table 1: The loglikelihood and accuracy of F-DNN and ASU-DNN for the test and train datasets

	Number of parameters	Train dataset		Test dataset	
		Loglikelihood	Accuracy	Loglikelihood	Accuracy
F_DNN	21,403	1330.17	98.79	8861.69	67.20
ASU_DNN	1,803	3519.97	74.92	1831.45	71.78

4. CONCLUSIONS

In this study, DNN models for choice modelling are analyzed using state-of-the-art interpretation techniques. It is crucial to clarify how predictions are formed by DNN models when it comes to

artificial intelligence in discrete choice modeling. The contribution of this research is the use of recent post-hoc approaches to uncover new insights from the recently developed DNN model based on RUM theory. We apply SHAP and LIME, two of the most recent interpretation approaches, to evaluate the performance of ASU-DNN and F-DNN. The contributions of each input variable in both models F-DNN and ASU-DNN are retrieved using SHAP and LIME. The interpretation analysis of DNN models shows that DNNs with a theory-based architecture (ASU-DNN) have more consistency with the RUM theory, in contrast with conventional DNN models (F-DNN). Additionally, the results revealed that ASU-DNN could reduce overfitting by eliminating unsupported connections.

This research indicates a new research direction of using post-hoc analysis in discrete choice modeling. Future studies can concentrate on extracting information from DNN models using other post-hoc approaches such as DeepLift.

REFERENCES

- Arrieta, A. B., N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina and R. Benjamins (2020). "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI." Information fusion **58**: 82-115.
- Assi, K. J., K. M. Nahiduzzaman, N. T. Ratrouf and A. S. Aldosary (2018). "Mode choice behavior of high school goers: Evaluating logistic regression and MLP neural networks." Case studies on transport policy **6**(2): 225-230.
- Bierlaire, M., K. Axhausen and G. Abay (2001). The acceptance of modal innovation: The case of Swissmetro. Swiss Transport Research Conference.
- Chen, C., B. Liu, S. Wan, P. Qiao and Q. Pei (2020). "An edge traffic flow detection scheme based on deep learning in an intelligent transportation system." IEEE Transactions on Intelligent Transportation Systems **22**(3): 1840-1852.
- Golshani, N., R. Shabanpour, S. M. Mahmoudifard, S. Derrible and A. Mohammadian (2018). "Modeling travel mode and timing decisions: Comparison of artificial neural networks and copula-based joint model." Travel Behaviour Society **10**: 21-32.
- Goodfellow, I., Y. Bengio and A. Courville (2016). Deep learning, MIT press.
- Lipton, Z. C. (2018). "The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery." Queue **16**(3): 31-57.
- Lundberg, S. M. and S.-I. Lee (2017). "A unified approach to interpreting model predictions." Advances in neural information processing systems **30**.
- Ribeiro, M. T., S. Singh and C. Guestrin (2016). "Why should i trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.
- Sifringer, B., V. Lurkin and A. Alahi (2020). "Enhancing discrete choice models with representation learning." Transportation Research Part B: Methodological **140**: 236-261.
- Train, K. E. (2009). Discrete choice methods with simulation, Cambridge university press.
- van Cranenburgh, S. and A. Alwosheel (2019). "An artificial neural network based approach to investigate travellers' decision rules." Transportation Research Part C: Emerging Technologies **98**: 152-166.
- Wang, S., B. Mo and J. Zhao (2020). "Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions." Transportation Research Part C: Emerging Technologies **112**: 234-251.
- Wang, Y., D. Zhang, Y. Liu, B. Dai and L. H. Lee (2019). "Enhancing transportation systems via deep learning: A survey." Transportation research part C: emerging technologies **99**: 144-163.

Wong, M. and B. Farooq (2021). "ResLogit: A residual neural network logit model for data-driven choice modelling." Transportation Research Part C: Emerging Technologies **126**: 103050.

Yang, Y.-Y., C.-N. Chou and K. Chaudhuri (2022). "Understanding rare spurious correlations in neural networks." arXiv preprint arXiv:2205.05189.

Zhao, X., X. Yan, A. Yu and P. Van Hentenryck (2020). "Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models." Travel behaviour society **20**: 22-35.

An open-source framework for the robust calibration of large-scale traffic simulation models

Vishal Mahajan¹, Guido Cantelmo², and Constantinos Antoniou¹

¹Technical University of Munich, Arcisstraße 21, Munich 80333, Germany

²Transport Division, Technical University of Denmark, Denmark

SHORT SUMMARY

Large-scale traffic simulation models are a crucial tool for simulating and evaluating different transport solutions. However, due to the scale and complexity of these models, numerous parameters exist that can significantly influence their outputs. The problem of estimating these parameters is referred to as the Dynamic Traffic Assignment (DTA) calibration problem. After more than 30 years of research, several algorithms have been proposed that can - with a certain degree of success - address this challenge, even for large instances or in the presence of noisy data. Two challenges, however, remain critical today and are addressed in this paper. From a purely methodological perspective, DTA calibration is a highly under-determined problem, meaning multiple plausible solutions exist. This is particularly relevant when calibrating demand parameters. Therefore, in this paper, we propose two techniques inspired by the field of computer science that allow for enhancing robustness: *bagging* and *Stochastic Parameter Averaging* (or *SPA*). The second contribution of this research is more practical. While many algorithms have been proposed, the source codes of these algorithms are often not shared with the scientific community. As a consequence, most papers still use as a benchmark model the SPSA, an algorithm proposed roughly 30 years ago. Therefore, this study introduces an end-to-end open-source framework for DTA calibration. The model can calibrate supply and demand parameters, include state-of-the-art optimizers (*W-SPSA*, *SPSA*, *Bayesian Optimization*), an auto-tuning option to calibrate their parameters, and the bagging/SPA extension already mentioned. The conceptual framework proposed in this research is general and includes a few algorithms already. It is currently linked with the open traffic simulator *SUMO* to demonstrate its effectiveness. Researchers can use this framework as a benchmark or extend it using new simulators and optimizers. The method is tested both in controlled settings, as well as using the real-world large scale network of Munich.

Keywords: DTA calibration, ensemble, large-scale optimization, OD estimation, open-source (Topics: Transport Network Modelling, Big data analytics, Operations research applications)

1 INTRODUCTION

A transportation system is made up of different parts and their interactions, which results in travel demand and supply of transport services (Cascetta, 2001). Dynamic Traffic Assignment (DTA) simulators are advanced tools commonly used by researchers and practitioners to represent the traffic flow variations and behavioral choices in a large-scale network (Ben-Akiva et al., 2012). Due to the scale and complexity of these models, numerous parameters exist that can significantly influence their outputs. DTA calibration is the process of estimating the value of these parameters so that the difference between the simulated data (counts, travel time, speed) and observed data is minimized. The resulting optimization problem, however, is notoriously complex to solve. One issue is whether to calibrate supply and demand parameters together or separately (Toledo et al., 2014). Second, it is virtually impossible to guarantee optimal solutions in real-life settings. That is because the problem is highly under-determined (due to a large number of variables compared to a relatively small amount of observations), highly nonlinear (due to congestion dynamics), and non-convex (non-unique optimal solution). For an extensive review of these issues, as well as their solutions, we refer to (Antoniou et al., 2016). In general, many algorithms have been proposed that can partially cope with the above-mentioned challenges. However, two aspects remain critical and are addressed in this study. First, many models proposed in the literature are difficult to reproduce. This is because the code is not publicly available, often due to restrictions in data availability or in the DTA model (e.g., commercial software). As a consequence, many authors

benchmark their models against the SPSA (an algorithm proposed in 2006) or spend considerable time reproducing algorithms from the literature. In addition, while the DTA calibration process has been tested in many real-life experiments, due to the non-convex nature of the problem, obtaining robust and reliable estimates is still an open challenge. The work proposed in this research aims at answering these two questions. First, we propose an end-to-end open-source framework for DTA calibration that includes state-of-the-art solvers. The framework is designed for the open source DTA model *SUMO* (Lopez et al., 2018) and can be used to benchmark new models. Second, we propose to use parameter ensembling techniques, most notably *bagging* and *Stochastic Parameters Averaging (SPA)* to obtain more robust estimates. The framework has been successfully tested on the large-scale network of Munich, showing that it can be deployed in practice.

2 METHODOLOGY

The methodology is divided into two parts. First, we introduce the framework for DTA calibration and its main features. Then, we introduce the algorithms for SPA and bagging.

End-to-end calibration framework

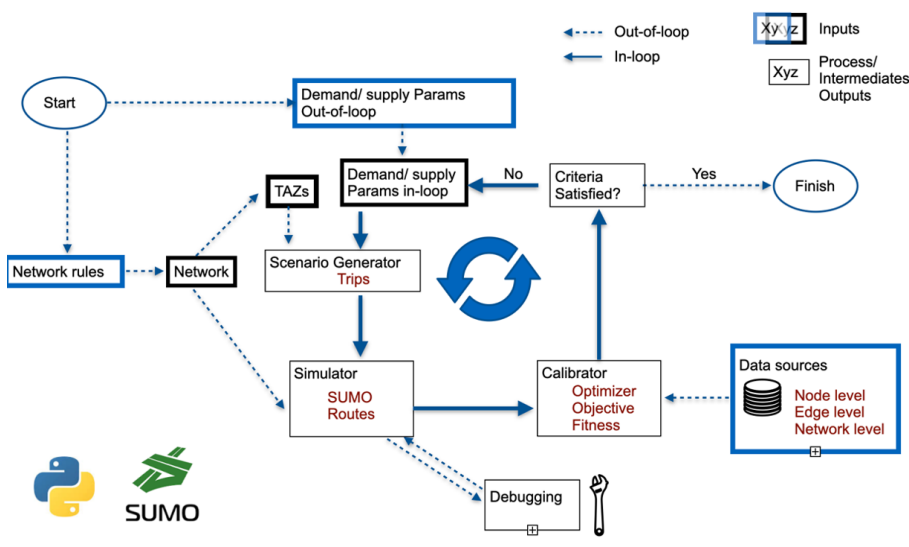


Figure 1: Experimental setup (Mahajan et al., 2023)

We developed a Python-based platform for the calibration of both demand and supply parameters of DTA models using the SUMO traffic simulator (Lopez et al., 2018). A schematic representation of the platform is shown in Figure 1. Given the simulation inputs (simulation network, traffic analysis zones, parameters), the platform estimates the parameters according to the proposed methodology. An initial Origin-Destination (OD) matrix generates trips between edges in different Traffic Analysis Zones (TAZs). The routing algorithm in SUMO assigns routes to these trips. Automatic or online routing is used for the traffic assignment. The parameters influencing the routing of vehicles are *re-routing probability*, *re-routing period*, and *re-routing adaptation steps*. Travel time of different edges can be scaled using the parameter *edge priority factor*. The parameters which affect the delays are *junctions flow penalty* at junctions and *unsignalized junction penalty*. We use Bayesian optimization for calibrating these selected supply parameters.

Concerning the demand parameters, we implemented a state-of-the-art optimizer (W-SPSA Antoniou et al. (2015)) by extending the Python SPSA implementation by Mayer (2017). Further, considerable time and manual effort is usually spent in fine-tuning the hyperparameters of the SPSA to enhance calibration performance. Therefore, we included in the framework an automatic tuning function to automatically optimize the SPSA hyper-parameters. The auto-tuning function uses the assignment matrix to create an analytical approximation of SUMO. Bayesian optimization is used to find the optimal parameters that reduce the error of the analytical model.

Finally, two additional extensions have been included, namely bagging and SPA. These methods, shortly described in the next sub-section, have been tested using W-SPSA. However, similarly to

most of the features described in this section, they can be combined with other optimizers once they are implemented in the framework. The complete platform is implemented using Python and is available on GitHub (<https://github.com/vishalmhjn/actrys>). In short, the model currently implements the following features:

- Three optimizers (Bayesian Optimization, SPSA, W-SPSA)
- A heuristic model to pre-process historical demand and remove bias
- The framework sequentially optimizes supply and demand parameters
- It includes an auto-tune function for the hyper-parameters of the optimizer
- It includes parameter ensembling to enhance robustness

Parameter ensembling to reduce estimators’ variance

This sub-section introduces the methodological contribution of this research. The methodology focuses only on improving the estimates for the demand parameters, therefore we refer to this sub-task as the *OD estimation* problem. Without loss of generality, the OD estimation problem can be operationalized as follows:

$$\underset{\mathbf{X}}{\text{minimize}} \sum_{t=1}^T [w_1 z_1(\mathbf{M}_t^o, \mathbf{M}_t^s) + w_2 z_2(\mathbf{X}_t, \mathbf{X}_t^a)] \quad (1)$$

In Equation 1, \mathbf{M}_t^o and \mathbf{M}_t^s refer to the observed and simulated traffic data. Similarly \mathbf{X}_t and \mathbf{X}_t^a refer to the estimated and a-priori values of the demand (OD) parameters. Finally z_1 , and z_2 are functions that measure the discrepancy between simulated and observed data (Goodness-of-Fit, GoF), while w_1 and w_2 are weights for these functions. The dependence between simulated traffic data and the OD matrices is directly obtained from the DTA traffic simulator (SUMO, in this study). Other constraints that are often applied in practice are ignored here for ease of reading.

The OD estimation problem is highly under-determined, meaning that many solutions exist that are theoretically feasible for a given optimization formulation. This also implies that even state-of-the-art models such as W-SPSA will unavoidably find a local solution for Equation 1, resulting in parameters with considerable variance. In this study, we hypothesize that, due to variance in the spatiotemporal demand patterns, variance in sampling distribution or measurement errors can be considered as a manifestation of the desired (or “true”) solution. Parameter averaging, such as in the bagging and averaging techniques, can help to cancel out some of the variances in the individual solution so that the averaged solution is closer to the desired one.

Therefore, we introduce two algorithms that can be used to combine bagging and SPA within the OD estimation problem.

W-SPSA with Bagging: when using bagging (*B-W-SPSA*) (Breiman, 1996), we run multiple estimators, such as W-SPSA, (in parallel or in serial order), and record the final estimates of each of the runs or cycles (since SPSA is stochastic in nature). To promote exploration, at each run we perturb the initial seed matrix - i.e., $\hat{X}^a \leftarrow X^a + \epsilon$, where ϵ is a normally distributed error term. Each run leads to different local optima. The final result is obtained as an average of these results

W-SPSA with Stochastic Parameter Averaging: The Stochastic Parameter Averaging (SPA) is a new algorithm applied in this research and inspired by the Stochastic Weight Averaging (SWA) (Izmailov et al., 2018), used in the field of computer science to find the weights of Deep Neural Networks (DNNs) while avoiding local minima. In SPA, the optimization is divided into two parts. In the first phase, the optimizer (W-SPSA, in this case) reduces the error in Equation 1. In this case, there is no difference between normal W-SPSA and W-SPSA with SPA. Phase two begins once the model achieves a local solution. In this step, the gain coefficients (i.e, the SPSA hyper-parameters) are reset. The next optimization cycle uses the iterate from the previous cycle as the initial parameters, and hence it is referred to as “warm restart”. Resetting of SPSA gain coefficients resembles the cyclic learning rate, and allows the algorithms to explore new solutions.

3 RESULTS AND DISCUSSION

We present the results of our optimization model for two case studies. We compare the results of three models, namely *W-SPSA*, *B-W-SPSA*, and *SPA*. Two scenarios are analyzed:

1. **Scenario 1:** Analytical simulator with synthetic sensor counts. A randomly generated assignment matrix is used for mapping OD flows (randomly sampled using a distribution function).
2. **Scenario 2:** SUMO and Munich regional network with synthetic sensor counts data. Given OD flows (Moeckel et al., 2020) are simulated and corresponding sensor counts are recorded as desired counts. The Munich regional network is divided into 73 zones resulting in 5256 OD pairs. The network consists of a total of 8761 links.

Scenario one serves to visually explain how bagging and SPA build on and improve the performances of W-SPSA. Scenario 2 tests how parameter ensembling performs when using SUMO on a large, real network. In both cases, the Weighted Average Percentage Error (WAPE) and the Root Mean Squared Error (RMSE) are used as evaluation criteria for OD fitness and count fitness.

In **Scenario 1**, we use a random perturbation R^x to introduce bias in the OD matrix. Then, we use the proposed algorithms to estimate the optimal parameters. The objective function minimizes the error with respect to historical OD flows and traffic counts. In Figure 2, we show the contours of the OD fitness errors for single W-SPSA estimates, SPA estimates, and bagged estimates - for selected ODs. Due to high dimensional optimization, fitness error is influenced by thousands of the demand parameters, so the plot shows the conditional error (because it depends on multiple parameters) region with the values of the pair of zones on X and Y-axes. The two columns in this figure correspond to levels of random perturbation R^x 30%, and 90%. It is evident, that in both cases, the single estimates are scattered in the region, but the averaged estimates from SPA and bagging are lying with the region of lower errors as compared to the single W-SPSA estimates. Intuitively, this illustrates how bagging/SPA help to reduce the variance in the estimates from single W-SPSA estimates. While the two models overall achieve the objective of reducing variance in the estimates, our results show that - for the same number of objective functions evaluation - bagging systematically outperforms SPA. Therefore, in Scenario 2 we will mostly focus on comparing bagging and traditional W-SPSA.

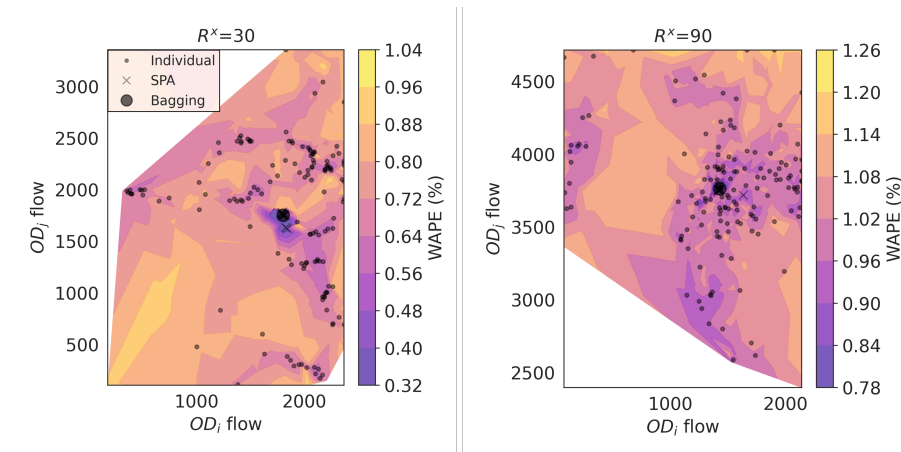


Figure 2: Contour plots showing the parameter values of the objective function for selected pair of the zones at different values of the R^x .

Scenario 2: We show the results of the calibration for the Munich scenario using the SUMO platform in Table 1. In all experiments, we assume a systematic bias $B^x = 0.60$ and a relatively smaller factor for randomness ($R^x = 20\%$). By default, we use only sensor counts in the objective function. However, we also simulate the case where we introduce artificial randomness in the sensor counts to mirror data errors. In this case, we also introduce speeds in the objective function to test the open-source framework when multiple data sources are available.

We use the W-SPSA with manual optimization of the hyper-parameters as the baseline. The corresponding improvement in speed fitness and OD fitness are 36.65% and -59.71%, respectively. A negative value of improvement tells that the estimated OD is worse than the initial OD values, which points to the ineffectiveness of the optimization. When using W-SPSA combined with bagging (*B-W-SPSA*), all error metrics substantially improve. The counts and speeds error improve by about 80% and 44%, respectively. More notably, the error in the OD flows is also reduced of about 20%. To further test our hypothesis that bagging can reduce variance by filtering noise, we can look at the experiments where some sensor noise was introduced. When the noise is low

Table 1: Results of the Munich scenario with synthetic data

Model	B^x	Sensor noise	% Improvement WAPE (RMSE)		
			Count	Speed	OD
W-SPSA	0.6	0	68.11 (69.01)	36.65 (50.27)	-59.71 (-82.84)
B-W-SPSA	0.6	0	82.34 (84.31)	44.36 (55.72)	20.08 (12.88)
B-W-SPSA*	0.6	15%	61.54 (59.53)	48.85 (65.22)	9.06 (-7.06)
B-W-SPSA*	0.6	30%	41.29 (39.06)	32.47 (33.56)	0.33 (-19.02)
B-W-SPSA*	0.6	45%	25.07 (23.79)	28.43 (14.43)	-0.38 (-38.48)

*: both counts and speeds are used in the objective function

(15%), the model still performs better than W-SPSA (except for the counts). When sensor noise increases, performances deteriorate. However, this is expected. More importantly, *B-W-SPSA* has a low error on the OD flows even for high sensor noise, which highlights the robustness of the framework when SPSA is combined with bagging.

4 CONCLUSIONS

Robust and efficient algorithms for DTA calibration are required for virtually any application where DTA models are deployed. However, obtaining robust estimates is still an open challenge in the research community. This research contributes to this direction in two ways. First, we provide an open-source framework for DTA calibration. The framework already includes several features, and can easily be extended to new DTA models and/or optimization algorithms. We hope that this can be useful to other researchers when developing new and better algorithms. Second, we propose two parameter ensembling techniques, the SPA and the bagging, and test them with the W-SPSA algorithm. Both techniques show that ensembling can lead to more robust estimates compared to W-SPSA. Conceptually, SPA and bagging are similar, as they both achieve better estimates by averaging results. However, SPA does this sequentially, while bagging runs independent optimizations. In practice, bagging can be more convenient when it is possible to run several simulations in parallel, while SPA can be more suited to explore different regions sequentially. We tested our framework on the large-scale network of Munich, Germany, showing how ensembling techniques allow filtering noise and estimating robust solutions. In the future, we hope to collaborate with other researchers, to include other algorithms, and interface the platform with other DTA models.

REFERENCES

- Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., ... others (2016). Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation. *Transportation Research Part C: Emerging Technologies*, 66, 79–98.
- Antoniou, C., Lima Azevedo, C., Lu, L., Pereira, F., & Ben-Akiva, M. (2015). W-spsa in practice: Approximation of weight matrices and calibration of traffic simulation models. *Transportation Research Part C: Emerging Technologies*, 59, 129-146. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0968090X15001710> (Special Issue on International Symposium on Transportation and Traffic Theory) doi: <https://doi.org/10.1016/j.trc.2015.04.030>
- Ben-Akiva, M. E., Gao, S., Wei, Z., & Wen, Y. (2012). A dynamic traffic assignment model for highly congested urban networks. *Transportation Research Part C: Emerging Technologies*, 24, 62-82. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0968090X12000186> doi: <https://doi.org/10.1016/j.trc.2012.02.006>
- Breiman, L. (1996, Aug 01). Bagging predictors. *Machine Learning*, 24(2), 123-140. Retrieved from <https://doi.org/10.1007/BF00058655> doi: 10.1007/BF00058655

- Cascetta, E. (2001). *Transportation systems engineering: Theory and methods*. Springer New York, NY. doi: <https://doi.org/10.1007/978-1-4757-6873-2>
- Izmailov, P., Podoprikin, D., Garipov, T., Vetrov, D., & Wilson, A. G. (2018). *Averaging weights leads to wider optima and better generalization*. arXiv. Retrieved from <https://arxiv.org/abs/1803.05407> doi: 10.48550/ARXIV.1803.05407
- Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., ... Wießner, E. (2018, November). Microscopic traffic simulation using sumo. In *The 21st ieee international conference on intelligent transportation systems* (pp. 2575–2582). IEEE. Retrieved from <https://elib.dlr.de/127994/>
- Mahajan, V., Cantelmo, G., & Constantinos, A. (2023). One-shot heuristic and ensembling for automated calibration of large-scale traffic simulations. *preprint*. Retrieved from <https://mediatum.ub.tum.de/node?id=1701188>
- Mayer, A. (2017). Noisyopt: A python library for optimizing noisy functions. *Journal of Open Source Software*, 2(13), 258. Retrieved from <https://doi.org/10.21105/joss.00258> doi: 10.21105/joss.00258
- Moeckel, R., Kuehnel, N., Llorca, C., Moreno, A. T., & Rayaprolu, H. (2020, Feb 25). Agent-based simulation to improve policy sensitivity of trip-based models. *Journal of Advanced Transportation*, 2020, 1902162. Retrieved from <https://doi.org/10.1155/2020/1902162> doi: 10.1155/2020/1902162
- Toledo, T., Kolechkina, T., Wagner, P., Ciuffo, B., Lima Azevedo, C., Marzano, V., & Flötteröd, G. (2014, 09). Network model calibration studies. In (p. 22). CRC Press.

Use of Origin-Destination data for calibration and spatialization of synthetic travel demand

Benoît Matet^{*1,2}, Etienne Côme¹, Angelo Furno², Sebastian Hörl^{1,3}, and Latifa Oukhellou¹

¹Université Gustave Eiffel, COSYS, GRETTIA, France

²Université Gustave Eiffel, COSYS, LICIT-ECO7, France

³IRT SystemX, France

SHORT SUMMARY

The dynamics of urban transportation can be understood with activity-based models, which rely on synthetic travel demand data to get a comprehensive understanding of urban mobility. These data are usually derived from small population samples and surveys, which may be expensive and do not adequately cover the spatial trajectories of the users. In this paper, we explore the use of a time-dependent origin-destination (OD) matrix derived from mobile phone data for the attribution of locations in a synthetic population for the city of Lyon, France. OD matrix data can also mitigate uncertainties or outdated information in travel surveys regarding flows by time of day and between zones. The resulting population enrichment is measured in terms of fit to the input mobility data.

Keywords: activity-based modeling, data fusion, multi-modal transportation, origin-destination matrix, synthetic demand

1 INTRODUCTION

One way of acquiring insight into a city's transportation system is to simulate an adequately generated synthetic population of agents. Assuming the synthetic population is representative of the real one and the simulator accurately describes how travelers make decisions, the result of the simulation should be a comprehensive description of the city's mobility dynamics. The synthetic population should match the known marginal distributions of the real population while staying as close as possible to the joint distribution of variables observed from an input population sample. In activity-based models, each agent must also have an agenda, stating a chain of activities (e.g. home, work, study, shopping, other), corresponding times and transport modes, and a chain of locations for each activity. The agendas should be likely at the individual level and ultimately match flows observed from other data sources.

Hörl & Balac (2021) propose a general pipeline to create a synthetic population with socioeconomic variables and activity chains from multiple data sources. In a first step, a population sample is used as input to set the socioeconomic variables of the agents. When the input sample is small, the approaches proposed by Sun & Erath (2015); Sun et al. (2018) allow modeling the underlying probability law and sampling the desired number of agents from it. In a second step, these agents can be extended with attributes that are not in the population sample, ensuring that the available marginal statistics are matched. Activity chains, without locations, are then assigned to the agents. These activity chains can either be available from a separate mobility survey or generated by approaches such as the ones proposed by Joubert & de Waal (2020); Anda et al. (2020). Activity locations are usually separated into primary locations, such as home or workplace, and secondary locations for other types of activity. While primary locations can be drawn in the first steps of the generation, secondary locations are usually sampled through a process aiming at mimicking the decision rules of an individual choosing where to go for a particular task (Ma & Klein (2017); Hörl & Axhausen (2021)).

In this work, we propose to expand the state-of-the-art pipeline with a novel data-driven approach for the location of activities, leveraging mobile data from the telecommunications provider Orange. Mobile data have recently become an interesting alternative to mobility surveys, as they are much cheaper and faster to generate. They can also be more representative regarding spatial features,

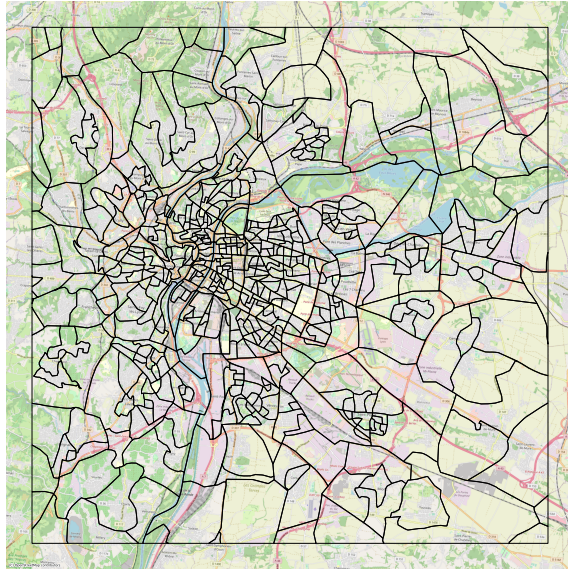


Figure 1: Map of the partitioning of the study area of the city of Lyon (France), Background: OpenStreetMap

as around 30% of people in France are customers of Orange, compared with 1% of persons reached by mobility surveys. Once re-scaled, the data gives us a reliable estimation of the trips of the total population. However, they lack the level of detail offered by surveys, which makes it interesting to use the data sources together in order to obtain a dataset that would be both detailed and easily adaptable to new observations.

Mobile data can take the form of full trajectories of users, on which rely approaches proposed by Zilske & Nagel (2015); M. Yin et al. (2017). However, full trajectories are computationally heavy and feature a significant privacy risk. We thus derive a time-dependent Origin-Destination (OD) matrix from mobile phone reconstructed trajectories (Bonnetain et al. (2021)), giving only a map of flows between the zones of our study area for each time step of the day. OD matrices are lighter and more manageable than whole trajectory datasets and can be fully anonymized as shown in L. Yin et al. (2015); Matet et al. (2021). This makes them more readily available and safer regarding the privacy of transportation users. As our first contribution, we interpret this OD matrix as a transition probability between two locations, which we use to sample the location chain. Second, we use the total number of trips in the OD matrix for each time of the day as a target that the agendas of the population should match.

In the case of France, valuable data sources for the complete generation process are available to researchers upon request. The census performed by the French statistics institute INSEE features socioeconomic variables of a population sample large enough to be representative of the whole population. The activity chains are taken from the Enquête Ménage Déplacements (EMD) performed by the French agency for urban planning Cerema (2015), which details the complete agendas of 25,203 persons in the Lyon region.

2 METHODOLOGY

The first steps of the synthetic travel demand generation are inspired from Hörl & Balac (2021), which focuses on Île-de-France and Paris, now with the target area being Lyon and its surroundings. Our study area takes the form of a square 25,600 meters wide centered around Lyon, divided into 515 distinct zones which either represent municipalities or spatial units (IRIS) common to many statistical analyses in France. Figure 1 shows the partitioning of the study area.

In the case of France, the census performed by the French statistics institute INSEE constitutes a population sample extensive enough so that no additional synthesis is required. Each of the 487,628 rows for our study zone features socioeconomic variables and a non-integer scaling coefficient for a total of 1,366,072 persons. To interpret the data as an integer number of agents, we apply the TRS approach (Truncate, Replicate, Sample) introduced by Lovelace & Ballas (2013). Each

coefficient is stochastically rounded up with probability equal to its decimal part, and rounded down otherwise. The result is a population of 1,372,566 agents described by the variables detailed in Table 1. Note that we choose not to retain the household structures of the population. In a generative approach, this information would be necessary in the downstream modeling steps. In our approach, activity chains are assigned by sociodemographic attributes, hence only individuals are relevant.

Table 1: Socioeconomic variables for the synthetic population

Variable	Modalities
Home zone	515 zones in the study area
Age	0-17; 18-29; 30-59; 60+
Gender	male or female
Occupation	jobless; farmer; independant; executive; employee; intermediate professions; worker; student; retired
Has car	yes or no
Home status	Owner; Tenant; Social housing

The assignment of activity chains from the EMD travel survey to agents from the census follows a process inspired from statistical matching (D’Orazio et al. (2006)). As in Hörl & Axhausen (2021), this process is close to a regular join between relational databases with a join key of multiple variables, except that each census row is required to match at least 20 EMD rows. For census rows for which this is not the case, we successively remove the last variable from the join key until more than 20 EMD rows match. Then, one of these EMD rows is randomly drawn, and its activity chain is assigned to the person described by the census row. This process guarantees diversity in the assignment of chains to population agents while joining as many variables as possible so that the agents have activity chains relevant to their characteristics. The variables used in the join key are described in Table 2.

Table 2: Variables used in the statistical matching to assign activity chains from EMD to agents from census

Variable	Number of modalities	% of agents with more than 20 matches
Age	4	100%
Gender	2	100%
Occupation	9	95%
Has car	2	94%
Home status	3	90%
Canton	8	75%

Note that although the census is the same, the EMD differs between Paris and Lyon. We cannot use the income in the join key as is done by Hörl & Balac (2021) because the EMD for Lyon does not feature it. We replace it with what is arguably a proxy for one’s wealth, i.e., whether the person is the owner, tenant of their home, or lives in social housing. Similarly, instead of the department used in Île-de-France, we use the canton, an intermediary administrative zoning system between our zoning and the department. The resulting population comprises agents, each equipped with a chain of trips specifying the purposes, time of day, and transport modes, as is detailed in Table 3. Note that the day is divided into time steps of one or more hours to be consistent with the division used for anonymization in the mobile data. The activity chains contained in the EMD feature up to 12 trips during the day.

Table 3: Variables describing trips

Variable	Modalities
Trip purpose	home; work; study; shopping; personal
Time step	0h-2h; 2h-5h; 5h-7h; 7h; 8h; 9h; 10h-12h; 12h-14h; 14h-16h; 16h; 17h; 18h; 19h; 20h-22h; 22h-0h;
Transport mode	foot; bicycle; car or motorcycle; public transport

Rescaling

While the previous steps correspond to existing approaches for travel demand synthesis, we propose the following steps that make direct use of a large-scale time-dependent Origin-Destination matrix that is featured in this research. Due to data discrepancies, the number of trips performed in each time step of the day by our agent population does not match the number of trips observed from mobile data. This is illustrated in Figure 2, where each point represents the marginal value of a joint socioeconomic attribute (in blue) or the total number of trips taken for a given time step (in yellow). The x-coordinate of the point is the truth value from the available data, i.e., the marginals from the census in the case of blue points and the number of trips from mobile data for yellow points. The y-coordinate is the total measured in our synthetic population. We perform rescaling via Iterative Proportional Updating (IPU) as introduced by Ye et al. (2009), with the adaptation that instead of a population composed of households divided into persons, we have a population of persons divided into trips. The resulting population is consistent both with the socioeconomic composition of the real population and the number of trips they take at each time step of the day.

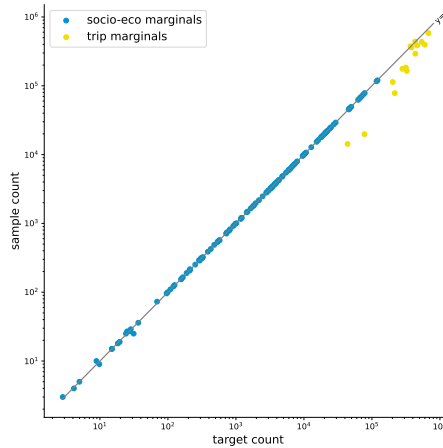


Figure 2: Marginals of the synthetic population versus what is expected from the data sources. Each blue dot is the volume for a socioeconomic modality. Each yellow dot is volume of trips for a given time step.

Spatialization

To be complete, the synthetic population requires locations for each activity in their agenda. We interpret the OD matrix as a probability table of destinations D given the origin O and time T : $P(D|O, T)$. For each agent, we model the chain of locations during the day as a Markov chain with transition probability $P(D|O, T)$.

The activity purpose may specify that the agent is going home, in which case the home area defined in the first step of the generation from the census is deterministically assigned as a location. In contrast with state-of-the-art methods, we do not rely on pre-defined work or study places, as we do for home. This is justified by the fact that, as 45% of trips in the activity chains of our population are commute trips, pre-defining the work and study places would not only half the use of the data we aim at exploiting but also invalidate it as our mobile data would still contain the distribution of the sum of commute and non-commute trips but be used only for non-commute trips.

However, even without pre-defining them, we still want to ensure that work and study activities have the same location when they appear more than once in the activity chain. The resulting probability law for the chain of locations is described by a Markov chain with fixed states for home and linked states for work and study. Because of these dependencies between states, we cannot sample from it as we would a regular Markov chain. We resort to Gibbs sampling to generate a chain of locations. This drawing process can only manage a fixed activity chain with a fixed pre-defined home zone, meaning that two agents with either a different activity chain or a different home require each a distinct Gibbs sampling. We obtain 478,963 distinct processes to run, which is computationally expensive because of the required warm-up phase: Gibbs sampling is considered

to yield the true joint distribution only after having discarded the first thousand samples. We reduce the computing time by observing that most activity chains have a low autonomy, i.e. they feature a limited number of unknown locations between two fixed home activities. Segments of the activity chain that are separated by a fixed state are independent of each other and follow a distribution of much lesser dimension. In fact, segments of only one unknown activity do not require Gibbs sampling, and segments of two activities only require a limited warm-up.

As a refinement, we consider the transport modes M from trips in the activity chains to obtain a transition probability $P(D|O, T, M)$. The transport mode is not specified in the mobile OD data, but can be integrated into the sampling process using Bayes rule: $P(D|O, T, M) \propto P(M|O, D, T) \times P(D|O, T)$. As an estimation of $P(M|O, D, T)$, we consider that the mode M depends only on the distance L between origin O and destination D . Using all trips from the EMD, we obtain an empirical probability distribution $P(M|L(O, D))$ discretized to individual 1 km bins.

3 RESULTS

As the synthetic population is expected to behave like the real population, we evaluate it on all the available indicators about the composition or mobility of the real population. Note that our various input sources can be inconsistent between themselves, e.g. on the number of trips for each time step or even the socioeconomic composition of the population between the census and the EMD. As such, it is already a satisfying result for the synthetic population to match indicators derived from our own input data.

Socioeconomic composition of the population

In Figure 3, we illustrate how the socioeconomic composition of our population matches the official census better than the transport survey EMD. As the EMD represents only a small number of people, it is normal that the totals have a higher variance. Our synthetic population can then be seen as a version of the EMD that agrees with the census on the socioeconomic distribution of the population.

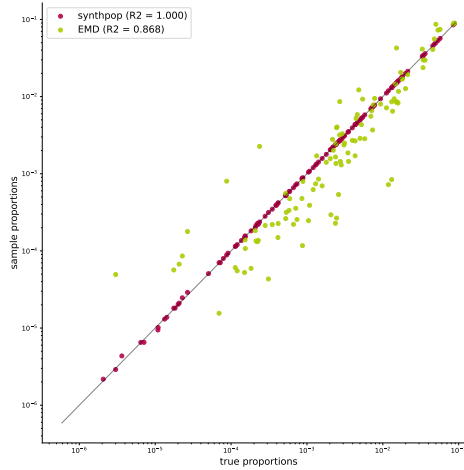


Figure 3: Matching of the socioeconomic marginals of our rescaled synthetic population (in red) compared to the transport survey EMD (in green).

Number of trips by hour

In Figure 4, we illustrate the distribution of trips during a typical day as measured from mobile data (in orange), our synthetic data (in magenta), and the survey EMD (in green). In this case, we consider the observed volumes from the mobile data to be the ground truth. We see that our population can also fit the ground truth source better than the official survey EMD. In particular, the EMD seems to overestimate the volumes of the morning, midday, and evening peaks while underestimating the volumes during the rest of the day.

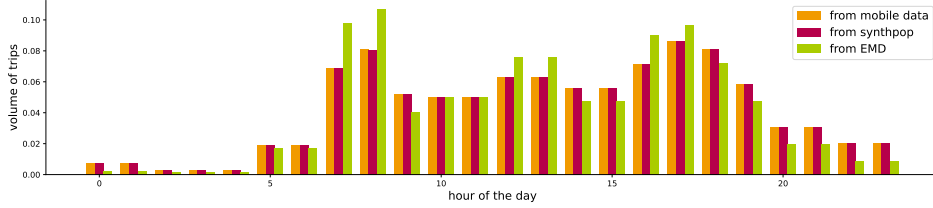


Figure 4: Distribution of trips taken during the day.

Conditional probability of destinations

In Figure 5, we illustrate how the probability of destinations given origin and hour fits the probability table derived from the OD matrix. Each point corresponds to a combination of destination, origin, and time of the day. Its x-coordinate is the probability $P(D|O, T)$ observed in the OD matrix, while its y-coordinate is the same probability as observed in the trips of our synthetic population. Interestingly, we observe a bi-modality in the ground truth: it seems that each origin has a restricted list of favorite destinations forming the upper-right cloud and a list of secondary destinations in the lower-left cloud. We see that this bi-modality is retrieved in our synthetic population and that, overall, our synthetic demand is highly correlated to the actual observations.

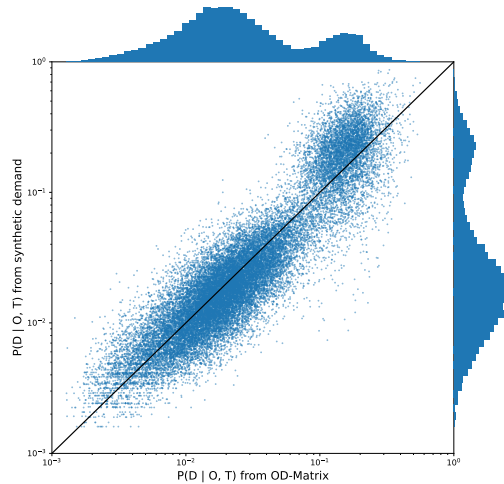


Figure 5: $P(D|O, T)$, from our synthetic data w.r.t. ground truth value from the OD matrix. The black line represents $y = x$.

Limitations

While it can be expected that our synthetic demand data matches well the distribution $P(D|O, T)$, we observe that the distribution of trips $P(O, D|T)$ is not the same as in the input OD matrix. This is because although we sample the destinations using $P(D|O, T)$ as a transition probability, we have no mechanism to make sure that the right number of agents leave each individual origin O at each time step T . This problem can be addressed by decomposing each time slice of the transition matrix into a sum of transition matrices depending on O , T , and on the time step of the next trip of the agent. This new decomposition of the OD matrix in our future work amounts to adding explanatory variables to the mobile data in the same fashion as we added the transport mode. By carefully choosing such a decomposition, we can make sure the agents taking a trip before time step T are assigned to destinations such that the map of agents leaving for a trip at time step T corresponds to the map of origins of trips in the OD matrix.

4 CONCLUSIONS AND PERSPECTIVES

This paper illustrates how a time-dependent OD matrix from mobile data can improve the generation process for synthetic travel demand. As we consider mobile data to be closer to the ground truth than surveys regarding the number of trips by the time of day and between zones, they are a valuable asset to take into account in addition to already available sources. Our approach makes use of the highly valuable activity chain structures that are already featured in the surveys and successfully improves on them with mobile data, using an OD matrix both as a rescaling target and a basis for spatialization.

As OD matrices can be anonymized, our approach is also relevant to leverage the richness of mobile data without the computational cost nor the privacy hazard of full trajectory data. However, a huge challenge lies in the fact that OD matrices derived from mobile phones are not dependent on anything other than time. As such, correlations between the spatial characteristics of the trips (such as average commute distance) and other variables are hard to capture. A more detailed OD matrix could be derived from the initial trajectories without additional data: for example, recurrent daily trips could be identified as commute patterns, resulting in two separated OD matrices for commute and non-commute.

ACKNOWLEDGEMENTS

This research is supported by the French ANR research projects MOBITIC (grant number ANR-19-CE22-0010) and PROMENADE (grant number ANR-18-CE22-0008).

REFERENCES

- Anda, C., Ordonez Medina, S. A., & Axhausen, K. (2020, 09). Synthesising digital twin travellers: Individual travel demand from aggregated mobile phone data. doi: 10.3929/ethz-b-000442517
- Bonnetain, L., Furno, A., El Faouzi, N.-E., Fiore, M., Stanica, R., Smoreda, Z., & Ziemlicki, C. (2021). Transit: Fine-grained human mobility trajectory inference at scale with mobile network signaling data. *Transportation Research Part C: Emerging Technologies*, 130, 103257.
- Cerema. (2015). lil-1023: Enquête ménage déplacement, lyon / aire métropolitaine lyonnaise.
- D’Orazio, M., Zio, M., & Scanu, M. (2006). *Statistical matching: Theory and practice*. doi: 10.1002/0470023554
- Hörl, S., & Axhausen, K. (2021, 10). Relaxation–discretization algorithm for spatially constrained secondary location assignment. *Transportmetrica A: Transport Science*. doi: 10.1080/23249935.2021.1982068
- Hörl, S., & Balac, M. (2021). Synthetic population and travel demand for paris and ile-de-france based on open and publicly available data. *Transportation Research Part C: Emerging Technologies*, 130, 103291. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0968090X21003016> doi: <https://doi.org/10.1016/j.trc.2021.103291>
- Joubert, J., & de Waal, A. (2020, 09). Activity-based travel demand generation using bayesian networks. *Transportation Research Part C Emerging Technologies*, 120.
- Lovelace, R., & Ballas, D. (2013, 09). ‘truncate, replicate, sample’: A method for creating integer weights for spatial microsimulation. *Computers Environment and Urban Systems*, 41. doi: 10.1016/j.compenvurbsys.2013.03.004
- Ma, T.-y., & Klein, S. (2017, 09). Bayesian networks for constrained location choice modeling using structural restrictions and model averaging. *European Journal of Transport and Infrastructure Research*, 18.
- Matet, B., Côme, E., Furno, A., Bonnetain, L., Oukhellou, L., & El Faouzi, N.-E. (2021, 01). A lightweight approach for origin-destination matrix anonymization. In (p. 487-492). doi: 10.14428/esann/2021.ES2021-56

- Sun, L., & Erath, A. (2015, 10). A bayesian network approach for population synthesis. *Transportation Research Part C Emerging Technologies*, 61, 49-62.
- Sun, L., Erath, A., & Cai, M. (2018). A hierarchical mixture modeling framework for population synthesis. *Transportation Research Part B: Methodological*, 114, 199-212.
- Ye, X., Konduri, K., Pendyala, R., Sana, B., & Waddell, P. (2009, 01). Methodology to match distributions of both household and person attributes in generation of synthetic populations.
- Yin, L., Wang, Q., Shaw, S.-L., Fang, Z., Hu, J., Tao, Y., & Wang, W. (2015, 10). Re-identification risk versus data utility for aggregated mobility research using mobile phone location data. *PLOS ONE*, 10(10), 1-23. Retrieved from <https://doi.org/10.1371/journal.pone.0140589> doi: 10.1371/journal.pone.0140589
- Yin, M., Sheehan, M., Feygin, S., Paiement, J.-F., & Pozdnoukhov, A. (2017, 05). A generative model of urban activities from cellular data. *IEEE Transactions on Intelligent Transportation Systems*, PP, 1-15. doi: 10.1109/TITS.2017.2695438
- Zilske, M., & Nagel, K. (2015). A simulation-based approach for constructing all-day travel chains from mobile phone data. *Procedia Computer Science*, 52, 468-475.

Evaluating the impact of Free Public Transport using agent-based modeling: the case-study of Luxembourg

Federico Bigi^{*1}, Nicola Schwemmler¹, and Francesco Viti²

¹PhD Student, Faculty of Science, Technology and Medicine (FSTM), University of Luxembourg,
Luxembourg

²Associate Professor, Faculty of Science, Technology and Medicine (FSTM), University of
Luxembourg, Luxembourg

SHORT SUMMARY

Luxembourg's high car ownership per household, combined with a strong population growth and high share of cross-border commuters, led to systematic traffic congestion issues all over the country. To address this problem, the government made public transport (PT) free in 2020 by eliminating second-class fares, a policy that attracted great interest but also led to controversial opinions on its real impact on car use. Notably, since free PT requires no ticketing, there is a problematic lack of passenger data that would allow assessing the policy impacts. Therefore, in this study a MATSim scenario was developed and used to evaluate the impact of this policy under realistic settings. The population was generated using data from a national travel survey collected in 2017 and the free PT scenario was compared against a benchmarking scenario where public transport was still not free, allowing to collect and analyse different KPIs. The simulation showed that the policy brought significant benefits in terms of Passenger Kilometer (PKT) and Hours (PHT) Travelled for PT, and a decrease in car usage in favor of PT especially for cross-border commuters for the parts of their journey that take place inside of Luxembourg. Nonetheless, this study found that the overall impact of Free PT was not significant enough to strongly impact the congestion levels, with the high car ownership rates being one possible reason for this resistance.

Keywords: Agent-based modeling; Free public transport; MATSim.

1 INTRODUCTION

The development of new on-demand services, the increasing popularity of car and bike sharing, and the growing trend towards multi-modal travel pose a series of challenges for research groups around the world. These trends, combined with new transport policies, have the potential to significantly impact travel behavior, making the forecast of the impact of such changes essential for policymakers to avoid investing substantial amounts of resources in ineffective measures.

This is particularly relevant for Luxembourg, where in March 2020 all second-class fares were canceled, becoming the first country in the world to provide nationwide free public transportation (PT). The goal of this policy was to reduce the high percentage of trips made by car, which was around 79% in 2017, with 33% being performed by cross-border workers who reside in France, Germany, and Belgium. This convergence of travelers onto specific segments of the network leads to severe congestion and significant increases in travel time during peak hours.

Understanding the impact of transport policies such as free-PT is a non-trivial task, since it requires estimating the mode choice elasticity for all trips performed by an individual in a context where many other factors may have caused an equal or even stronger effect (e.g. the COVID19 pandemic). Moreover, such policies are expected to have an impact on the whole trip chain, hence requiring more sophisticated approaches than the traditional trip-based approach. Conversely, agent-based demand modeling designed to further test new transport models allows a better forecast of traffic demand on the network, and a more accurate policy evaluation. The main goal of this study is to investigate the impact of the policy put into action in 2020, using the MATSim agent-based simulation (Horni et al. (2016)) scenario developed for Luxembourg.

The developed scenario simulates the main modes of transportation including cars, PT, biking, and walking. It provides detailed schedules for public transport services such as buses, trams, and trains. The transportation network generation was set up using two open data sources: network information from OpenStreetMap (OSM) and public transport schedules generated from the

General Transit Feed Specification (GTFS). The population generation and trip chains, as well as the origin-destination matrices, were obtained from the LuxMobil travel survey conducted in 2017. After the population was generated and the MATSim scenario was calibrated using the data from the travel survey, as well as traffic counts, the impact of the Free Public Transport policy in Luxembourg was analyzed. The aim of this study is therefore to evaluate the Free Fare Public Transport (FFPT) policy in Luxembourg. The study will tackle the evaluation of this policy from various perspectives, such as travel time reduction, travel distance decrease, modal split shift, and cost-benefit analysis.

Literature Review

Free public transport has been a topic of interest for many cities and countries worldwide, seen as a means to reduce traffic congestion, improve air quality, and provide affordable transportation to citizens. Many studies have been conducted on this topic. Kębłowski (2017) presents an analysis of the various models of Free Public Transport implemented globally. It categorizes these models into two categories, the FFPT and Partial FFPT. FFPT is defined as PT systems in which the free service is widely available and accessible to the majority of users throughout the majority of the time. On the other hand, Partial FFPT is a ticket-free system that is limited in some way, such as being restricted by space, time, or in terms of user groups. Guelton & Poinot (2020) focuses on fare-free public transportation in France and its increasing availability, as well as working financial models for this policy in the context of limited public funds. According to Duhamel (2004), the FFPT can be considered as a viable option only for cities with low ticketing revenue and passenger volumes prior to implementation.

Given the complexity of impacts that this policy can have in different contexts, the academic community is divided in their opinion. Studies by Studenmund & Connor (1982) and Brown et al. (2003) support the notion that ticket-free systems can attract car users to public transportation, even after fares are reinstated, settling the obtained shift in the modal split towards Public Transport in favor of cars. Bull et al. (2021) proves with their study the effectiveness of this policy in increasing the modal share of PT in Santiago, Chile. Conversely, Volinski (2012) highlights FFPT's limited achievements in generating a modal shift from private vehicles to PT in terms of cost and resources to be sustained by the companies. The study of Cats et al. (2017) highlights the limited impact of FFPT on sustainable development and suggests that it may cause financial instability in public transportation systems, resulting in irrational travel patterns and unproductive mobility. In the case of Luxembourg, this could result in PT becoming a competitor with walking as a mode of transportation for short trips of less than 500 meters.

Nonetheless, several cities have implemented free public transport policies, with varying degrees of success. Many European municipalities justify FFPT as a strategy working towards reducing car usage (e.g. Avesta, Sweden) and car-related pollution and noise (e.g. Livigno, Italy) and thereby increasing the liveability of the city itself. A study from the city of Tallinn provides an empirical evaluation of the impact of FFPT on service performance, passenger demand, and accessibility, showing that this policy led to an increase of 1.2% in passenger demand for PT, while highlighting that the relatively small impact could be attributed to the previous price level and public transport share, as well as the short-term impact, Cats et al. (2017). The important difference with the case of Luxembourg is that a ticketing system is still applied, where the residents don't pay for the PT whereas tourists and visitors have to pay standard fares. In a similar context, the french city of Dunkerque implemented a Partial FFPT policy in 2018, Briche & Huré (2018): they showed that this policy is sustainable and synergizes with urban development while also being mindful of the needs of low-income populations. In this study, free transportation is proven technically and financially feasible for an urban area of 200,000 residents, debunking the hypothesis that it is only applicable to medium-sized cities.

2 METHODOLOGY

Survey Data

In this work, the data used to generate the agents' population is extracted from the LuxMobil travel survey, which was conducted in 2017, whose aim was to draw a picture of mobility throughout the country of Luxembourg, and contained information of both residents as well as cross-border workers. In the survey, a total of 33,207 individuals were asked to provide their socio-demographic

information and a travel diary for a typical day, with detailed trip chain information such as main modes of transport for each trip, travel time, origin, and destination. The respondents were 66% from Luxembourg, 15% from France, 10% from Germany, and 9% from Belgium. This study included 218 zones, with 71 of those being from Belgium, France, and Germany.

Data cleaning

After excluding all individuals who declared not to perform any activities, the population sample size was reduced to 22,199 respondents (67% of all interviewees) without altering substantially the distribution of the population in terms of resident/crossborderers. Considering the total population moving in the area of the sample (around 197.000 cross borders, together with 645000 in Luxembourg Residents)¹, the MATSim synthetic population represents approximately 2% of the population, with agents carrying out activities of various types such as home, work, school, shopping, leisure, and others.

From the travel survey, we extracted 85506 trips, each with an associated main mode of transport, and estimated the modal split for a weekday, which is divided as follows: 79% Car, 16% PT, 1% bicycle legs, and 4% walk legs. Demand generation and activity location assignment were based on 24-hour activity OD (Origin-Destination) matrices generated from the travel survey data. The source data for the MATSim Luxembourg road and PT network was obtained from OSM and GTFS files using the pt2matsim tool Poletti (2016). To simplify the analysis, the 216 areas were merged into 146 zones, with 3 of these zones representing the three neighboring countries. This was due to the limited population density data available for some of the foreign areas in the travel survey and the focus on the Luxembourgish road and PT network. To try to maintain as much accuracy as possible on the data regarding cross-border commuters, a "bee-line" moving approach was utilized, which involved relocating their house location to the nearest centroid and adjusting their first and last activity times based on their declared mode of transport and average speed. This was calculated using the beeline distance between the declared position of the house and the closest centroid. Moreover, regarding the cross-border commuters, since there were no interviews performed on people under the age of 18 in the survey, we assumed that they would all hold valid driving licenses. Car availability was then determined based on the area and household aggregation. While this approach could alter travel behavior if coupled with a strong mode replanning strategy, it ultimately proved to be an effective solution as the data remained consistent.

Calibration of the MATSim simulation

In order to calibrate the scenario prior to the free PT policy (Pay-PT), we run different MATSim simulations with 10% of the overall population (around 64000 agents) to reach a steady state. This was achieved by performing multiple simulations, each with 250 iterations, starting from the parameters used for the Berlin scenario, Ziemke et al. (2019). The scoring parameters were adjusted after each simulation until convergence was achieved with the traffic count data, as well as with the modal split, average travel time, and average distance per mode per leg from the Luxmobil travel survey. The volume on the main congested roads that are leading to the foreign centroids was then compared to the real traffic counts, showing adherence with the initial data. In order to reproduce the FFTP system, the monetary constant was set to 0. Nonetheless, the calibration process for the FFTP Scenario turned out to be a challenging task. The biggest obstacle encountered was the absence of a ticketing system for every PT, making it difficult especially for busses to determine the actual number of passengers, paired with the presence of multiple PT providers. However, for trains and trams, data is available regarding the passenger-km performed², which was then matched with the corresponding data from the simulation.

¹<https://luxembourg.public.lu/en/society-and-culture/population/demographics.html>

²<https://gouvernement.lu/fr/publications/rapport-activite/minist-mobilite-travaux-publics/departement-mobilite-transport/2021-rapport-activite-dmt.html>

3 RESULTS AND DISCUSSION

Table 1 and Table 2 present the results of the effect of the Free PT policy, segmented between Luxembourg Residents and Crossborderers, and the overall comparison between the Pay-PT scenario and the Free-PT scenario. The asterisk in the Tables (*) indicates that the sample size for these modes of transport was too small, leading to unrealistic results in the simulation.

	Comparison - Paid PT vs Free PT				
	AVG Leg TT	AVG Leg Distance	PKT	PHT	Modal Split
	Resident				
Car	-5%	-4%	-12%	-13%	*
Walk	3%	-2%	*	*	*
PT	-11%	-9%	8%	5%	*
Bike	18%	18%	*	*	*
	Crossborderers				
Car	-1%	-2%	2%	3%	*
Walk	17%	6%	*	*	*
PT	-13%	-12%	61%	60%	*
Bike	7%	7%	*	*	*
	Overall				
Car	-2%	-1%	-8%	-8%	-10%
Walk	9%	2%	*	*	0%
PT	-12%	-9%	19%	16%	10%
Bike	12%	12%	*	*	0%

Table 1: Analysed KPIs - Pay vs Free PT

Table 1 compares the difference between the Pay-PT and Free-PT scenarios in terms of average Leg travel time and distance, person-kilometers traveled (PKT), and person-hours traveled (PHT). The policy appears to impact all transportation modes: for Residents, the use of public transportation increased, as seen as an increase in the PKT and PHT, with a corresponding decrease in leg travel time and distance. The policy had a greater impact on the Crossborderers, resulting in a 60% rise in PKT and PHT, demonstrating that public transportation became a more appealing option for their daily journeys. It has to be noted that these numbers are likely overestimated, as the FFPT policy is available only for PT trips performed *within* Luxembourg, not for trips originating from the Greater Region. Nevertheless, these results show that extending the FFPT policy to cover cross-border trips from external countries to Luxembourg could have a substantial impact, given that a large proportion of commuters travel long distances. Overall, the Free PT affected the agent's travel behavior, as seen in an increase in PKT and PHT for PKT, together with a Modal Split change that presents a 10% of car users switching to public transportation.

Table 2 presents information on the changes in the trips' mode of transportation. For trips performed by Residents, PT was favored over cars for all the considered segments, which led to a decrease in car usage. On the other hand, the free-PT policy had a noteworthy impact on cross-border in terms of a reduction of trips performed by cars in all the considered segments. Nonetheless, considering cross-border commuters were the target group for the policy as they travel on average more than 25 km to their main activity (an average of 35.3 km as per our travel survey), the impact is very limited. The reluctance of these commuters to switch modes of transportation is shown by a slight increase in the PT choice for all the segments, and especially the limited impact is highlighted by the slight PT trips increase in the > 25 km trip segment, which accounts for the largest portion of trips. This is also linked to the high PKT and PHT for PT for Crossborderers presented in Table 1, suggesting that cross-border travelers could find PT to be convenient, and willing to travel further to reach their destination. This could be due to several reasons, including the lack of effective connections to foreign countries and the high car dependency of the country with 676 passenger cars for every 1000 inhabitants and being the country with one of the highest mobility expenditures in Europe.

% of trip difference for specific distance segments - Paid PT vs Free PT				
Resident				
trip <5 km			5 km <trip <10km	
	<i>trips (Paid PT)</i>	<i>Paid PT vs Free PT</i>	<i>trips (Paid PT)</i>	<i>Paid PT vs Free PT</i>
Car	30449	-4%	16361	-2%
Walk	10835	1%	90	0%
PT	520	1%	335	1%
Bike	448	2%	56	1%
10 km <trip <25 km			trip >25 km	
Car	21670	-2%	27761	-3%
Walk	60	0%	15	0%
PT	904	1%	3019	1%
Bike	68	1%	120	2%
Crossborders				
trip <5 km			5 km <trip <10km	
	<i>trips (Paid PT)</i>	<i>Paid PT vs Free PT</i>	<i>trips (Paid PT)</i>	<i>Paid PT vs Free PT</i>
Car	678	-4%	988	-10%
Walk	934	2%	68	3%
PT	30	2%	113	7%
Bike	19	0%	45	1%
10 km <trip <25 km			trip >25 km	
Car	3787	-5%	22694	-5%
Walk	79	0%	41	0%
PT	60	1%	734	2%
Bike	64	4%	248	3%
Overall				
trip <5 km			5 km <trip <10km	
	<i>trips (Paid PT)</i>	<i>Paid PT vs Free PT</i>	<i>trips (Paid PT)</i>	<i>Paid PT vs Free PT</i>
Car	31127	-4%	17348	-3%
Walk	11768	1%	158	0%
PT	550	1%	448	2%
Bike	467	2%	102	1%
10 km <trip <25 km			trip >25 km	
Car	25457	-3%	50455	-4%
Walk	139	0%	56	0%
PT	964	1%	3753	2%
Bike	132	1%	369	2%

Table 2: % of trip difference within specific trip segments - Paid PT vs Free PT

4 CONCLUSIONS

This paper analyses the impact of Free Public Transport in the context of Luxembourg, based on a MATSim scenario calibrated through survey data. The results demonstrate the effects of this policy on the country, where benefits can be seen in the increase of the overall choice for PT as a main mode of transport, seen as an increase in the PKT and PHT, together with a shift in the modal split. However, the policy slightly impacted cross-border travelers, with a significant increase in their travel time if using public transport, and indeed resistance has been observed in the expected modal shift.

Future research includes the investigation of the impact of the 1 PT-stop trips and how these affect the whole system, given that before 2020 Luxembourg already included an integrated ticketing system so without a real time-gain on the ticket validation to be done in-vehicle, especially for busses; a deeper study on the cross-border behavior and which policies could be put in place to reduce the still very high car mode share in the country; the implementation of the different on-demand services which are already present in the country (such as bike and car-sharing, taxi, car-pooling etc,..) and a comparison with the obtained data for the PT scenario with the real data

that will be made available in 2023 by the national statistic agency STATEC.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Mobility and Public Works for sharing the Lux-mobil dataset.

REFERENCES

- Briche, H., & Huré, M. (2018). Dunkerque, nouveau « laboratoire » de la gratuité des transports.
- Brown, B. B., Werner, C. M., & Kim, N. (2003, dec). Personal and Contextual Factors Supporting the Switch to Transit Use: Evaluating a Natural Transit Intervention. *Analyses of Social Issues and Public Policy*.
- Bull, O., Muñoz, J. C., & Silva, H. E. (2021, jan). The impact of fare-free public transport on travel behavior: Evidence from a randomized controlled trial. *Regional Science and Urban Economics*, 86.
- Cats, O., Susilo, Y. O., & Reimal, T. (2017). The prospects of fare-free public transport: evidence from Tallinn. *Transportation*, 44(5), 1083–1104. doi: 10.1007/S11116-016-9695-5/FIGURES/6
- Duhamel, Y. (2004). *Gratuité des transports publics urbains et répartition modale, rapport pour le Predit* (Tech. Rep.). GO1, Paris ADEME.
- Guelton, S., & Poinot, P. (2020, jul). La gratuité des réseaux de transports collectifs urbains : un modèle de financement particulier ? *Transports urbains*, 136(1), 12–17. doi: 10.3917/TURB.136.0012
- Horni, A., Nagel, K., & Axhausen, K. (Eds.). (2016). *Multi-agent transport simulation matsim*. London: Ubiquity Press. doi: 10.5334/baw
- Kębłowski, W. (2017). More than just riding without a ticket ? Exploring the geography of fare-free public transport . (November).
- Poletti, F. (2016). *Public Transit Mapping on Multi-Modal Networks in MATSim* . ETH Zurich.
- Studenmund, A., & Connor, D. (1982, jul). The Free-Fare Transit Experiments. *Transportation Research Part A: Policy and Practice*, 16A(4).
- Volinski, J. (2012). Implementation and Outcomes of Fare-Free Transit Systems. *Implementation and Outcomes of Fare-Free Transit Systems*.
- Ziemke, D., Kaddoura, I., & Nagel, K. (2019). The matsim open berlin scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. *Procedia Computer Science*, 151, 870-877. doi: 10.1016/j.procs.2019.04.120

Influence of station characteristics, urban surroundings and perceived safety on satisfaction and public transport ridership

Başaran, Gülin Göksu*^{1,2}, Ingvardson, Jesper Bláfoss³, and Nielsen, Otto Anker⁴

¹PhD Student, Department of Management, Technical University of Denmark, Denmark

²Copenhagen Metro, Denmark

³Associate Professor, Department of Management, Technical University of Denmark, Denmark

⁴Professor, Department of Management, Technical University of Denmark, Denmark

SHORT SUMMARY

Public transport (PT) is essential to fulfill travel needs in urban areas. Predictors of PT ridership and satisfaction provide a good understanding of how new users can be attracted and existing users can be retained. Among these predictors, perceived safety and built environment (BE) attributes of stations and their surroundings still require further research. Using data from a tailor-made survey on train trips in East Denmark (1,004 respondents), we investigate the relationship between perceived safety, satisfaction and PT ridership and highlight the influence of BE attributes. Based on a structural equation model, we find a significant relationship between (i) perceived safety and satisfaction with trip-ends, and (ii) satisfaction with trip-ends and overall trip satisfaction. Lighting, maintenance and wayfinding are some of the essential attributes of stations, and their surroundings should not have isolated areas. No significant effect was found for trip satisfaction on PT ridership, but further research will consider this relationship.

Keywords: Public transport; Perceived safety; Satisfaction; Station design; Structural equation modelling

1 INTRODUCTION

Public transport (PT) is an important travel mode in urban areas, serving a heterogeneous set of users and contributing to more sustainable transport and social equity in cities. While attracting new PT users requires a good understanding of predictors of PT ridership, ensuring existing users' satisfaction is crucial to retain their ridership (van Lierop et al., 2018). Both PT ridership and satisfaction are linked to the built environment (BE) attributes, including those of station environments and their surroundings (Iseki et al., 2007; Susilo & Cats, 2014; Taylor & Fink, 2013). In addition, perceived safety is a crucial factor in predicting both ridership and satisfaction with the overall trip and can be enhanced by improving the stations and their surroundings as well (Ingvardson & Nielsen, 2021; Iseki et al., 2007; Susilo & Cats, 2014). However, further research is necessary to investigate which of these BE attributes have the highest importance for satisfaction and ridership.

Satisfaction with PT is influenced by a high number of factors, among which on-board cleanliness, comfort and staff behaviour are the most common ones according to a recent review (van Lierop et al., 2018). While these are trip attributes, station attributes such as wayfinding (Nielsen et al., 2021), real-time information (Chowdhury & Ceder, 2013), maintenance and cleanliness (Eboli et al., 2018) also improve users' experience. There are, however, few studies on PT satisfaction which incorporate urban characteristics.

Perceived safety is one of the most crucial needs which has to be fulfilled for a PT trip to take place, as otherwise, one might have to alter the time or mode of the trip, or cancel the trip completely (Loukaitou-Sideris et al., 2009; Lubitow et al., 2017). Focusing on lighting, maintenance, real-time information and staff presence can help achieve safe station environments which can encourage PT users (Cozens et al., 2003; Rahaman et al., 2016). That said, urban design around stations is at least as important as at the station. Providing trees (Basu et al., 2022), good lighting and human activity (Iseki et al., 2007) enhance perceived safety around stations while the presence of isolated areas or unused parking lots negatively affect it (Iseki et al., 2007).

This study analyzes the influence of station characteristics, urban surroundings and perceived safety on satisfaction as well as PT ridership, focusing on train travel in East Denmark. Using a

tailor-made online survey incorporating a detailed list of built environment attributes, we estimate a comprehensive structural equation model (SEM). First, we examine which attributes of train stations and their surroundings improve perceived safety and satisfaction. We do this separately at the home and activity ends of the trip to consider explicitly potential differences in user perceptions. Second, we investigate whether overall trip satisfaction increases with higher levels of satisfaction with individual attributes at both trip ends. Lastly, we explore the relationship between trip satisfaction and PT ridership. The SEM framework allows for analysing these relationships in detail through both the direct and indirect effects. Our final data set comprises 1,004 train trips made by a large sample of PT users in East Denmark in June-July 2022.

2 METHODOLOGY

Survey design and data collection

We designed an online survey in Danish, with 35-40 questions in three parts: (i) travel patterns and preferences towards the attributes of stations and their surroundings, (ii) details of respondents' latest train trip, and (iii) background questions.

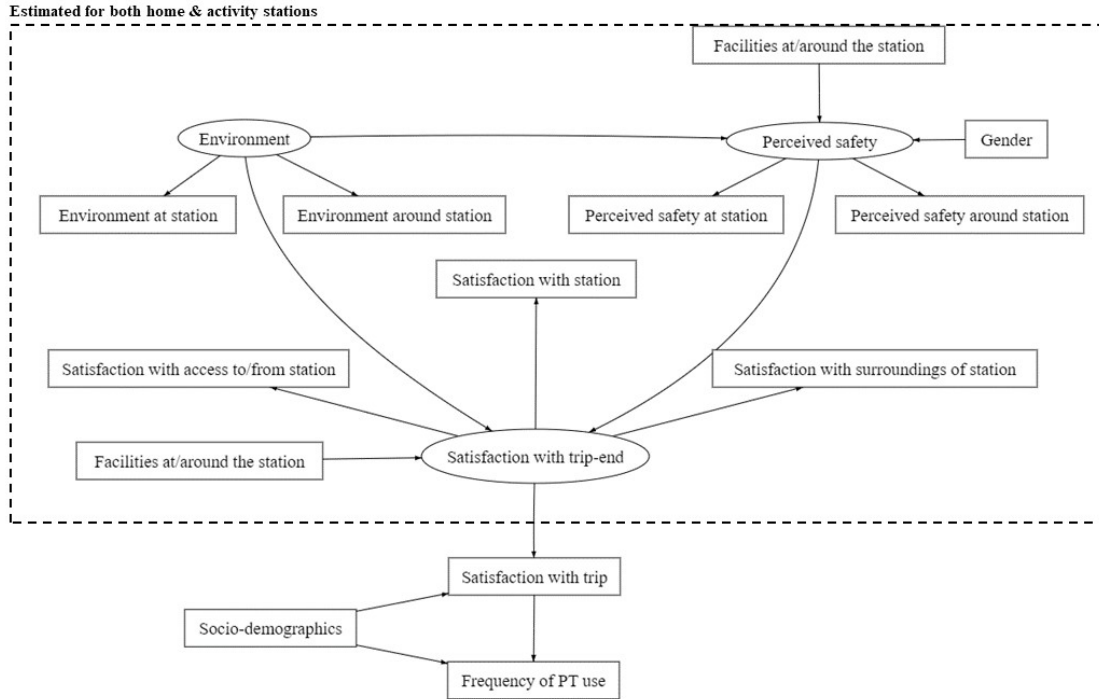
In the first part, we included station facilities (e.g. escalators, information screens, wayfinding), station surroundings (e.g. human activity, large parking lots), environment at/around stations (noise, air quality, lighting, maintenance, cleanliness), and access paths to stations (e.g. tunnels, pedestrian streets). We measured the importance of the selected attributes on a 5-point Likert scale (1: Very unimportant, 5: Very important). In the second part, we collected which of these attributes were present on the specific trip, and measured respondents' perceived safety level. To provide a concrete scenario, we asked the respondents to rank how safe they would feel after dark at/around their stations on a 5-point Likert scale. We also measured satisfaction with: (a) station, (b) station surroundings, (c) access to/from the station, and (d) the entire trip, all measured on a 5-point Likert scale. Items (a) to (c) were repeated for start and end stations. In the third part, we asked about gender, age, education, and access to transport resources.

We distributed the online survey through a panel of PT users from the Danish consumer watchdog for public transport (Passenger Pulse). We targeted PT users over 18-years old who reside in East Denmark, an area which includes the city of Copenhagen and has more than 2 million residents. Collecting data in June-July 2022, we reached 1,314 complete responses in the final data set.

Analysis method

Figure 1 shows our framework which covers the relationship between perceived safety, satisfaction variables and frequency of PT use, while also exploring the effects of station characteristics and urban surroundings. Socio-demographic variables are included to account for different user groups' needs and experiences. We followed a structural equation modelling approach, as it allows for simultaneous consideration of the dependent variables in a consistent manner. As we expected a difference in individuals' preferences at home and activity ends of the trip, we created latent variables and estimated models for both trip ends in the same SEM model.

Figure 1: Structural equation modelling framework where circles represent latent variables.



In the SEM model, we introduced variables in 5-point scales as continuous variables in the model, and employed dummy variables for the yes/no questions or categorical variables (e.g. age, education).

In the measurement part of the SEM model, we applied confirmatory factor analysis (CFA) to create latent variables from items which were measured both at and around the stations. The five environment variables and perceived safety underwent this procedure separately at home and activity ends. We also loaded the three satisfaction items (i.e. satisfaction with access conditions, satisfaction with station surroundings, satisfaction with the station) into a single latent variable at each trip end.

In the structural part, we estimated models explaining these latent variables at both trip ends as well as satisfaction with the trip and frequency of PT use to test several hypotheses. First, we expected that maintenance and lighting, along with BE attributes and gender, significantly affected perceived safety at both trip ends. Second, we expected a positive significant relationship between perceived safety and satisfaction with trip ends, in addition to the significant effect of attributes such as wayfinding and crowdedness. Third, we hypothesised that satisfaction with trip ends would significantly contribute to the satisfaction with the overall trip and that there would be significant differences based on socio-demographics. Lastly, we expected a significant positive effect of satisfaction with the overall trip on frequency of PT use, in addition to socio-demographics.

3 SAMPLE STATISTICS

To test whether respondents have different preferences at home and activity ends of their trips, we converted start & end stations into home-end & activity-end stations depending on whether the trip started or ended at home. After removing trips with missing values, our modelling data set includes 1,004 trips. These 1,004 trips cover 180 stations out of the 297 in the region, and the largest transport hubs are represented. The remainder of the stations are mostly smaller local train stations with few daily users.

Table 1 describes the full sample (N= 1,314), and the sample used in the SEM model (N=1,004). In addition, we show the sample of PT users in the Danish National Travel Survey (Transportvaneundersøgelsen) between 2018-2022 for comparison, as this survey is representative for Denmark (Christiansen & Skougaard, 2015). In all three samples, women are slightly overrepresented. Given their low percentage, we merged "nonbinary" and "other" categories with women in the rest of the analysis. Both of our samples stand out with their age distribution from the TU sample, with a higher share of individuals over 50-years-old. This is a result of the age bias in the Passenger

Pulse panel, which has an average age of 61. Furthermore, our sample is more educated and has higher income levels than the TU sample. In terms of access to transport resources, the samples resemble each other.

Table 1: Sample description, compared to the description of PT users in the Danish NTS

Variable	Full sample	SEM sample	Danish NTS PT users
Gender			
Female	55.1%	54.7%	53.5%
Male	43.8%	45.0%	46.5%
Nonbinary	0.1%	0.1%	-
Prefer not to say	1.1%	-	-
Other	0.2%	0.2%	-
Age			
18-29	2.4%	2.6%	39.7%
30-39	5.0%	4.8%	19.9%
40-49	9.8%	10.0%	12.5%
50-59	21.9%	21.6%	12.2%
60-69	29.3%	29.6%	7.9%
70-79	26.7%	26.7%	5.9%
>80	4.9%	4.8%	1.9%
Education			
Primary school	2.7%	3.2%	14.7%
High school	4.6%	4.8%	17.6%
Vocational	13.4%	13.8%	9.8%
Short-term higher education (1.5-2 years)	7.2%	6.8%	4.6%
Medium-term higher education (2-5 years)	36.2%	36.6%	27.1%
Long-term higher education (5+ years)	35.8%	34.9%	26.2%
Income			
0-99.999 DKK	1.9%	1.8%	9.2%
100.000-199.999 DKK	7.3%	7.5%	9.8%
200.000-299.999 DKK	16.8%	17.6%	9.3%
300.000-399.999 DKK	18.3%	18.9%	12.9%
400.000-499.999 DKK	15.4%	15.4%	8.9%
More than 500.000 DKK	25.0%	25.4%	11.7%
NA	15.2%	13.3%	38.2%
Car availability			
Yes	55.7%	56.0%	47.7%
No	44.3%	44.0%	52.3%
Bicycle availability			
Yes	73.8%	84.8%	76.4%
No	26.2%	15.2%	23.6%
Driving licence			
Yes	84.3%	74.6%	70.8%
No	15.7%	25.4%	29.2%
No. of obs	1314	1004	1850

We describe the frequency of using different transport modes in table 2. 55.6% of the SEM sample travels at least 3-4 times a week with PT. Similarly, more than half of the respondents walk or cycle frequently, while the share of car drivers and passengers are quite low. E-scooters, shared bicycles and shared cars are almost never used.

Table 2: Frequency of travelling with different modes in the SEM sample (N=1004)

Transport mode	Never	Less than once a month	1-3 times a month	1-2 times a week	3-4 times a week	5 times or more a week
Public transport	0.5%	4.6%	18.3%	21.0%	24.6%	31.0%
Car driver	45.6%	9.7%	10.0%	16.0%	10.6%	8.2%
Car passenger	17.2%	33.0%	26.1%	18.2%	3.8%	1.7%
Walking	2.3%	2.4%	4.6%	9.7%	14.6%	66.4%
Own bicycle	22.3%	9.5%	7.7%	9.1%	15.4%	36.1%
E-scooter	98.0%	1.6%	0.3%	-	-	0.1%
Shared bicycle	98.3%	1.4%	0.2%	0.1%	-	-
Shared car	93.8%	3.9%	1.2%	0.8%	0.2%	0.1%

Table 3 describes the environment at and around stations at both trip ends. Noise levels are perceived higher at the activity end, which is expected given that stations with high activity levels are mostly located in the centre of Copenhagen while quieter suburbs are represented more in the home stations. Air quality shows a similar pattern with a higher share of respondents reporting bad conditions at activity ends. At least 40 % of respondents state good lighting conditions in all cases. Maintenance and cleanliness distributions are similar in all four cases, achieving high rankings from approximately 30 % of the sample.

Table 3: Description of environment variables in the SEM sample (N=1,004) (^a The noise variable is reverse coded where 1: very quiet, 5: very loud. For all other variables, 1: very bad, 5: very good.)

Environment	1	2	3	4	5	Average	Std.Dev
At home station							
Noise ^a	8.6%	33.8%	45.4%	11.2%	1.1%	2.62	0.83
Air quality	1.6%	7.5%	43.8%	34.3%	12.9%	3.49	0.87
Lighting	1.5%	10.2%	38.8%	42.4%	7.1%	3.43	0.82
Maintenance	5.5%	23.4%	37.3%	30.0%	3.9%	3.03	0.95
Cleanliness	4.2%	24.4%	39.8%	27.3%	4.3%	3.03	0.92
Around home station							
Noise ^a	6.3%	32.5%	40.6%	18.6%	2.0%	2.78	0.89
Air quality	1.6%	10.6%	44.5%	32.6%	10.8%	3.40	0.87
Lighting	1.2%	11.6%	41.3%	41.6%	4.3%	3.36	0.79
Maintenance	3.3%	19.5%	44.7%	29.7%	2.8%	3.09	0.85
Cleanliness	3.5%	22.2%	45.7%	25.6%	3.0%	3.02	0.86
At activity station							
Noise ^a	5.1%	17.3%	44.5%	27.2%	5.9%	3.11	0.93
Air quality	3.6%	19.8%	46.2%	25.2%	5.2%	3.09	0.89
Lighting	1.3%	10.8%	42.8%	40.4%	4.7%	3.36	0.79
Maintenance	4.9%	15.4%	44.5%	30.5%	4.7%	3.15	0.91
Cleanliness	5.7%	19.2%	42.7%	28.8%	3.6%	3.05	0.92
Around activity station							
Noise ^a	3.8%	14.4%	38.8%	35.9%	7.2%	3.28	0.93
Air quality	3.5%	22.9%	48.2%	21.1%	4.3%	3.00	0.87
Lighting	0.9%	9.7%	46.8%	38.8%	3.8%	3.35	0.74
Maintenance	3.3%	15.5%	50.0%	27.3%	3.9%	3.13	0.84
Cleanliness	5.2%	21.9%	46.1%	24.2%	2.6%	2.97	0.88

The overall perceived safety levels at and around stations are high, as almost 50 % of the sample rank perceived safety 4 or higher (Figure 4). The distribution of scores at and around each trip end are quite similar.

Table 4: Description of perceived safety in the SEM sample (N=1004)

Perceived safety	1 - Very unsafe	2	3	4	5 - Very safe	Average	Std. dev.
Around home station	4.6%	13.7%	24.9%	43.2%	13.7%	3.48	1.04
At home station	5.0%	13.8%	26.0%	42.4%	12.9%	3.44	1.04
Around activity station	2.7%	12.1%	35.9%	39.7%	9.7%	3.42	0.92
At activity station	2.9%	12.8%	33.9%	41.5%	9.0%	3.41	0.92

As table 5 shows, the sample has quite high satisfaction levels, the highest being trip satisfaction with 74.4% respondents stating that they are satisfied or very satisfied, leading to an average score of 3.85/5.

Table 5: Description of satisfaction variables in the SEM sample (N=1004)

Satisfaction with:	1 - Very unsatisfied	2	3	4	5 - Very satisfied	Average	Std. dev.
Access to/from the home station	2.8%	7.6%	18.7%	49.1%	21.8%	3.80	0.96
Surroundings of home station	3.8%	12.3%	27.0%	42.3%	14.6%	3.52	1.01
Home station	3.4%	8.3%	35.0%	41.7%	11.7%	3.63	0.96
Access to/from the activity station	2.7%	11.3%	21.4%	49.5%	15.1%	3.76	0.91
Surroundings of activity station	2.1%	7.1%	22.2%	50.2%	18.4%	3.50	0.92
Activity station	2.5%	8.8%	25.7%	51.5%	11.6%	3.61	0.89
Trip	3.3%	3.7%	18.6%	53.8%	20.6%	3.85	0.90

Lastly, tables 6 and 7 describe the facilities at/around stations from the SEM model. As the activity stations are at more central locations, they have a lower share of large parking lots, closed facades and isolated areas, while the share of urban life is also higher. In most cases (86%), respondents did not experience problems seeing information screens at stations. A similar percentage of respondents report good wayfinding at stations while crowdedness appears to be a bigger problem at the activity-end.

Table 6: Description of attributes at and around home/activity stations in the SEM sample (N=1004)

Attribute	Home-end		Activity-end	
	No	Yes	No	Yes
Shops around	55.8%	44.2%	50.8%	49.2%
Urban life around	58.6%	41.4%	44.0%	56.0%
Large parking lots around	67.3%	32.7%	81.7%	18.3%
Closed facades around	81.2%	18.8%	84.2%	15.8%
Isolated areas around	76.7%	23.3%	85.4%	14.6%
Trees around	55.5%	44.5%	70.7%	29.3%
Problem seeing information screens	86.5%	13.6%	86.8%	13.3%
Respondent used elevator	83.6%	16.4%	84.9%	15.1%
Respondent used escalator	89.2%	10.8%	78.1%	21.9%
Respondent used stairs	31.7%	68.3%	35.0%	65.0%
Access via pedestrian street	64.2%	35.8%	57.4%	42.6%
Access via tunnel	72.2%	27.8%	74.2%	25.8%
Access via bridge	86.5%	13.6%	87.3%	12.8%
Access via bike path	70.5%	29.5%	78.9%	21.1%

Table 7: Description of attributes at home/activity stations (N=1004) (^b: The crowdedness variable is reverse coded where 1: very crowded, 5: not crowded at all)

Attribute	1	2	3	4	5
Home station					
Wayfinding	1.0%	3.4%	6.3%	27.0%	62.4%
Crowdedness ^b	4.8%	26.6%	24.4%	32.6%	11.7%
Activity station					
Wayfinding	0.9%	5.6%	7.2%	31.0%	55.4%
Crowdedness ^b	12.5%	29.4%	24.6%	24.0%	9.6%

4 RESULTS AND DISCUSSION

To test our hypotheses, we estimated a SEM model. When designing the model, we looked at the correlation between the variables described so far, and found that most of the attributes measured both at and around the station were highly correlated. For example, two variables measuring perceived safety at/around the home stations had a statistically significant correlation coefficient of 0.86. This was also the case for the three satisfaction variables (at, around, access) and environment variables. This informed the decision to create latent variables. Maintenance and cleanliness variables were also highly correlated in all cases (>0.70). Therefore, we defined residual covariances between these variables in the SEM model. The fit indices show acceptable model fit with RMSEA and SRMR both below 0.08. However, CFI is slightly below the required level of 0.90 (Table 8).

Table 8: Goodness-of-fit measures of the SEM model

Num of. obs	χ^2	DoF	P-value	Comparative Fit Index	RMSEA	SRMR
1004	4901.87	1712	0.00	0.862	0.043	0.064

We present the measurement part of our SEM model in table 9, and the structural part in tables 10 to 12. All indicators in the CFA have loaded into their corresponding latent variables with acceptable loading, and the Cronbach's alpha values were all above 0.70 thus suggesting good internal consistency (Miller, 1995). We introduced these variables as explanatory, and in some cases, also dependent variables in the structural part.

Table 9: Measurement model (N=1,004)

Latent variable (Cronbach's alpha)	Indicator	Std. Coef.	Std.Err	P-value	Sig.
Satisfaction_home (0.84)	Satisfaction with the home station		1	0.83	
	Satisfaction with access to/from the home station	0.72	0.04	0.00	***
Noise_home (0.80)	Satisfaction with surroundings of home station	0.75	0.04	0.00	***
	Noise at home		1.00	0.64	
Air_home (0.85)	Noise around home	0.87	0.06	0.00	***
	Air quality at home		1.00	0.74	
Lighthing_home (0.82)	Air quality around home	0.88	0.04	0.00	***
	Lighting at home		1.00	0.69	
Maintenance_home (0.80)	Lighting around home	0.81	0.04	0.00	***
	Maintenance at home		1.00	0.79	
Cleanliness_home (0.83)	Maintenance around home	0.79	0.03	0.00	***
	Cleanliness at home		1.00	0.82	
Safety_home (0.93)	Cleanliness around home	0.81	0.03	0.00	***
	Perceived safety at home station		1.00	0.99	
Satisfaction_activity (0.83)	Perceived safety around home station	0.89	0.03	0.00	***
	Satisfaction with the activity station		1.00	0.75	
Noise_activity (0.84)	Satisfaction with access to/from the activity station	0.67	0.04	0.00	***
	Satisfaction with surroundings of activity station	0.67	0.04	0.00	***
Air_activity (0.86)	Noise at activity		1.00	0.76	
	Noise around activity	0.88	0.04	0.00	***
Lighthing_activity (0.78)	Air quality at activity		1.00	0.77	
	Air quality around activity	0.87	0.03	0.00	***
Maintenance_activity (0.83)	Lighting at activity		1.00	0.65	
	Lighting around activity	0.77	0.04	0.00	***
Cleanliness_activity (0.84)	Maintenance at activity		1.00	0.76	
	Maintenance around activity	0.83	0.03	0.00	***
Safety_activity (0.92)	Cleanliness at activity		1.00	0.78	
	Cleanliness around activity	0.86	0.03	0.00	***
	Perceived safety at activity station		1.00	0.85	
	Perceived safety around activity station	0.91	0.04	0.00	***

***: $p \leq 0.001$, **: $0.001 < p \leq 0.01$, *: $0.01 < p \leq 0.05$, . : $0.05 < p \leq 0.1$

The models explaining the latent perceived safety variables at home and activity ends identify isolated areas around stations, lighting conditions and gender as common significant predictors in both cases (Table 10). While increasing levels of isolation reduce perceived safety, good lighting conditions improve the experience. Men feel significantly safer than non-male respondents. At home-end, urban life and trees also have positive and significant parameter estimates.

Table 10: Structural model - perceived safety (N=1,004)

Dependent variable	Explanatory variable	Std. Coef.	Std.Err	P-value	Sig.
Safety_home	Shops around home station	0.03	0.06	0.38	
	Urban life around home station	0.07	0.06	0.04	*
	Large parking lots around home station	-0.03	0.06	0.29	
	Closed facades around home station	-0.03	0.07	0.37	
	Isolated areas around home station	-0.11	0.07	0.00	***
	Trees around home station	0.06	0.06	0.05	.
	Lighting_home	0.46	0.07	0.00	***
	Maintenance_home	0.05	0.06	0.31	
	Male (Ref: Female and other)	0.14	0.06	0.00	***
Safety_activity	Shops around activity station	0.04	0.05	0.20	
	Urban life around activity station	0.03	0.06	0.32	
	Large parking lots around activity station	-0.04	0.07	0.26	
	Closed facades around activity station	-0.01	0.07	0.76	
	Isolated areas around activity station	-0.13	0.08	0.00	***
	Trees around activity station	0.02	0.06	0.49	
	Lighting_activity	0.30	0.09	0.00	***
	Maintenance_activity	0.11	0.08	0.13	
	Male (Ref: Female and other)	0.13	0.05	0.00	***

***: $p \leq 0.001$, **: $0.001 < p \leq 0.01$, *: $0.01 < p \leq 0.05$, . : $0.05 < p \leq 0.1$

For satisfaction with trip ends, maintenance, perceived safety, wayfinding and problems with seeing screens are significant at both ends (Table 11). Except for problems with seeing information screens, these common predictors significantly increase satisfaction with trip ends. Satisfaction with the

home station significantly reduces if the respondent has used escalators or stairs. This could imply a dispreference towards level changes or an unpleasant experience due to dirty elevators or long stairs.

Table 11: Structural model cont'd - satisfaction with stations (N=1,004)

Dependent variable	Explanatory variable	Std. Coef.	Std.Err	P-value	Sig.
Satisfaction_home	Noise_home	0.00	0.07	0.94	
	Air_home	0.01	0.06	0.93	
	Lighthing_home	0.02	0.07	0.71	
	Maintenance_home	0.53	0.09	0.00	***
	Cleanliness_home	0.04	0.07	0.53	
	Safety_home	0.24	0.03	0.00	***
	Access via pedestrian street	0.08	0.05	0.01	**
	Access via tunnel	0.02	0.05	0.54	
	Access via bridge	0.00	0.07	0.95	
	Access via bike path	0.11	0.05	0.00	***
	Wayfinding_home	0.15	0.03	0.00	***
	Problems seeing information screens	-0.11	0.06	0.00	***
	Used escalator	-0.06	0.07	0.04	*
	Used stairs	-0.08	0.06	0.01	*
	Used elevator	0.02	0.07	0.52	
	Crowdedness	-0.02	0.02	0.43	
	Satisfaction_activity	Noise_activity	-0.07	0.07	0.34
Air_activity		0.03	0.08	0.67	
Lighthing_activity		0.06	0.09	0.40	
Maintenance_activity		0.54	0.11	0.00	***
Cleanliness_activity		0.03	0.09	0.73	
Safety_activity		0.14	0.03	0.00	***
Access via pedestrian street		0.13	0.05	0.00	***
Access via tunnel		0.02	0.05	0.52	
Access via bridge		0.03	0.06	0.30	
Access via bike path		0.05	0.06	0.13	
Wayfinding_home		0.14	0.02	0.00	***
Problems seeing information screens		-0.09	0.06	0.00	***
Used escalator		0.06	0.06	0.08	.
Used stairs		-0.01	0.06	0.82	
Used elevator		0.03	0.07	0.43	
Crowdedness		0.04	0.02	0.16	

***: $p \leq 0.001$, **: $0.001 < p \leq 0.01$, *: $0.01 < p \leq 0.05$, . : $0.05 < p \leq 0.1$

As table 12 shows, satisfaction with the station conditions in both trip ends has a significant positive relationship with trip satisfaction. There is also a slight gender effect where men are less satisfied than non-male respondents. Respondents without a car are also slightly more likely to be satisfied with their trip. Education, which was added as a proxy for income, does not show a strong effect. Lastly, all age categories are less satisfied than respondents over 70-years-old, however only one of these categories is significant.

Unlike our expectation, satisfaction with the trip does not significantly influence frequency of PT use. However, age and not owning a car have strong effects.

Table 12: Structural model cont'd - trip satisfaction and frequency of PT use (N=1,004)

Dependent variable	Explanatory variable	Std. Coef.	Std.Err	P-value	Sig.
Satisfaction with the trip	Satisfaction_home	0.28	0.05	0.00	***
	Satisfaction_activity	0.32	0.05	0.00	***
	Male (Ref: Female and other)	-0.05	0.05	0.06	.
	No cars (Ref: Car owner)	0.05	0.05	0.07	.
	Education - Primary school (Ref: 2+ year higher ed.)	-0.03	0.14	0.34	
	Education - Highschool (Ref: 2+ year higher ed.)	0.05	0.12	0.11	
	Education - Vocational (Ref: 2+ year higher ed.)	0.07	0.07	0.01	*
	Education - Shrt higher education (Ref: 2+ year higher ed.)	-0.02	0.10	0.58	
	Age - 18-29 (Ref: 70+)	-0.06	0.16	0.04	*
	Age - 30-39 (Ref: 70+)	-0.06	0.12	0.06	.
	Age - 40-49 (Ref: 70+)	-0.05	0.09	0.12	
	Age - 50-59 (Ref: 70+)	-0.06	0.07	0.08	.
	Age - 60-69 (Ref: 70+)	-0.08	0.06	0.02	*
	Frequency of PT use	Satisfaction with the trip	-0.02	0.04	0.52
Male (Ref: Female and other)		0.00	0.08	0.95	
No cars (Ref: Car owner)		0.20	0.08	0.00	***
Education - Primary school (Ref: 2+ year higher ed.)		-0.04	0.22	0.17	
Education - Highschool (Ref: 2+ year higher ed.)		0.08	0.19	0.02	*
Education - Vocational (Ref: 2+ year higher ed.)		0.03	0.11	0.28	
Education - Shrt higher education (Ref: 2+ year higher ed.)		0.02	0.15	0.49	
Age - 18-29 (Ref: 70+)		0.01	0.25	0.77	
Age - 30-39 (Ref: 70+)		0.09	0.19	0.00	**
Age - 40-49 (Ref: 70+)		0.11	0.14	0.00	***
Age - 50-59 (Ref: 70+)		0.21	0.11	0.00	***
Age - 60-69 (Ref: 70+)		0.12	0.10	0.00	***

***: $p \leq 0.001$, **: $0.001 < p \leq 0.01$, *: $0.01 < p \leq 0.05$, . : $0.05 < p \leq 0.1$

Discussion

This study found a high correlation between attributes measured at and around the stations. The experience found on the way to a station might have a lasting influence on the experience at the station and vice versa, resulting in similar measurements. This should encourage planners to have a more holistic approach when designing stations and urban environments. Respondents might also have had difficulty differentiating between stations and their surroundings. However, as we explicitly mentioned the difference in each question, we believe this to be less likely.

The perceived safety analyses confirm gender effects in line with the literature. Among the hypothesised attributes, the presence of isolated areas and lighting were significant as expected while it is interesting that maintenance, for example, does not influence perceived safety.

It is initially unexpected that trip satisfaction does not significantly affect the frequency of PT use as, for example, Ingvardson & Nielsen (2019) found a positive significant relationship. One key difference between the two studies is that our sample consists heavily of captive PT users who never drive. Therefore, they might have to use PT even though they are dissatisfied as van Lierop, Badami & El-Geneidy also highlight in their review (2018). Susilo & Cats (2014) also state that choice users are often more satisfied than captive users.

As our structural model shows, home and activity ends have many significant parameters in common, although more factors are significant at the former. That said, many of the attributes at both ends, such as perceived safety, were correlated and this might have influenced model findings. One reason for this correlation can be the representation of the same stations with high passenger volumes in both ends. Another reason can be respondents' strong personal preferences. This will be considered explicitly in future work.

While the SEM results provide interesting insights, the model can and will be further developed. First, the SEM model lacks some important predictors of PT use such as service headway and trip duration in comparison to other alternatives. We will include such variables to get a clearer outcome and potentially improve model fit.

Second, the results are based on a relatively large sample of PT users, acting as a valuable source to understand actual users' needs and preferences. However, the age bias in the sample might have had an effect on the findings as younger individuals' preferences were under-represented. In March 2023, we will send the survey to a subset of respondents from the Danish NTS between 2018-2022 to achieve a larger and more representative sample. This way, we can also include more attributes such as bicycle parking which were left out due to having too many missing values. By including less frequent PT users in the data set, we can also test whether the relationship between satisfaction and ridership holds among different, and more representative, user groups.

5 CONCLUSIONS

This study analyzed the influence of station characteristics, urban surroundings and perceived safety on satisfaction as well as PT ridership, in East Denmark. Using 1,004 observations from our tailor-made survey, we employed a SEM model and first, we created latent variables from environment attributes which are highly correlated at and around stations as well as perceived safety and satisfaction variables. Second, we showed a significant relationship between (i) perceived safety and satisfaction with trip-ends, and (ii) satisfaction with trip-ends and overall trip satisfaction. While doing so, we identified which BE attributes should be present at stations and surroundings. Lastly, we could not demonstrate a significant effect of trip satisfaction on frequency of PT use. However, we expect that to be due to our sample's high share of captive PT users. As part of our future work, we will collect more data to achieve a more representative sample and incorporate more trip-related attributes to our SEM model to further investigate the relationship between the satisfaction constructs and PT ridership.

ACKNOWLEDGEMENTS

This study was conducted as a part of Seamless sustainable everyday urban mobility (EASIER), funded by ERA-NET Cofund Urban Accessibility and Connectivity (Grant no. 875022) and Innovation Fund Denmark (E.grant no. 0205-00002B). The authors would like to thank these parties for funding.

REFERENCES

- Basu, N., Oviedo-Trespalacios, O., King, M., Kamruzzaman, M., & Mazharul Haque, M. (2022). The influence of the built environment on pedestrians' perceptions of attractiveness, safety and security. *Transportation Research Part F: Psychology and Behaviour*, 87, 203–218. Retrieved 2022-05-02, from <https://doi.org/10.1016/j.trf.2022.03.006> doi: 10.1016/j.trf.2022.03.006
- Chowdhury, S., & Ceder, A. (2013). The Effect of Interchange Attributes on Public-Transport Users' Intention to Use Routes Involving Transfers. *Psychology and Behavioral Sciences*, 2(1), 5–13. Retrieved 2022-04-28, from <http://www.sciencepublishinggroup.com/j/pbs> (Publisher: Science Publishing Group ISBN: 20130201.12) doi: 10.11648/J.PBS.20130201.12
- Christiansen, H., & Skougaard, B. Z. (2015). *Documentation of the Danish National Travel Survey* (Tech. Rep.). Technical University of Denmark Department of Transport. (ISBN: 978-87-7327-288-6 ISSN: 1601-9458)
- Cozens, P., Neale, R., Whitaker, J., & Hillier, D. (2003). Managing crime and the fear of crime at railway stations-A case study in South Wales (UK). *International Journal of Transport Management*, 1, 121–132. (Publisher: Elsevier Ltd) doi: 10.1016/J.IJTM.2003.10.001
- Eboli, L., Forciniti, C., & Mazzulla, G. (2018). Spatial variation of the perceived transit service quality at rail stations. Retrieved 2022-04-28, from <https://doi.org/10.1016/j.tra.2018.01.032> doi: 10.1016/j.tra.2018.01.032
- Ingvardson, J. B., & Nielsen, O. A. (2019, August). The relationship between norms, satisfaction and public transport use: A comparison across six European cities using structural equation modelling. *Transportation Research Part A: Policy and Practice*, 126, 37–57. Retrieved 2023-03-02, from <https://www.sciencedirect.com/science/article/pii/S0965856417313046> doi: 10.1016/j.tra.2019.05.016
- Ingvardson, J. B., & Nielsen, O. A. (2021). The influence of vicinity to stations, station characteristics and perceived safety on public transport mode choice: a case study from Copenhagen. *Public Transport*, 1–22. Retrieved from <https://doi.org/10.1007/s12469-021-00285-x> doi: 10.1007/s12469-021-00285-x
- Iseki, H., Ringler, A., Taylor, B. D., Miller, M., & Smart, M. (2007). *Evaluating Transit Stops and Stations from the Perspective of Transit Users* (Tech. Rep.). California Department of Transportation.

- Loukaitou-Sideris, A., Bornstein, A., Fink, C., Samuels, L., & Gerami, S. (2009). *How To Ease Women's Fear of Transportation Environments: Case Studies and Best Practices* (Tech. Rep.). Mineta Transportation Institute.
- Lubitow, A., Rainer, J., & Bassett, S. (2017). Exclusion and vulnerability on public transit: experiences of transit dependent riders in Portland, Oregon. *Mobilities*. doi: 10.1080/17450101.2016.1253816
- Miller, M. B. (1995). Coefficient Alpha: A Basic Introduction From the Perspectives of Classical Test Theory and Structural Equation Modeling. *Structural Equation Modeling-a Multidisciplinary Journal*, 2(3), 255–273. Retrieved 2023-03-02, from <https://www.tandfonline.com/doi/abs/10.1080/10705519509540013> (Place: Abingdon Publisher: Routledge Journals, Taylor & Francis Ltd WOS:000208062100005) doi: 10.1080/10705519509540013
- Nielsen, O. A., Eltved, M., Anderson, M. K., & Prato, C. G. (2021). Relevance of detailed transfer attributes in large-scale multimodal route choice models for metropolitan public transport passengers. *Transportation Research Part A: Policy and Practice*, 147(February), 76–92. Retrieved from <https://doi.org/10.1016/j.tra.2021.02.010> (Publisher: Elsevier Ltd) doi: 10.1016/j.tra.2021.02.010
- Rahaman, M., Currie, G., & Muir, C. (2016, January). Development and Application of a Scale to Measure Station Design Quality for Personal Safety:. *Transportation Research Record: Journal of the Transportation Research Board*, 2540, 1–12. Retrieved 2021-08-11, from <https://journals.sagepub.com/doi/10.3141/2540-01> (Publisher: SAGE PublicationsSage CA: Los Angeles, CA) doi: 10.3141/2540-01
- Susilo, Y. O., & Cats, O. (2014, September). Exploring key determinants of travel satisfaction for multi-modal trips by different traveler groups. *Transportation Research Part A: Policy and Practice*, 67, 366–380. Retrieved 2023-01-09, from <https://linkinghub.elsevier.com/retrieve/pii/S0965856414001827> doi: 10.1016/j.tra.2014.08.002
- Taylor, B. D., & Fink, C. N. (2013). Explaining transit ridership: What has the evidence shown? *Transportation Letters*, 5(1), 15–26. Retrieved 2021-11-13, from <https://www.tandfonline.com/doi/abs/10.1179/1942786712Z.0000000003> (Publisher: Taylor & Francis) doi: 10.1179/1942786712Z.0000000003
- van Lierop, D., Badami, M. G., & El-Geneidy, A. M. (2018, January). What influences satisfaction and loyalty in public transport? A review of the literature. *Transport Reviews*, 38(1), 52–72. Retrieved 2023-01-04, from <https://doi.org/10.1080/01441647.2017.1298683> (Publisher: Routledge _eprint: <https://doi.org/10.1080/01441647.2017.1298683>) doi: 10.1080/01441647.2017.1298683

Assessing the Impacts of Public Transport-Based Crowdshipping: A Case Study in Nørrebro District in Copenhagen

Rong Cheng*¹, Andreas Fessler², Allan Larsen³, Otto Anker Nielsen⁴, Yu Jiang⁵

¹ PhD candidate, DTU Management, Technical University of Denmark, Denmark

² Advisor, Atkins, Denmark

³ Professor, DTU Management, Technical University of Denmark, Denmark

⁴ Professor, DTU Management, Technical University of Denmark, Denmark

⁵ Associate professor, DTU Management, Technical University of Denmark, Denmark

SHORT SUMMARY

This paper proposes a public transport-based crowdshipping concept as a complementary solution to the traditional parcel delivery system, where public transport users utilize their existing trips to carry out crowdsourced deliveries. To analyze the impact of public transport-based crowdshipping, we conduct a case study in Nørrebro district in Copenhagen using real-world data. Three scenarios with varying percentages of crowdshipped parcels are developed to be compared with the traditional distribution mode. For each scenario, the distribution of non-crowdshipped parcels is formulated as a capacitated vehicle routing problem and solved by the adaptive large neighborhood search metaheuristic. Results show that applying public transport-based crowdshipping could reduce the total vehicle kilometers traveled, the total working time of drivers, and the number of used vans (drivers) to perform last-mile deliveries. Moreover, public transport-based crowdshipping has great potential to reduce the total costs including driving costs, external costs, labor costs, and compensation.

Keywords: City logistics; Last-mile delivery; Crowdshipping; Impact assessment.

1. INTRODUCTION

Facing the exponential growth of E-commerce, both logistics service providers (e.g., DHL) and E-retailers (e.g., Amazon) have experimented with crowdshipping as a complementary solution to provide efficient last-mile delivery. In such a system, ordinary people utilize their free capacity regarding time and/or space to perform parcel delivery with monetary compensation.

Crowdshipping can be implemented in different ways. The main body of prior research and practical applications related to crowdshipping has focused on private personal vehicle use, where dedicated trips or detours are more or less unavoidable (Allahviranloo and Baghestani, 2019; Punel and Stathopoulos, 2017). Such personal vehicle-based concepts often entail rebound effects resulting in emission increases instead of decreases (Buldeo Rai et al., 2018). Meanwhile, sharing economic concepts has often been criticized for undermining workers' rights and creating a 'gig-economy' precariat (Paus, 2018).

To balance these considerations, this paper proposes a public transport (PT)-based crowdshipping concept (see Figure 1). In PT-based crowdshipping, automated parcel lockers (APLs) are

installed in some PT stations to store small parcels. A proportion of parcels are transported from the depot to PT stations by trucks and then delivered by crowdshippers. The crowdshippers are PT users, who pick up the parcels from APLs installed in the PT stations at PT users' origins before starting their PT trip and deliver the parcels to APLs installed in the PT stations at PT users' destinations. The final recipients take the parcels from the APLs in the stations near the parcels' destinations. The crowdshippers are compensated with credit for the transit system. This could ensure that only trips that would be taken anyway are utilized and that the task of crowdshipping cannot evolve into creating a new precarious job market lacking workers' rights.

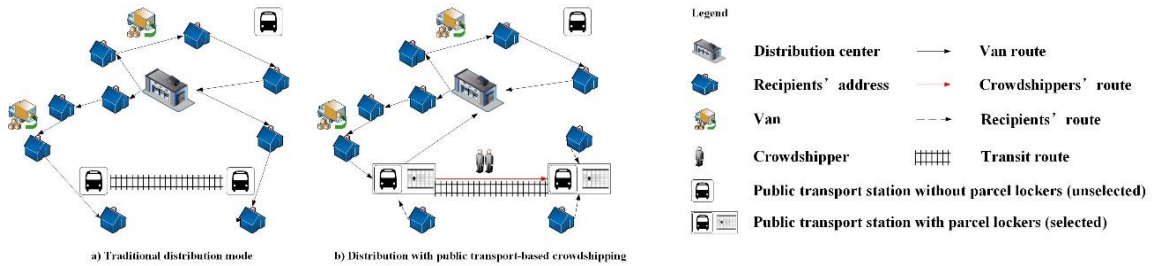


Figure 1 Comparison of traditional distribution mode and public transport-based crowdshipping

Three studies have examined the potential impacts of such a novel last-mile delivery system. All of them found that PT-based crowdshipping has positive impacts on reducing greenhouse gas emissions, vehicle kilometers traveled, etc. The PT-based crowdshipping concept in Gatta et al. (2018) and Karakikes and Nathanail (2022) is the same as our idea presented in Figure 1. Zhang et al. (2022) proposed another type of PT-based crowdshipping where PT users pick up parcels from APLs located in metro stations and deliver the parcels to the parcel's end destination instead of an APL in the PT user's destination stop.

Different from Gatta et al. (2018) and Karakikes and Nathanail (2022) which consider only installing APLs at metro stations, our study investigates a PT-based crowdshipping system with a denser network of APLs by extending the APLs network to some major bus stops so that more parcels could be delivered by crowdshippers. We determined the bus stops to install APLs and the vehicle routes to deliver the packages unserved by crowdshippers by solving a set covering problem and capacitated vehicle routing problem, respectively. The impacts of the PT-based crowdshipping were evaluated using real-world data in Copenhagen.

2. METHODOLOGY

To assess the impact of PT-based crowdshipping, three crowdshipping scenarios with varying percentages of crowdshipped parcels are simulated to be compared with the base scenario. The study area in this research is Nørrebro district in Copenhagen, the capital of Denmark. Nørrebro is located northwest of the city center, with an area of 3.82 km² and a population of 71891. It has a high population density and good public transport coverage.

Demand

The parcel delivery data is provided by PostNord – the largest logistics service provider in Denmark. The study period is from October 11th to October 17th, 2021, representing a normal

operation week. On average, 864 parcels with 492 delivery points are delivered per weekday and 480 parcels with 146 delivery points are delivered on the weekend.

The demand for crowdshipped parcels is influenced by many factors, e.g., goods' attributes (category, weight, size), recipients' social demographics (age, job), etc. Unfortunately, we do not have this information from the parcel data. Thus, we randomly select a certain number of parcels as crowdshipped parcels to reflect the ultimate effects of various influencing factors on demand for crowdshipping. When a sender places an order, he/she should specify whether the item to be delivered is a home parcel, collect parcel, or crowdshipped parcel.

Supply

Parcels distributed to Nørrebro are sorted in Brøndby distribution center, from where vans with small capacity depart, visit customers in Nørrebro, and finally return to the distribution center. The geographic distribution of the Brøndby distribution center and the study area is illustrated in Figure 2.



Figure 2 Study area

- Location of APLs

The main public transport means in Denmark include bus, metro, S-train, and train. S-train serves the Copenhagen metropolitan area. It has 86 stations that connect the suburban and urban areas. The S-train system carries more than 357000 passengers a day. Two S-train stations near Brøndby distribution center, i.e., Glostrup station and Brøndbyøster station, are selected to install APLs for crowdshippers to pick up the parcels. Crowdshipped parcels are placed in the APLs installed in the two stations before 7:30 am and will be completed on the same day.

We assume recipients only accept crowdshipped parcels delivered at PT stations within 500 meters of their original delivery addresses (i.e., home address of home delivery parcels and collect points of collect parcels). APLs are available at eight railway stations (i.e., S-train and metro stations) within or near Nørrebro district. Since the eight railway stations cannot serve all recipients, seven bus stops are selected, by solving a set covering problem proposed by Toregas et al. (1971), to serve recipients that railway stations cannot serve.

The capacity of APLs at each selected PT station is unlimited by installing sufficient APLs. This assumption makes sense considering the following reasons. First, APLs are easy to install – one person can complete them in several minutes. Second, APLs are inexpensive. The long-life battery inside the APL could work for ten years.

- Crowdshippers

The supply of crowdshipping is influenced by many factors, e.g., goods' attributes (category, weight, size), crowdshippers' social demographics (age, job), compensation, and extra time needed to perform delivery. Based on data from the Danish Rejsekort and the national Danish traffic model, around 400 daily trips are made from Brøndbyøster station and Glostrup station to Nørrebro. Although an additional effort is required to determine the number of trips to the selected PT stations where APLs are installed, this provides an estimation of the overall volume of passenger traffic. According to Fessler et al. (2022), when the compensation is 10 DKK per parcel, a passenger's probability of bringing a parcel is more than 30%. Increasing the compensation would attract more passengers working as crowdshippers and encourage crowdshippers to bring more parcels on their trips. To estimate the maximum potential benefits of PT-based crowdshipping, we assume that all crowdsourced parcels could be performed by crowdshippers by allowing crowdshippers to bring multiple parcels on their trips and increasing the compensation.

Scenario description

In the base scenario (S0), all parcels are distributed by the logistics companies using their vans. In the crowdshipping scenarios S1, S2, and S3, 10%, 20%, and 30% of the parcels are randomly selected and assigned to crowdshippers, while the rest are distributed by the delivery vans. The routes of the vans are determined by solving a standard capacitated vehicle routing problem using the adaptive large neighborhood search (ALNS) metaheuristic (Ropke and Pisinger, 2006). We generate 15 samples for each scenario of the daily parcel data to eliminate the stochastic effects of randomly selecting crowdshipped parcels. S1 is easy to achieve when most crowdshippers bring one parcel and the compensation is 10 DKK per parcel, while the realization of S2 and S3 requires more passengers to bring multiple parcels or more compensation paid to crowdshippers.

3. RESULTS AND DISCUSSION

Using the methodology introduced in section 2, we simulate the delivery operation of PostNord under different scenarios. Three indicators, vehicles kilometers traveled per day (including the travel distance of trucks that deliver crowdsourced parcels from the depot to selected PT stations), total working time of drivers (time spent by the driver from departure to return to the distribution center), and the number of used vans to serve Nørrebro, are used to describe the performance of each scenario. It is worth noting that the simulation result of the base scenario is validated by PostNord, which means the aforementioned three indicators obtained from our simulation are close to their actual operation on those days. The value of each indicator for each scenario is equal to the average value of the 15 samples of the scenario.

- Impacts on vehicle kilometers traveled

Figure 3 presents the percentage change of vehicle kilometers traveled during the study period under different crowdshipping scenarios. All the signs are negative, indicating that using PT-based crowdshipping as a complementary solution to last-mile delivery could reduce the vehicle kilometers traveled to deliver the parcels, even if some distances are needed to transport the crowdsourced parcels from the depot to PT stations. Moreover, the more parcels delivered by crowdshippers, the more percentage reduction of vehicle kilometers traveled occurred. First, it is shown that the average percentage reduction of vehicle kilometers traveled is 6%, 11%, and 20% under scenarios S1, S2, and S3, respectively. Second, the percentage reduction of vehicle kilometers traveled on the weekdays (8%, 14%, and 25% for scenarios S1, S2, and S3, respectively) is more significant than that on the weekend (2%, 4%, and 6% for scenarios S1, S2, and S3, respectively).

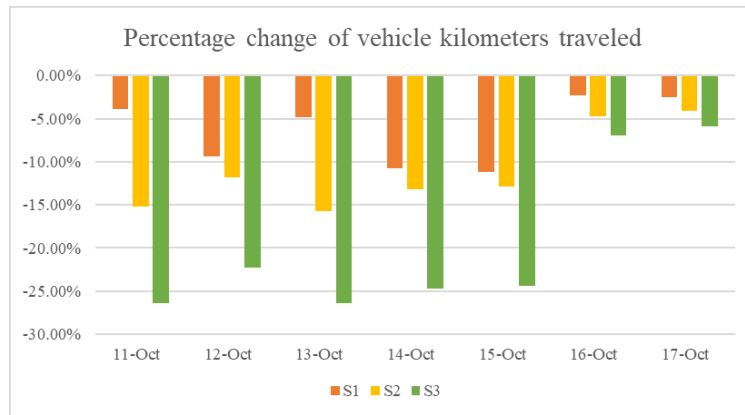


Figure 3: Percentage change of vehicle kilometers traveled under different scenarios

- Impacts on total working time of drivers

Figure 4 demonstrates the percentage change in drivers' total working time under different scenarios. On average, drivers' total working time could be reduced by 11%, 20%, 30% on weekdays and 7%, 15%, 21% on the weekend under scenarios S1, S2, and S3, respectively. This could reduce the increasing labor intensity of drivers.

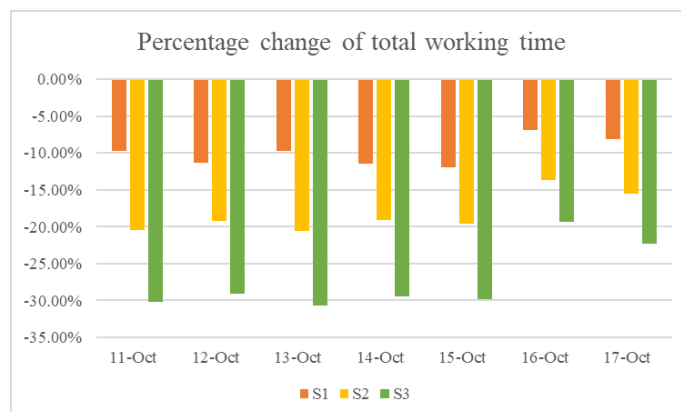


Figure 4: Percentage change of total working time under different scenarios

- Impacts on the number of used vans

Figure 5 shows the change of the number of used vans to serve Nørrebro. The simulation results are in line with our intuition that when some parcels are transferred from vans to crowdshippers, the number of used vans should be less than or equal to that in the base scenario. The reduction of used vans on weekend is zero, because only one van is used in the base scenario. The parcels cannot be completely delivered by crowdshippers. Therefore, one van is still used in crowdshipping scenarios. The number of used vans keeps unchanged in S1 on October 11th and October 13th. This indicates that the number of required vans (drivers) could be saved only when enough parcels are removed from vans to crowdshippers. Generally, if 20% of the parcels could be delivered by crowdshippers, one van (driver) is released. If the percentage of crowdsourced parcels reaches 30%, two vans (drivers) are released.

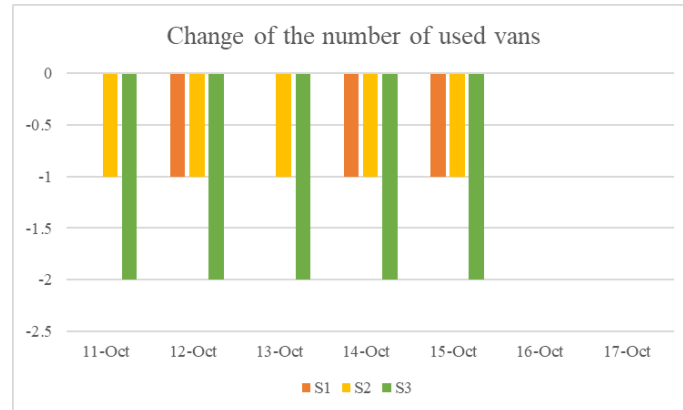


Figure 5: Change of the number of used vans under different scenarios

- Cost analysis

Four types of costs are related to the PT-based crowdshipping, i.e., driving costs of vans and trucks, external costs of traffic (e.g., marginal costs of air pollution, traffic congestion), drivers' salary, and compensation paid to crowdshippers. This section presents the potential benefits of PT-based crowdshipping based on the transport economic unit prices (TEUP) of 2022 prepared by Transport DTU and COWI for the Ministry of Transport (Denmark) (<https://www.man.dtu.dk/forskningsbaseret-raadgivning/teresa-og-transportoekonomiske-enhedspriser>).

The driving costs of vans and trucks include costs for fuel, tires, repair and maintenance, and depreciation. These costs are split into fixed and variable costs per hour and per kilometer, respectively, in TEUP. The fixed costs for vans and trucks are 529 DKK/hour and 542 DKK/hour, respectively. The variable costs for vans and trucks are 1.82 DKK/km and 4.19 DKK/km, respectively.

The negative externalities of transport include air pollution, climate change, noise, accidents, congestion, and wear on the infrastructure. The marginal external costs are used to estimate the cost per kilometer for the external effects. The marginal external costs for vans and trucks are 1.46 DKK/km and 6.01 DKK/km, respectively.

The average salary for a postal delivery worker is 24274 DKK per month (<https://www.paylab.com/dk/salaries-in-country?lang=en>), and the compensation for crowdshipper is 10 DKK per parcel as Fessler et al. (2023) did in the field test.

Table 2 presents the four types of costs under different crowdshipping scenarios. The distribution of each type of cost under different scenarios is similar. The driving and external costs account for 25% and 1% of the total costs, respectively. Labor costs account for about 70% of the total costs, and the compensation for crowdshippers accounts for 0%-5% of the total costs. Compared with the base scenario, on average, the total costs of S1, S2, and S3 are reduced by 8%, 13%, 24% on weekdays and 1%, 3%, 4% on the weekend, respectively. Based on Table 2, we conclude that by providing small compensation, PT-based crowdshipping has great potential to reduce last-mile delivery's driving costs. Meanwhile, it could benefit logistics companies by reducing labor costs, but it might negatively impact laborers in terms of salary.

Table 2: Cost analysis of public transport-based crowdshipping under different scenarios

		11-Oct	12-Oct	13-Oct	14-Oct	15-Oct	16-Oct	17-Oct
Driving costs (DKK)	S0	16123	17275	17010	16239	15230	2717	7747
	S1	14117	15431	14882	14502	13558	2487	6965
	S2	12569	13648	13228	12887	12034	2256	6228
	S3	11239	12192	11806	11529	10780	2082	5618
External costs (DKK)	S0	288	314	292	282	279	58	140
	S1	281	289	283	255	252	57	138
	S2	253	286	255	253	251	56	137
	S3	225	259	228	225	223	56	137
Labor costs (DKK)	S0	42480	48548	42480	42480	42480	6069	18206
	S1	42480	42480	42480	36411	36411	6069	18206
	S2	36411	42480	36411	36411	36411	6069	18206
	S3	30343	36411	30343	30343	30343	6069	18206
Compensation (DKK)	S0	0	0	0	0	0	0	0
	S1	870	970	870	830	800	120	360
	S2	1730	1930	1740	1650	1600	240	720
	S3	2600	2900	2610	2480	2390	360	1080
Total costs (DKK)	S0	58890	66137	59782	59000	57989	8843	26092
	S1	57747	59169	58514	51998	51021	8732	25668
	S2	50962	58344	51634	51201	50297	8621	25291
	S3	44407	51761	44987	44576	43735	8566	25041
Percentage change of total costs	S1	-2%	-11%	-2%	-12%	-12%	-1%	-2%
	S2	-13%	-12%	-14%	-13%	-13%	-3%	-3%
	S3	-25%	-22%	-25%	-24%	-25%	-3%	-4%

4. CONCLUSIONS

This study investigated the impact of applying PT-based crowdshipping as a complementary solution to traditional last-mile delivery. We select the study area of the Nørrebro district in Copenhagen because of its high population density and good public transport coverage. Post-Nord provides the parcel data. Three crowdshipping scenarios with different percentages of crowdsourced parcels are created to compare with the base scenario and identify the effects of

PT-based crowdshipping under different development stages. We use three indicators, i.e., vehicle kilometers traveled, total working time of drivers, and the number of used vans, to evaluate the performance of different scenarios. All values of the indicators are reduced with the increasing percentage of crowdshipped parcels. In the most optimistic scenario where 30% of the parcels are delivered by crowdshippers, the reduction percentage of the first two indicators reaches 20% and 27% on average, and two vans (drivers) are released. The cost analysis shows that significant potential savings on driving and labor costs could be achieved by transferring some parcels to crowdshippers and providing them with small compensation.

ACKNOWLEDGEMENTS

This research is supported by China Scholarship Council. We thank PostNord for providing the data.

REFERENCES

- Allahviranloo, M., & Baghestani, A. (2019). A dynamic crowdshipping model and daily travel behavior. *Transportation Research Part E: Logistics and Transportation Review*, 128, 175-190.
- Buldeo Rai, H., Verlinde, S., & Macharis, C. (2018). Shipping outside the box. Environmental impact and stakeholder analysis of a crowd logistics platform in Belgium. *Journal of Cleaner Production*, 202, 806-816.
- Fessler, A., Cash, P., Thorhauge, M., & Haustein, S. (2023). A public transport based crowdshipping concept: Results of a field test in Denmark. *Transport Policy*.
- Fessler, A., Thorhauge, M., Mabit, S., & Haustein, S. (2022). A public transport-based crowdshipping concept as a sustainable last-mile solution: Assessing user preferences with a stated choice experiment. *Transportation Research Part A: Policy and Practice*, 158, 210-223.
- Gatta, V., Marcucci, E., Nigro, M., Patella, S. M., & Serafini, S. (2018). Public transport-based crowdshipping for sustainable city logistics: Assessing economic and environmental impacts. *Sustainability*, 11(1), 145.
- Karakikes, I., & Nathanail, E. (2022). Assessing the Impacts of Crowdshipping Using Public Transport: A Case Study in a Middle-Sized Greek City. *Future Transportation*, 2(1), 55-83.
- Paus, E. (Ed.). (2018). *Confronting dystopia: The new technological revolution and the future of work*. Cornell University Press.
- Punel, A., & Stathopoulos, A. (2017). Modeling the acceptability of crowdsourced goods deliveries: Role of context and experience effects. *Transportation Research Part E: Logistics and Transportation Review*, 105, 18-38.
- Ropke, S., & Pisinger, D. (2006). An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation science*, 40(4), 455-472.
- Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations research*, 19(6), 1363-1373.
- Zhang, M., Cheah, L., & Courcoubetis, C. (2022). Exploring the Potential Impact of Crowdshipping Using Public Transport in Singapore. *Transportation Research Record*, 03611981221123246.

Enhancing Evacuation Planning and Management through Vehicular Communication

Hassan Idoudi^{1,2}, Mostafa Ameli¹, Cyril Nguyen Van Phu¹, Mahdi Zargayouna¹,
and Abderrezak Rachedi²

¹University Gustave Eiffel, COSYS, GRETTIA, Paris, France

²University Gustave Eiffel, LIGM, Paris, France,

SHORT SUMMARY

The current study presents a novel framework that aims to solve dynamic population evacuation (DPE) problems, divided into two phases: planning and online evacuation management, utilizing vehicular communication. During the planning phase, an initial evacuation plan is created by dynamically solving the shelter allocation problem (SAP) to determine destination choices and dynamic traffic assignment (DTA) to choose the best path to the selected destinations. Once the evacuation process begins, the vehicular ad hoc network (VANET) enables communication between evacuees, providing an opportunity to update initial decisions in real-time using VANET under the vehicular cloud computing (VCC) architecture, which considers the dynamic evolution of the hazard and traffic congestion levels. We apply the proposed online DPE framework to a test case in Luxembourg City to benchmark with existing planning methods. The results demonstrate that the proposed framework surpasses existing solution methods by more than 10% in network clearance time. Furthermore, the proposed framework's performance is evaluated by changing the penetration rate of connected vehicles in VANET, which provides additional insight into the framework's effectiveness.

Keywords: Network evacuation, online disaster management, Telecommunication network, VANET, shelter allocation, dynamic traffic assignment.

1 INTRODUCTION

According to Supian & Mamat (2022), the population residing in areas prone to natural disasters and catastrophes faces danger as the frequency of such incidents is on the rise due to climate change, leading to an increase in human casualties and environmental destruction. Effective evacuation orders are necessary to reduce the impact of these disasters, which can adapt to changing hazards and the needs of evacuees in real-time. This can be achieved through telecommunication technology, particularly dynamic population evacuation (DPE), using vehicle-to-everything (V2X) communication Pan et al. (2016); J. Wang et al. (2019).

Determining the best evacuation plans requires considering the disaster's characteristics, as the evacuation orders' objectives may vary depending on the type of disaster. The most common goals of evacuation orders are to minimize the mean evacuation time or the total time Supian & Mamat (2022).

To determine the best and most effective evacuation plans, it is crucial to consider the characteristics of the disaster at hand. The type of disaster plays a significant role in determining the objectives of evacuation orders given to evacuees. Typical objectives of such orders include minimizing the mean evacuation or the total time Hajjem et al. (2017); Bayram & Yaman (2018); Bayram et al. (2015), minimizing the network clearance time Hsu & Peeta (2014); Lim et al. (2015); Zhao et al. (2016), and minimizing the total traveled distance Sheu & Pan (2014); Alçada-Almeida et al. (2009).

Since the 1970s, DTA models have been used to analyze long-term, and short-term planning problems Han et al. (2015). The reactive nature of both SAP and DTA limits their effectiveness during the evacuation process, i.e., they are more contributing to the planning phase compared to online management Pan et al. (2013). In this context, adding telecommunication technologies moves one step forward by providing effective methods for proactive rerouting when an emergency is predicted based on real-time traffic information. Since traffic conditions are very time-varying

during an evacuation process, updating evacuation guidance messages frequently and quickly is critical.

With the emergence of intelligent and connected vehicles, vehicular networks, particularly vehicular ad hoc networks (VANET), were introduced in 2001 as a part of ad hoc mobile networks Olariu et al. (2011); Zeadally et al. (2012). VANET have received much attention from research communities in the last few years since it opened new doors of research (e.g., on vehicle and road safety, traffic efficiency, etc.) in intelligent transportation systems (ITS) Al-Sultan et al. (2014); Hartenstein & Laberteaux (2008).

With the growing demands of drivers, vehicles require empowering themselves in processing power, computing resources, and storage space. Despite all the efforts made to satisfy all these requirements, VANET shows some disadvantages, such as the high costs generated by communication between vehicles due to the high mobility of vehicles Qin et al. (2012). To support and serve all drivers' needs and ensure their comfort and safety, we have to increase the resources of VANET. As a result, the concept of vehicular cloud computing (VCC) has emerged Gerla (2012); Mekki et al. (2017) to enable vehicles to harness the benefits of cloud computing to satisfy certain requirements. VCC concept refers to the use of cloud computing in VANET Gerla (2012); Mekki et al. (2017). VCC allows vehicles to use the cloud resources required for a particular period, representing the time they need to achieve their goals.

In this study, we realize that there is no study in the literature about considering vehicle rerouting in the DPE context. However, the replanning decision is a critical part of the evacuation process and can impact the success of the evacuation. This study proposes an online evacuation framework to solve the DPE problem. The proposed methodology can dynamically assign evacuees to the best shelter considering the current traffic conditions. Our model uses an initial plan for evacuation that represents the output of solving SAP and DTA based on Idoudi et al. (2022). In our model, we consider two phases of the evacuation process:

- Planning phase, considering the initial evacuation plan solving both SAP and DTA problems, and
- Online evacuation management phase, which employs vehicular cloud computing technology to modify the initial evacuation plan by shelter reallocation and rerouting evacuees according to the dynamics of the network and evolution of the risk due to the disaster status.

Our methodology includes rerouting evacuees based on their distance from the risky zone and the density of vehicles on the way toward the shelters, considering their communication capacity. We implement the designed framework for a city-scale real test case to validate the model and compare the evacuation results in the presence and absence of telecommunication technology. In addition, we perform a sensitivity analysis on the penetration rate of equipped vehicles that can use the VANET.

The rest of the paper is organized as follows. In the next section, we present the framework to solve the evacuation problem. 3 is dedicated to presenting the case study and optimization scenarios. We discuss the results in 4 and present the concluding remarks in 5.

2 METHODOLOGY

The resolution of the DPE problem through our model involves two primary steps: constructing an evacuation plan by addressing the issues of SAP and DTA and providing real-time guidance to reroute vehicles as necessary in congested areas. In this section, we will provide a detailed account of the sequential process for executing each step of our formulation.

To provide an initial plan for planning purposes and in the dynamic setting, we adopt the methodology used in Idoudi et al. (2022). However, we aim to modify the planning model to adopt a stochastic user equilibrium (SUE) instead of a pure user equilibrium (UE) solution. We introduce a network layer for vehicular communication to capture network congestion for the online evacuation management phase. Using this communication network, we can re-plan evacuation routes and shelter locations during the evacuation process and provide real-time instructions to evacuees. We employ a cloud computing scheme to implement this methodology, which is advantageous due to its low implementation cost compared to fog or edge architectures, as noted in Gaouar & Lehsaini (2021).

Figure 1 depicts the proposed methodology of our study in the Plan-Do-Check-Act (PDCA) diagram format. The steps of the framework are detailed as follows:

Table 1: The steps of the methodological process described in Figure 1

<i>Start:</i>	
Step 1.	Initial evacuation plan: This step corresponds to solving the multi-level DTA and SAP to generate an evacuation plan. The SAP is going for system optimal (SO), and the DTA is formulated under SUE Idoudi et al. (2022).
<i>Plan:</i>	
Step 2.	Simulation for the current time step and set $t=t+1$: This step corresponds to simulating the evacuation process that could be the same as proposed by the plan, or new events could occur due to several decisions made by evacuees in the previous time step. We have also to increment the simulation time index.
Step 3.	Data collection: This is the first part of our cloud computing architecture wherein each vehicle (node) broadcasts data messages, using their OBU, to RSUs that send it to the cloud server.
Step 4.	Aggregation: In this step, we aggregate messages from different RSUs. An evacuee could be connected to more than one RSU and broadcast his message to all RSUs in his range of communication
<i>Do:</i>	
Step 5.	Risk update: In this step, we update the risk based on data from step 3. The considered risk consists of two main components: the vehicle's distance from a hazardous area and the congestion levels of the vehicle's location.
Step 6.	Prediction of new travel times: In this step, the travel time of edges might change according to the risk and congestion evolving by Step 4. In this step, we use a prediction model to predict new travel times.
<i>Check:</i>	
Step 7.	Check for replanning: This step is for deciding whether a user i is concerned by the rerouting process or not. For user i we estimate edge density, including the road speed and traffic density based on the Greenshield model Pan et al. (2016).
Step 8.	Evacuees selection for replanning: This step corresponds to selecting vehicles that to go to another safe destination or have to be rerouted before getting inside a congested edge (road). For shelter reallocation, we select vehicles if there is congestion in front of their original destinations, and the server asks them to go to a less congested destination.
<i>Act:</i>	
Step 9.	Shelter reallocation and rerouting: In this step, we prepare a message to the targeted users to ask them to reroute to the path with the current shortest travel time having their planned shelter as a safe destination.
Step 10.	Sending notification to evacuees: The step represents the second essential part of our cloud computing scheme where the cloud server sends its decisions to RSUs that forward the results to vehicles to react accordingly.
<i>End:</i>	
Step 11.	Check stopping condition: This step checks if all the demand is evacuated, go to 12 otherwise go to 2.
Step 12.	End of the simulation: In this step, we end the simulation of the evacuation process,
Step 13.	Result calculation : In this step and after ending the simulation, all results are then calculated in terms of packet delay ratio, end-to-end delay, and other measures.

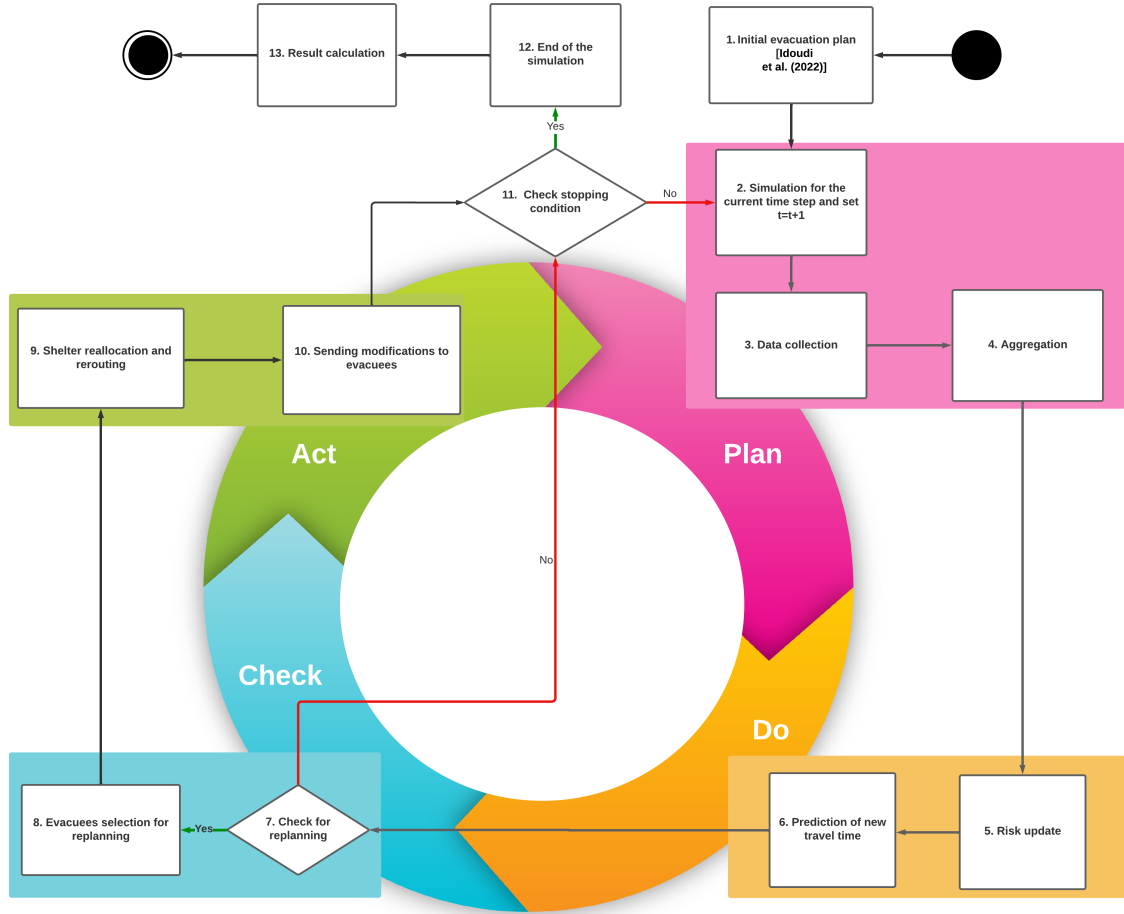


Figure 1: flowchart of the solving the DPE problem

3 NUMERICAL EXPERIMENTS

In the preceding section, we introduced our framework for addressing the online DPE problem. This section will apply the methodology to an actual network to validate the proposed solution. We will begin by describing the selected test case, followed by discussing the experimental design.

Case study

We implement our proposed solution on a realistic network of Luxembourg city. The LuST scenario provides the network. We have used a laptop with 1.7 GHz. and 16 GB of RAM to generate all the results. We employed a solution method using the simulation-based DTA. For this, we performed all simulations by SUMO simulator, and we calculated the C-logit model and the travel time prediction by SUMO Lopez et al. (2018). In addition, we used ILOG CPLEX version 12.9 to implement the SAP model and solve it. To simulate the scenario considering vehicular communication, we used the Veins/Omnet++ simulator and a cloud computing architecture based on works done in Z. Wang et al. (2020).

We applied our methodology to the realistic network of Luxembourg city Codeca et al. (2015). Please refer to Idoudi et al. (2022) for the network and evacuation scenario characteristics. Figure 2(a) presents the evacuation network map of Luxembourg. Figure 2(b) presents the real network of Luxembourg with the size of 155.95 km² and the traffic network graph considered by Veins for dynamic simulation. We examine a hypothetical threat in the center zone affecting people of that region colored in red in Figure 2(b). We do not assume any super source nodes (risky nodes) in this study. Four origin nodes are considered evacuation sources in the risk zone (see Figure 2c in Idoudi et al. (2022)). Vehicles carrying people should be evacuated to safe destinations (shelters), colored in green in Figure 2(b), and placed at the border of the network. We set the duration of each planning departure time interval (η) to 20 minutes for the simulation, considering the network's size. The demand at each node is 200 vehicles at each period. We have selected four

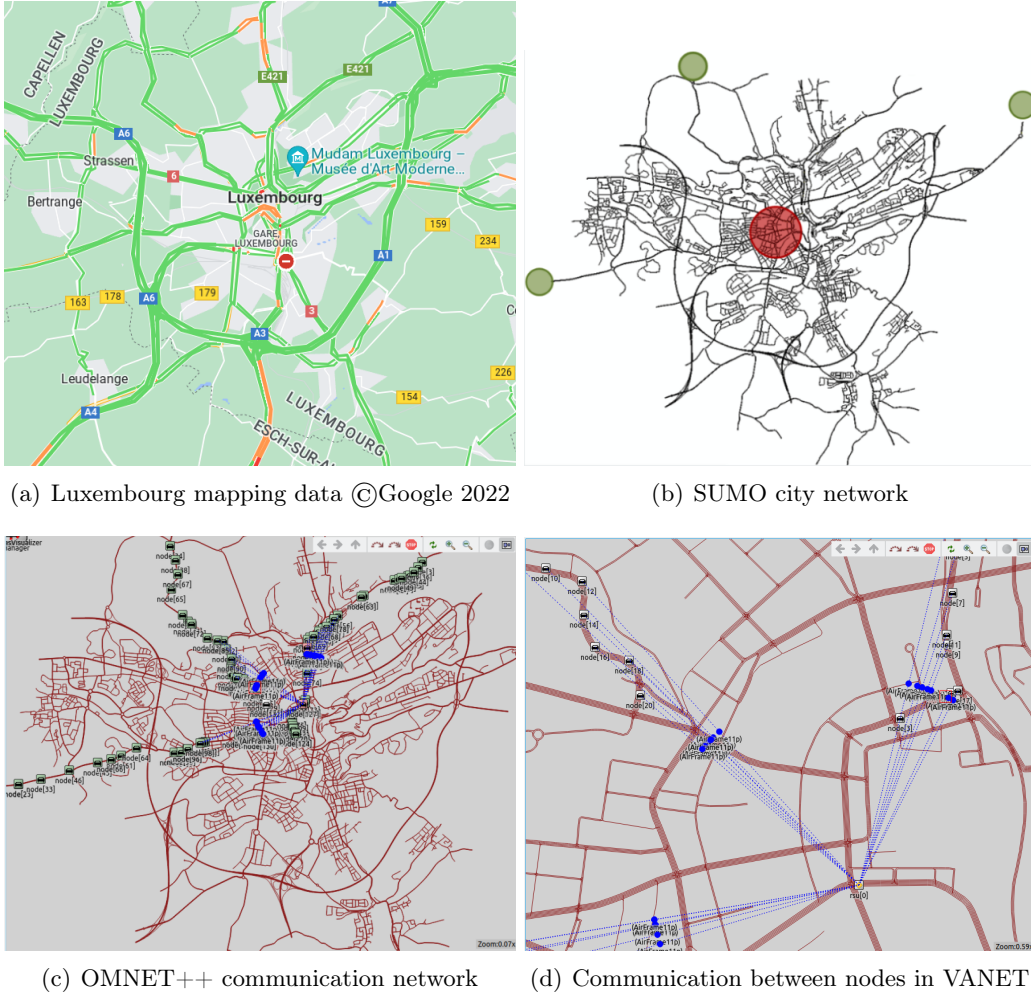


Figure 2: Vehicular communication map of Luxembourg city

origin nodes and four shelters, each with a capacity to hold 1500 evacuees. Therefore, the total demand is 600 vehicles per origin for the planning horizon (H). Figure 2(c) shows the vehicular communication network in the OMNET++ simulator. Figure 2(d) illustrates the message exchange process between vehicles and the network infrastructure.

Study optimization scenarios

In this study, we design four scenarios to investigate the impact of planning and online orders on the DPE problem. The scenarios are detailed below:

- **Scenario P+C: Scenario with both planning and vehicular communication:** This scenario follows the proposed framework (demonstrated in Figure 1).
- **Scenario P: Scenario with the initial plan only:** This scenario illustrates the case of just planning for evacuation without any communication between vehicles or vehicles to RSUs. It means that we do not reroute evacuees during the evacuation process; they just follow the initial plan.
- **Scenario C: Scenario with vehicular communication only:** This scenario is the same as Figure 1) except in step 1 where evacuees consider the nearest shelter and choose their routes following the SUE.
- **Scenario N: Naive scenario without any optimal plan and vehicular communication:** This scenario represents the case where the system operators do not provide guidelines for evacuees. It means the evacuees choose the nearest shelter and their routes following the SUE.

4 RESULTS

In this section, the results for the four mentioned scenarios were executed on the synthetic demand profile. Table 2 presents the results for the four scenarios. The results show a significant improvement in the quality of the final solution obtained by scenario P+C wherein we used both planning and online guidance models. For instance, the reduction of more than 18 minutes (39%) in the network clearance time compared to the naive scenario. Also, there is an improvement of more than 3 minutes (10%) between scenario P+C and scenario P. Results show that scenario P represents the second-best solution. The comparison between scenarios P+C and P proves that new orders handling new events not expected in planning create a more successful evacuation operation. Besides, scenario C provides a better solution than scenario N, meaning that using the telecommunication network can improve the evacuation solution, even without any planning phase. This observation could prove the effectiveness of online communication and highlights the importance of giving new orders to evacuees to revise their route choice during the evacuation process. Inspecting the result for scenario P and scenario C, we can observe that planning contributes more than telecommunication during the evacuation operation. One of the reasons behind this observation is that in scenario C the shelter allocation was done without considering the congestion level. We have monitored scenario C to have a better view and understand more of the effect of online evacuation guidance. We observe that allocating all users to the same nearest shelters in all evacuation operation generate congestion that cannot be escaped even by using online vehicle rerouting. That is why different shelters, like in scenario P in each state, will ensure that we assign evacuees to the closest destinations in terms of time-dependent shortest path and not distance measure.

Table 2: Performance metrics

Metrics / Scenario	P+C	P	C	N
Network clearance time(s)	1775.00	1980.00	2765.00	2835.0
Mean evacuation time(s)	1071.54	1093.70	1407.92	1447.61
Average travel delay (ATD)	205.47	220.62	341.63	349.78
Average evacuation delay (AED)	241.32	366.65	366.65	392.12

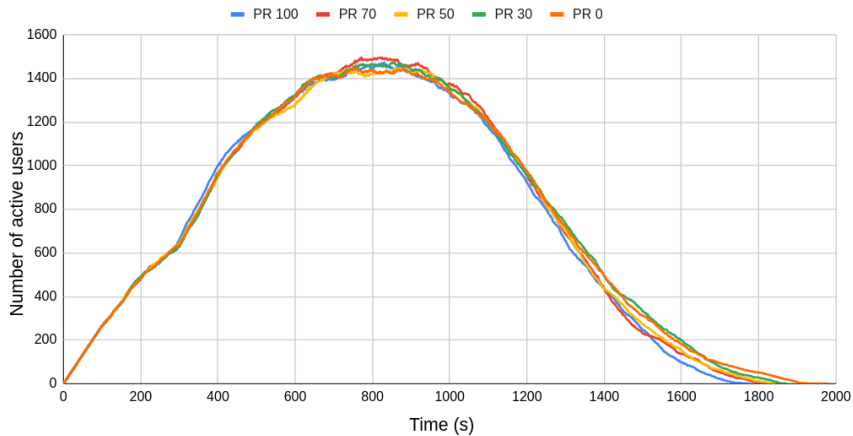
The decrease in mean evacuation time in 2 shows that the online DPE improves the evacuation solution. Compared to the second best, the proposed model used in scenario P+C generates better ATD for evacuees with more than 6% of reduction. The improvement is remarkable for AED (34%). We mention that including telecommunication network gives us some errors and delays in sending and receiving messages. In both cases, P+C and C, we have around 205.30 ms for the end-to-end delay and PDR around 74%.

Sensitivity analysis on penetration rate

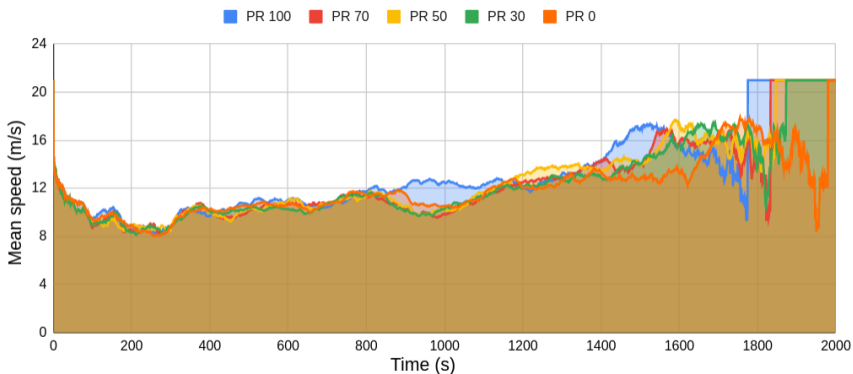
The sensitivity analysis on the penetration rate is performed on the Luxembourg city map. Assuming that 100% of the evacuees are using connected vehicles is not currently realistic, and it can be reachable in the future. That is why we should consider multiple penetration rate values. In the case of x% of penetration rate, we select connected vehicles with a random distribution. Only this x% is sending positioning information and receiving online orders. Thus, the cloud server sees and guides only this x% of the vehicles.

Figure 3(a) illustrates the change in the number of vehicles evacuating in the network for five scenarios. The curves shown in this figure represent different penetration rate values.

Figure 3(b) follows the results of the distribution of accumulation of users in the network. In addition, Figure 3(b) depicts the evolution of the mean speed in the evacuation operation. The maximum network speed limit is the free-flow speed (21 m/s) attained when the network does not have any vehicles. The network speed illustrated by Figure 3(b) shows that having a 100% penetration rate is the fastest curve by arriving at the free-flow speed in the shortest time. Also, the figure presents the result of the mean speed variation of other penetration rates showing that there is not a huge difference between 70% and 30% penetration rate on network clearance time (the arrival to the free-flow speed). Figure 3(b) shows that adding the communication layer, even with different penetration rates, positively affects the evacuation process. It means that the online solving of DPE uses the network's capacity better than just planning. We conclude that using 30% of the penetration rate is more realistic, and its results are comparable to having 100%.



(a) Number of active users in the network variation



(b) Network mean speed variation

Figure 3: Performance measures variation over different penetration rates

5 CONCLUSIONS

The timely evacuation of affected populations during a disaster is critical in reducing the overall impact of the event. In this paper, we focus on the dynamic population evacuation (DPE) problem and propose a framework for effectively modeling and optimizing the evacuation process to save as many lives as possible in a faster and more efficient manner. We divide the evacuation problem into two parts: the first part involves creating an optimal evacuation plan that considers dynamic shelter allocation and traffic assignment, while the second part involves considering new orders for the online guiding system.

Our framework captures the dynamics of the evacuation process by using a traffic simulator to build an evacuation planning process to determine shelters and routes. We then perform an online management procedure during the evacuation, allowing vehicles to send and receive data to update their routes. To achieve this, we use a cloud computing architecture comprising vehicles, roadside units (RSUs), and a distant cloud server.

We apply our methodology to the real-world networks of Luxembourg and show that our proposed model outperforms the model with only evacuation planning by reducing the network clearance time by more than 10% in the medium-scale network of Luxembourg. Our framework also effectively improves network capacity in terms of speed even at a low penetration rate of connected vehicles. We consider only rerouting and shelter reallocation to manage the online evacuation process in this study, and currently, we are working on a real large-scale test case of evacuation in California state in USA. We also plan to perform a sensitivity analysis on the shelter allocation objective to search for the best objective that minimizes more the clearance time measure. As future work could consider departure times before and during the evacuation, as well as the behavioral reactions of users to evacuation orders. We also aim to extend our framework to include other modes of transport, such as buses. Additionally, we aim to improve our framework by implementing a more accurate travel time predictor.

REFERENCES

- Alçada-Almeida, L., Tralhão, L., Santos, L., & Coutinho-Rodrigues, J. (2009). A multiobjective approach to locate emergency shelters and identify evacuation routes in urban areas. *Geographical analysis*, 41(1), 9–29.
- Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H., & Zedan, H. (2014). A comprehensive survey on vehicular ad hoc network. *Journal of network and computer applications*, 37, 380–392.
- Bayram, V., Tansel, B. Ç., & Yaman, H. (2015). Compromising system and user interests in shelter location and evacuation planning. *Transportation research part B: methodological*, 72, 146–163.
- Bayram, V., & Yaman, H. (2018). Shelter location and evacuation route assignment under uncertainty: A benders decomposition approach. *Transportation science*, 52(2), 416–436.
- Codeca, L., Frank, R., & Engel, T. (2015). Luxembourg sumo traffic (lust) scenario: 24 hours of mobility for vehicular networking research. In *2015 ieee vehicular networking conference (unc)* (p. 1-8). doi: 10.1109/VNC.2015.7385539
- Gaouar, N., & Lehsaini, M. (2021). Toward vehicular cloud/fog communication: A survey on data dissemination in vehicular ad hoc networks using vehicular cloud/fog computing. *International Journal of Communication Systems*, 34(13), e4906.
- Gerla, M. (2012). Vehicular cloud computing. In *2012 the 11th annual mediterranean ad hoc networking workshop (med-hoc-net)* (pp. 152–155).
- Hajjem, M., Bouziri, H., Talbi, E.-G., & Mellouli, K. (2017). Intelligent indoor evacuation guidance system based on ant colony algorithm. In *2017 ieee/acs 14th international conference on computer systems and applications (aiccsa)* (pp. 1035–1042).
- Han, K., Szeto, W., & Friesz, T. L. (2015). Formulation, existence, and computation of boundedly rational dynamic user equilibrium with fixed or endogenous user tolerance. *Transportation Research Part B: Methodological*, 79, 16–49.
- Hartenstein, H., & Laberteaux, L. (2008). A tutorial survey on vehicular ad hoc networks. *IEEE Communications magazine*, 46(6), 164–171.
- Hsu, Y.-T., & Peeta, S. (2014). Risk-based spatial zone determination problem for stage-based evacuation operations. *Transportation research part C: emerging technologies*, 41, 73–89.
- Idoudi, H., Ameli, M., Van Phu, C. N., Zargayouna, M., & Rachedi, A. (2022). An agent-based dynamic framework for population evacuation management. *IEEE Access*, 10, 88606–88620.
- Lim, G. J., Rungta, M., & Baharnemati, M. R. (2015). Reliability analysis of evacuation routes under capacity uncertainty of road links. *Iie Transactions*, 47(1), 50–63.
- Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., ... Wießner, E. (2018). Microscopic traffic simulation using sumo. In *The 21st ieee international conference on intelligent transportation systems*. IEEE. Retrieved from <https://elib.dlr.de/124092/>
- Mekki, T., Jabri, I., Rachedi, A., & ben Jemaa, M. (2017). Vehicular cloud networks: Challenges, architectures, and future directions. *Vehicular Communications*, 9, 268–280.
- Olariu, S., Khalil, I., & Abuelela, M. (2011). Taking vanet to the clouds. *International Journal of Pervasive Computing and Communications*, 7(1), 7–21.
- Pan, J., Popa, I. S., & Borcea, C. (2016). Divert: A distributed vehicular traffic re-routing system for congestion avoidance. *IEEE Transactions on Mobile Computing*, 16(1), 58–72.
- Pan, J., Popa, I. S., Zeitouni, K., & Borcea, C. (2013). Proactive vehicular traffic rerouting for lower travel time. *IEEE Transactions on vehicular technology*, 62(8), 3551–3568.

- Qin, Y., Huang, D., & Zhang, X. (2012). Vehicloud: Cloud computing facilitating routing in vehicular networks. In *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (pp. 1438–1445).
- Sheu, J.-B., & Pan, C. (2014). A method for designing centralized emergency supply network to respond to large-scale natural disasters. *Transportation research part B: methodological*, *67*, 284–305.
- Supian, S., & Mamat, M. (2022). Insurance as an alternative for sustainable economic recovery after natural disasters: A systematic literature review. *Sustainability*, *14*(7), 4349.
- Wang, J., Shao, Y., Ge, Y., & Yu, R. (2019). A survey of vehicle to everything (v2x) testing. *Sensors*, *19*(2), 334.
- Wang, Z., Zheng, S., Ge, Q., & Li, K. (2020). Online offloading scheduling and resource allocation algorithms for vehicular edge computing system. *IEEE Access*, *8*, 52428–52442.
- Zeadally, S., Hunt, R., Chen, Y.-S., Irwin, A., & Hassan, A. (2012). Vehicular ad hoc networks (vanets): status, results, and challenges. *Telecommunication Systems*, *50*(4), 217–241.
- Zhao, X., Ren, G., & Huang, Z.-f. (2016). Optimizing one-way traffic network reconfiguration and lane-based non-diversion routing for evacuation. *Journal of Advanced Transportation*, *50*(4), 589–607.

Using Explainable Machine Learning to Interpret the Effects of Policies on Air Pollution: COVID-19 Lockdown in London

Liang Ma*¹, Daniel J. Graham¹, Marc E.J. Stettler¹

¹ Centre for Transport Studies, Department of Civil and Environmental Engineering,
Imperial College London

Short Paper submitted for presentation at the
11th Symposium of the European Association for Research in Transportation
Zurich, 2023

SHORT SUMMARY

Activity changes during the COVID-19 lockdown brought an unprecedented opportunity to understand the likely effectiveness of prospective air quality management policies on reducing air pollution. Using a regression discontinuity design for causal analysis, we show that the first UK national lockdown led to unprecedented decreases in road traffic yet incommensurate and heterogeneous responses in air pollution in London. At different locations, changes in air pollution attributable to the lockdown ranged from -50% to 0% for NO₂, 0% to +4% for O₃, -5% to +0% for PM₁₀ and there was no response for PM_{2.5}. Using explainable machine learning, we show that the degree to which NO₂ pollution was reduced in an area was correlated with spatial features (including road freight traffic and proximity to a major airport and the city centre), and that existing inequalities in air pollution exposure were exacerbated: pollution reductions were greater in places with more affluent residents and better access to public transport services.

Keywords: Air pollution; Causal analysis; COVID-19; Explainable machine learning.

1. INTRODUCTION

Various interventions have been implemented in cities to improve air quality. However, the impact pathways from an intervention to air quality can incorporate various factors and complicated interactions among factors, which presents challenges to both isolating the intervention effect and quantifying the contribution of different factors to the net effect. Recently, the rapid development of interpretation methods for machine learning (ML) models to achieve explainable ML has provided an opportunity to gain insights into complicated relationships among high-dimensional variables (Molnar et al., 2022). Unlike black-box ML models, explainable ML seeks to make the outputs and processes of ML models more interpretable and understandable to humans and has been applied in various fields to support decision-making (Guidotti et al., 2018; Molnar et al., 2022).

Activity changes during the COVID-19 lockdown brought an unprecedented opportunity to understand the likely effectiveness of prospective emission control policies in improving air quality. Different methods have been applied to quantify the air quality impacts of lockdowns, including comparing air quality levels before and after the lockdown, using bottom-up simulations, predicting business-as-usual concentrations with ML models, and applying causal

* Proposed presenting author. Email: liang.ma13@imperial.ac.uk

inference methods (Dang & Trinh, 2021; Jephcote et al., 2021; Menut et al., 2020; Venter et al., 2020). Compared with other methods applied, causal inference methods generally have advantages in data requirement, model building, and the interpretation of effect estimates. However, due to the complexity of atmospheric processes, previous studies of this type had a typical limitation by using a parametric confounding control for weather conditions. Meanwhile, several studies found heterogeneous air quality effects of lockdowns (Fassò et al., 2021; Vadrevu et al., 2020). However, most of them focused on depicting the variation in effects across subsamples of cities or countries; few compared effects within a city or further evaluated the contribution of factors to the heterogeneity of effects.

In this paper, we provide an analysis of the changes in London’s air quality attributable to the first UK COVID-19 lockdown with a causal inference approach. To overcome the typical limitation in confounding control, the meteorological normalisation technique is applied to non-parametrically control for weather conditions and seasonality effects. To further understand the spatial heterogeneity in the lockdown impacts, we additionally evaluate the contribution of different factors to the level of lockdown impacts by interpreting a predictive ML model with SHapley Additive exPlanations (SHAP) values. Various factors are considered, such as economy, demographics, transport demand, public transport supply, and geographical location. The identification of key factors affecting pollution reduction within a city can provide further guidance in improving air quality and help to shape our city with better and more equitable design, planning, and management.

2. METHODS

The research framework of this paper is shown in Figure 1. The method for quantifying the causal air quality impacts mainly follows Ma et al. (2021b). The evaluation is applied at individual air quality monitoring sites within London for regulated air pollutants including NO_2 , O_3 , $\text{PM}_{2.5}$, and PM_{10} , and a non-regulated pollutant NO_x . Specifically, meteorological normalisation is first applied to control important confounders. A Gradient Boosting Decision Trees (GBDT) is used to consider complex relationships among model variables. A normalised concentration time series is derived by removing the variation in the observed concentrations that can be explained by input confounders. Change point detection (CPD) is then conducted to detect structural changes in the normalised concentration time series; the results are used to support the research period specification and test a key model assumption for the next step. A sharp regression discontinuity design (RDD) is then specified on the normalised concentrations where a site had at least one change point around the start of the lockdown T_0 , i.e. showed a response. As a causal inference method, sharp RDD goes beyond comparing air quality before and after T_0 . The inference is based on a trend function approximation and quantification of effect at the discontinuity of a trendline either side of T_0 , which makes it less vulnerable to random noise and unrelated events (Ma et al., 2021b). With the estimated RDD coefficients, the lockdown effect τ is derived by stacking the impact from the current daily period and those from lagged periods. The interval estimates of τ is provided with a Monte Carlo simulation in Ma et al. (2021a). The τ estimates at different sites are aggregated with a bootstrapping in Ma et al. (2021a) for a city-wide mean.

To further interpret the spatial heterogeneity of impacts, 124 spatial features are evaluated based on their contribution in predicting the pollution reduction due to lockdown, using explainable ML (details in Figure 1). The evaluation is conducted at the Middle Layer Super Output Area (MSOA) level and focuses on the changes in annual mean concentrations of NO_2 in 2020 caused by the lockdown. Particularly, the annual mean concentration is estimated in two scenarios at individual sites: with and without lockdown. Mapping methods are then applied to estimate the pollution

reductions attributable to the lockdown at MSOAs; the mapping model mainly follows Horálek et al. (2019), which combines a linear regression and ordinary kriging of residuals. A GBDT model is then built on the estimated pollution reduction with the features at MSOAs, separately for absolute and relative reductions. The key hyperparameters of GBDT are automatically tuned with Bayesian optimisation. As the widely used feature importance metrics (such as total gains) do not directly indicate how the level of the output variable is affected by features, we use the SHAP values (an approximation of Shapley values) to interpret our GBDT models. The Shapley value is a key solution in cooperative game theory and can also be classified as an additive feature attribution method for a local explanation of complex predictive models (Lundberg et al., 2020). An additive feature attribution method is characterised by assigning a contribution ϕ_i to each feature, with the sum of $\{\phi_i\}$ approximating the original model predict (Lundberg et al., 2020). Therefore, ϕ_i is in the same unit as the model’s output and, consequently, can provide a better interpretation compared with using total gains. Moreover, Shapley values have advantages over the other methods of this class, as they can provide a single unique solution to assigning contributions with desirable properties (Lundberg et al., 2020).

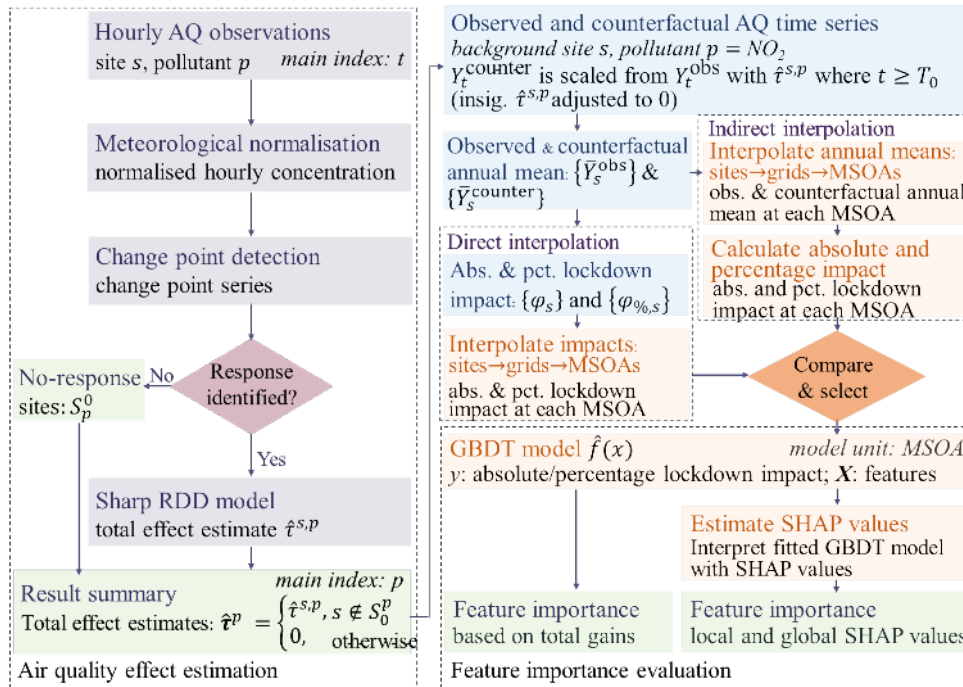


Figure 1. Graphical summary of the methodology for air quality effect estimation (left) and feature evaluation (right). The output/conclusion is coloured green. The left column (purple) goes through each concentration time series. The right column is either conducted on each monitoring site (blue) or focused on MSOAs (orange).

3. RESULTS AND DISCUSSION

COVID-19 lockdown effects on air quality

Road traffic in London dropped by up to 65% during the lockdown (Transport for London, 2020). However, our results show that the relative change in NO_2 caused by the lockdown ranged from -50% to 0% at different monitoring sites. Among them, 44% of the 45 sites close to roads

(roadside sites) and 70% of the 33 sites measuring town-wide pollution levels (background sites) showed a small reduction (<10%) or a null response. Aggregating the effects across London, the lockdown reduced NO₂ concentrations by 12% at roadside sites and 7% at background sites on average. For other pollutants, the relative change attributable to the lockdown ranged from -62% to 0% for NO_x, 0% to 4% for O₃, -5% to +0% for PM₁₀, and there was no response for PM_{2.5}. Unlike NO₂ and NO_x, the regional contribution to PM and O₃ is substantial (Greater London Authority, 2020). Our results imply that reducing transport activities and restricting exhaust emissions are not sufficient to tackle air pollution, particularly for those at background locations and for pollutants that are largely affected by regional emission sources.

Our estimated air quality impacts of the lockdown are generally consistent with previous studies in that we find more marked reductions in NO₂ concentrations yet less significant changes in other regulated pollutants. However, our estimates are not as large as those of Jephcote et al. (2021), who reported an average reduction of 38% and 17% respectively in NO₂ and PM_{2.5} concentrations and an average increase in O₃ concentrations of 8% across the UK during the first lockdown. Jephcote et al. (2021) estimated the impacts by comparing the air quality observations between the lockdown period and the same period in previous years. However, this approach may be biased by differing meteorological conditions and long-term trends in air quality. Particularly, air quality in the UK (both NO₂ and PM) has improved year on year in most major cities, including London (Department for Environment Food & Rural Affairs, 2020; Ma et al., 2021b). Therefore, comparing air pollution levels across different years is likely to overestimate the air quality impacts attributable to the lockdown.

Another study by Shi et al. (2021) reported abrupt but smaller than expected changes in air quality attributable to the lockdown in 11 cities globally. For London, they found that the lockdown changed background concentrations of NO₂, O₃, and PM_{2.5} respectively by $-8 \pm 8\%$, $-2 \pm 8\%$, and $+11 \pm 17\%$. While our results for NO₂ and O₃ are similar, the results for PM_{2.5} differ. Shi et al. (2021) estimated the lockdown impact by the relative change in air pollution before and after the lockdown in 2020 after subtracting the relative change over the same period in the average concentrations across the previous 4 years. Their estimated impacts for PM could be biased by PM episodes during their research period, between March and April in 2016-2020. PM episodes are a regular feature in springtime in western Europe; particularly, two episodes due to regional pollution transport were recorded in London in their specified post-lockdown period (Imperial College London, 2021). Although they applied a meteorological normalisation, this technique may not be effective to control for abrupt natural events or regional pollution transport (Shi et al., 2021). Consequently, the pollution increase estimated in Shi et al. (2021) for PM_{2.5} may incorporate the influence of these recorded episodes. In contrast, our study focuses on the time around the start of the lockdown, using CPD for response identification and subsequently a sharp RDD for effect estimation, and are therefore less susceptible to the influence of pollution episodes.

Factors affecting lockdown effects

The first UK national lockdown is estimated to have caused spatially heterogeneous impacts on NO₂ concentrations in London. By using explainable ML, we find that the degree to which NO₂ was reduced in an area was mostly correlated with the proportion of heavy goods vehicles (HGVs) in road traffic before the lockdown, the distance to London Heathrow Airport (LHR), and the distance to the Central Activities Zone (city centre); see Figure 2. This finding is generally consistent with Yang et al. (2021), who found the NO₂ reduction in Los Angeles during the lockdown was primarily due to the changes in HGVs' activities. In addition to feature ranking, our results also show that a lower proportion of HGVs in road traffic and closer proximity to LHR

are associated with greater pollution reductions in most MSOAs; however, different areas can have opposite effects from a short distance to the city centre.

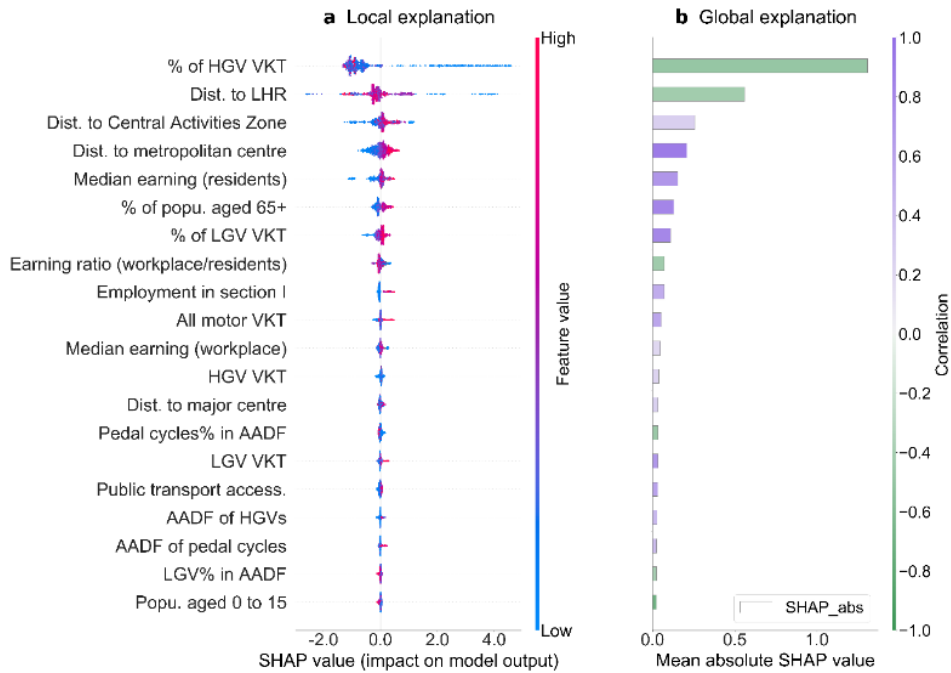


Figure 2. Estimated local SHAP values (left) and global feature importance (right) quantifying the contribution of each feature to the relative reduction in NO₂ due to lockdown. **a** Each point is a SHAP value specific to a feature and an MSOA; the colour of a point indicates the feature value for that MSOA; points are spread out in the y-axis to avoid overlap. **b** The global importance is calculated as the mean absolute SHAP value across MSOAs; the colour indicates the Pearson correlation between feature values and SHAP values across MSOAs; for example, a negative correlation (green) indicates that a higher feature value is generally associated with a smaller pollution reduction. Features are sorted by global feature importance; the top 20 features are plotted.

Furthermore, our results reveal that the existing health inequalities within the city were exacerbated during the lockdown: areas with lower-income residents, which have also been exposed to high levels of air pollution before the pandemic (Barnes et al., 2019), typically experienced smaller NO₂ pollution reductions during the lockdown. Particularly, areas with median incomes below the 10th percentile experienced a mean decrease in pollution reduction of 0.6 percentage points (pp), while those with median incomes above the 90th percentile experienced a mean increase in pollution reduction of 0.3 pp (Figure 3). Moreover, we find places with higher public transport accessibility levels were generally associated with a larger pollution reduction (Figure 3). Therefore, our results highlight the link between existing social inequalities and the effects of air pollution control policies.

While close proximity to the city centre is estimated to enhance pollution reduction in several cases (Figure 2), the areas that are closer to metropolitan town centres were commonly less affected compared with the areas further away (Figure 3). As people tended to make local trips during the lockdown, our result implies the potential of alleviating air pollution in the city centre by revitalising local town centres to reduce the need for travel. Particularly, a shift towards remote/hybrid working may continue in the UK after the pandemic (Transport for London, 2020), which is likely to further enhance the role of town centres. However, policymakers should be

mindful of the potential deterioration of air quality around town centres following the development.

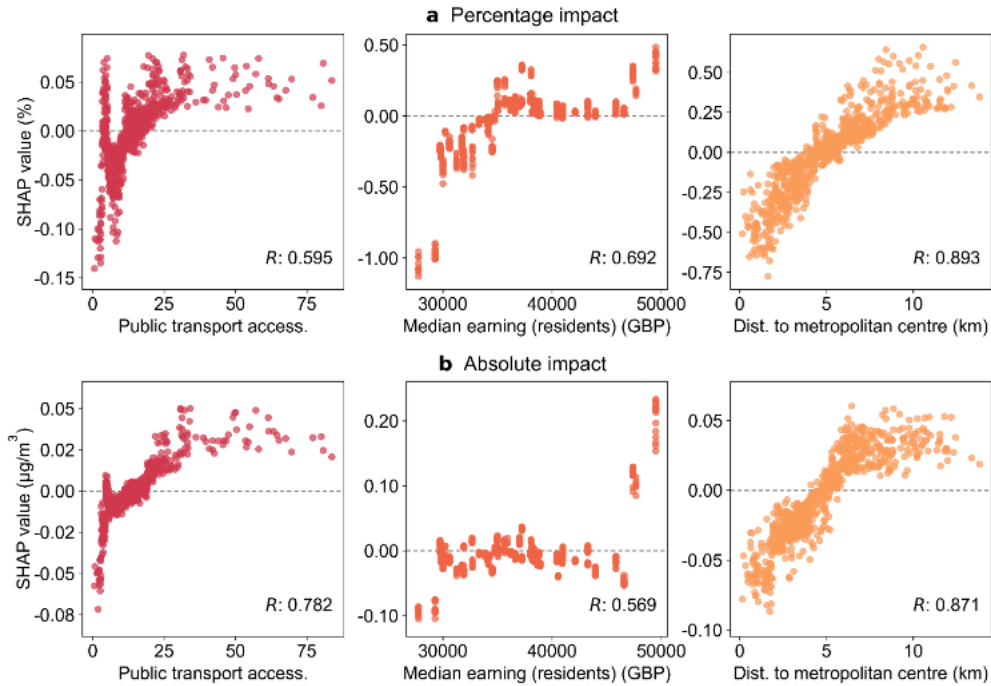


Figure 3. Relationship between public transport accessibility level, median gross annual pay for residents, the distance to the nearest metropolitan town centre, and the SHAP value for **a** the relative impact (pp) and **b** the absolute impact ($\mu\text{g}/\text{m}^3$). Each point shows the feature value (x-axis) and corresponding SHAP value (y-axis) at a particular MSOA. The sign of the SHAP value indicates how the feature value contributes to the predict of the lockdown impact at that MSOA (positive: increasing pollution reduction; negative: decreasing pollution reduction). The Pearson correlation (R) between the feature value and the corresponding SHAP value across MSOAs is labelled on each plot.

4. CONCLUSIONS

Our study shows that the unprecedented decrease in transport activities following the COVID-19 lockdown led to an incommensurate reduction in air pollution. Improving air quality in cities requires a multi-faceted set of policies to control emissions across sectors and full consideration of inequalities. Within the transport sector, a sustained effort is necessary to consistently guide the city for health and equality, such as by reducing emissions for freight transport and airport-related activities, facilitating and encouraging active travel and public transport, and integrating transport planning with land use planning.

ACKNOWLEDGMENTS

L M was funded by the Dixon and Skempton Scholarships from the Department of Civil and Environmental Engineering, Imperial College London.

REFERENCES

- Barnes, J. H., Chatterton, T. J., & Longhurst, J. W. S. (2019). Emissions vs exposure: Increasing injustice from road traffic-related air pollution in the United Kingdom. *Transportation Research Part D: Transport and Environment*, 73, 56–66. <https://doi.org/10.1016/j.trd.2019.05.012>
- Dang, H. A. H., & Trinh, T. A. (2021). Does the COVID-19 lockdown improve global air quality? New cross-national evidence on its unintended consequences. *Journal of Environmental Economics and Management*, 105, 102401. <https://doi.org/10.1016/j.jeem.2020.102401>
- Fassò, A., Maranzano, P., & Otto, P. (2021). Spatiotemporal variable selection and air quality impact assessment of COVID-19 lockdown. *Spatial Statistics*, 100549. <https://doi.org/10.1016/j.spasta.2021.100549>
- Greater London Authority. (2020). *Air Quality In London 2016-2020*. https://www.london.gov.uk/sites/default/files/air_quality_in_london_2016-2020_october2020final.pdf
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A Survey of Methods for Explaining Black Box Models. *ACM Comput. Surv*, 51(93). <https://doi.org/10.1145/3236009>
- Horálek, J., Schreiberová, M., Schneider, P., Kurfürst, P., Schovánková, J., & Ďoubalová, J. (2019). *European air quality maps for 2016* (Issue 9).
- Imperial College London. (2021). *LAQN Pollution Episodes*. <https://www.londonair.org.uk/london/asp/publicisodes.asp>
- Jephcote, C., Hansell, A. L., Adams, K., & Gulliver, J. (2021). Changes in air quality during COVID-19 ‘lockdown’ in the United Kingdom. *Environmental Pollution*, 272, 116011. <https://doi.org/10.1016/j.envpol.2020.116011>
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
- Ma, L., Graham, D. J., & Stettler, M. E. J. (2021a). Air quality impacts of new public transport provision: A causal analysis of the Jubilee Line Extension in London. *Atmospheric Environment*, 245. <https://doi.org/10.1016/j.atmosenv.2020.118025>
- Ma, L., Graham, D. J., & Stettler, M. E. J. (2021b). Has the ultra low emission zone in London improved air quality? *Environmental Research Letters*, 16(12), 124001. <https://doi.org/10.1088/1748-9326/ac30c1>
- Menut, L., Bessagnet, B., Siour, G., Mailler, S., Pennel, R., & Cholakian, A. (2020). Impact of lockdown measures to combat Covid-19 on air quality over western Europe. *Science of the Total Environment*, 741, 140426. <https://doi.org/10.1016/j.scitotenv.2020.140426>
- Molnar, C., König, G., Herbringer, J., Freiesleben, T., Dandl, S., Scholbeck, C. A., Casalicchio, G., Grosse-Wentrup, M., & Bischl, B. (2022). General Pitfalls of Model-Agnostic Interpretation Methods for Machine Learning Models. In *Lecture Notes in Computer Science* (Vol. 13200, pp.

39–68). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-04083-2_4

Shi, Z., Song, C., Liu, B., Lu, G., Xu, J., Van Vu, T., Elliott, R. J. R., Li, W., Bloss, W. J., & Harrison, R. M. (2021). Abrupt but smaller than expected changes in surface air quality attributable to COVID-19 lockdowns. *Science Advances*, 7(3), 6696–6709. <https://doi.org/10.1126/sciadv.abd6696>

Transport for London. (2020). *Travel in London Report 13*. <http://content.tfl.gov.uk/travel-in-london-report-13.pdf>

Vadrevu, K. P., Eaturu, A., Biswas, S., Lasko, K., Sahu, S., Garg, J. K., & Justice, C. (2020). Spatial and temporal variations of air pollution over 41 cities of India during the COVID-19 lockdown period. *Scientific Reports*, 10(1), 1–15. <https://doi.org/10.1038/s41598-020-72271-5>

Venter, Z. S., Aunan, K., Chowdhury, S., & Lelieveld, J. (2020). COVID-19 lockdowns cause global air pollution declines. *Proceedings of the National Academy of Sciences of the United States of America*, 117(32), 18984–18990. <https://doi.org/10.1073/pnas.2006853117>

Yang, J., Wen, Y., Wang, Y., Zhang, S., Pinto, J. P., Pennington, E. A., Wang, Z., Wu, Y., Sander, S. P., Jiang, J. H., Hao, J., Yung, Y. L., & Seinfeld, J. H. (2021). From COVID-19 to future electrification: Assessing traffic impacts on air quality by a machine-learning model. *Proceedings of the National Academy of Sciences of the United States of America*, 118(26), 2102705118. <https://doi.org/10.1073/pnas.2102705118>

Faster estimation of discrete choice models via weighted dataset reduction

Nicola Ortelli^{*1,2}, Matthieu de Lapparent¹, and Michel Bierlaire²

¹School of Management and Engineering Vaud, HES-SO, Switzerland

²École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

SHORT SUMMARY

When estimating discrete choice models, the prospect of using ever-larger datasets is limited by the poor scalability of maximum likelihood estimation. This paper proposes a simple and fast dataset reduction method that is specifically designed to preserve the richness of observations originally present in a dataset, while reducing its size. Our approach leverages locality-sensitive hashing to create clusters of similar observations, from which representative observations are then sampled and weighted. We demonstrate the efficacy of our approach by applying it on a real-world mode choice dataset; the obtained results confirm that a carefully selected and weighted subsample of observations is capable of providing close-to-identical estimation results while being, by definition, less computationally demanding.

Keywords: discrete choice models, maximum likelihood estimation, dataset reduction, sample size, locality-sensitive hashing

1 INTRODUCTION

When estimating DCMs, the use of ever-larger datasets raises two issues: (i) the number of possible model specifications exponentially grows with the number of potential explanatory variables, implying that analysts must spend more time searching for good models; and (ii) the computational cost of maximum likelihood estimation increases with the number of observations, quickly becoming intractable even for basic model structures. While the first issue has spurred great interest (van Cranenburgh et al., 2021), the second has received much less attention: to deal with the increased computational cost associated with large datasets, effort has mostly been dedicated to improving the optimization methods used to estimate DCMs (Lederrey et al., 2021; Rodrigues, 2022) or to enhancing their implementation (Molloy et al., 2021; Arteaga et al., 2022).

This study explores an alternative approach, which consists in reducing the size of datasets by sampling their observations. Removing observations from a dataset is usually advised against by econometricians and choice modelers, but has nevertheless become common practice when training machine learning models on large amounts of data: given the iterative nature of model specification, the use of a smaller sample that provides good approximations of the model’s quality allows for early modeling decisions to be taken significantly faster (Park et al., 2019).

To the best of our knowledge, only two studies explore this same idea in the context of discrete choice modeling. The first one is van Cranenburgh & Bliemer (2019): their proposed method scales down any dataset to a predefined fraction of its original size while iteratively minimizing an estimate of the D -error, obtained by means of a simplified version of the model of interest. The second direct precedent of this study is Schmid et al. (2022), in which the authors use the k -means algorithm to identify clusters of similar observations and sample from those as a pre-processing step. Both methods are computationally heavy, which severely limits their usage.

We propose a simple and extremely fast dataset reduction technique that is designed to introduce as little bias as possible in the parameter estimates of the model of interest. Our approach leverages locality-sensitive hashing (LSH) to create clusters from which representative observations are sampled, similar to Schmid et al. (2022). The observations obtained in such way are then given weights that are proportional to the sizes of the clusters they represent, so as to mimic the full dataset during the model estimation process. As argued in the following sections, we believe that a carefully selected *and weighted* subsample of observations is capable of providing close-to-identical estimation results while being, by definition, less computationally demanding.

2 METHODOLOGY

Intuition

Consider a choice dataset of N observations (x_n, i_n) , each consisting of a vector x_n of explanatory variables associated with individual n , together with the observed choice i_n of that same individual among J alternatives. In its simplest form, a discrete choice model $P(i | x_n; \theta)$ calculates the probability that individual n chooses alternative i as a function of x_n and θ , where θ is a vector of model parameters to be estimated from the data.

The values of the model parameters are typically determined through maximum likelihood estimation, which consists in finding the values of θ that maximize the joint probability of replicating all observed choices in the dataset. In practice, the logarithm of the likelihood is usually maximized instead, for numerical reasons. The *log likelihood* function is therefore defined as

$$\mathcal{L}(\theta) = \sum_{n=1}^N \log P(i_n | x_n; \theta). \quad (1)$$

Let us now assume that the dataset contains some observations that are identical in all explanatory variables and in the observed choice. By partitioning the observations into $G < N$ groups of identical observations, we may rewrite (1) as

$$\mathcal{L}(\theta) = \sum_{g=1}^G N_g \cdot \log P(i_g | x_g; \theta), \quad (2)$$

where N_g denotes the size of group g and i_g and x_g are the observed choice and explanatory variables shared by all observations in group g , respectively. (1) and (2) are equivalent and, as such, yield the exact same parameter estimates when maximized. However, since $G < N$, the computational cost of evaluating (2) is smaller, by a ratio of approximately $\frac{G}{N}$.

The idea behind our dataset reduction method is to extend this “factorization trick” to observations that are nearly identical. The number of distinct groups is thereby further reduced and so is the computational time associated with evaluating the log likelihood function and its gradient. This comes at the cost of degrading the estimation results, because part of the information contained in the dataset is lost; still, the use of an adequate clustering scheme limits said degradation while granting our method a certain reliability. The clustering technique chosen for this purpose is locality-sensitive hashing (LSH), which we introduce now.

Locality-sensitive hashing

LSH is an efficient method for gathering similar data points into clusters — or *buckets*. It achieves this goal by combining the outcomes of several hashing functions, designed in such way that pairs of items are more likely to be hashed to the same bucket if they are close to each other in their original space than if they are far apart. A considerable advantage of LSH over other clustering techniques is that its computational complexity is linear in the number of items to be hashed.

A *family* of LSH functions $\mathcal{H} = \{h : (M, d) \rightarrow \mathbb{Z}\}$ is a collection of functions h that map elements of a metric space (M, d) onto the set of integers \mathbb{Z} (Leskovec et al., 2020). Each integer represents a different bucket, and two data points x_p and x_q belong to the same bucket of function h if and only if $h(x_p) = h(x_q)$. For instance, a well-known family of LSH functions is given by

$$h_{a,b}(x) = \left\lfloor \frac{a \cdot x + b}{w} \right\rfloor, \quad (3)$$

where $\lfloor \cdot \rfloor$ denotes the floor function, a is a vector whose entries are independently drawn from a normal distribution $\mathcal{N}(0, 1)$, b is a real value chosen uniformly from the range $[0, w)$ and w is the bucket width (Arnaiz-González et al., 2016). One may see (3) as a projection of all data points onto a line whose direction is given by vector a ; an offset equal to b is added to all projected points before the line is discretized into uniform intervals of size w . All data points that fall in the same interval are therefore assigned to the same bucket.

The value of w is context-dependent. By changing the bucket width, one can choose an appropriate degree of similarity within buckets: a sufficiently small w only groups points that are exactly identical, whereas greater values result in fewer buckets that contain larger amounts of increasingly dissimilar points.

Another way of improving the discriminative power of LSH is to combine several hash functions. In the case of the family defined by (3), suppose a and b are drawn R times: now, two data points x_p and x_q belong to the same bucket if and only if they are hashed together by all R functions, *i.e.*:

$$H_{A,B}(x_p) = H_{A,B}(x_q) \iff h_{a_r,b_r}(x_p) = h_{a_r,b_r}(x_q) \quad \forall r = 1, \dots, R, \quad (4)$$

where $A = (a_1, \dots, a_R)$ and $B = (b_1, \dots, b_R)$ gather the R realizations of a and b , respectively. Increasing R reduces the joint probability that two data points are grouped together by all projections.

LSH-based dataset reduction (LSH-DR)

Our dataset reduction algorithm has three ingredients: (i) an LSH function or a combination of LSH functions capable of partitioning a dataset of size N into buckets that contain similar observations; (ii) a sampling strategy, based on which some observations are selected from each bucket; and (iii) a weighting scheme that assigns a weight N_g to each selected observation (x_g, i_g) . The G observations obtained in such way, together with their associated weights N_1, \dots, N_G , constitute the outcome of our method. Any model of interest may then be estimated on the obtained subsample rather than on the whole dataset by using the log likelihood function of (2), with i_g and x_g now referring to the observed choice and explanatory variables associated with the g -th selected observation, respectively.

Clustering Our method uses the family of LSH functions introduced in (3) with, as parameters, the discretization step w and the number of projections R . It is crucial that prior to hashing, all explanatory variables are normalized such that their values are between 0 and 1. The individuals' choices are not taken into account during the clustering, which implies that the buckets might be heterogeneous, *i.e.*, observations with different chosen alternatives might end up in the same bucket.

Sampling The current version of our method randomly selects one observation from each alternative in each bucket.

Weighting Each selected observation (x_g, i_g) is given a weight N_g that is equal to the number of observations that share the same bucket $H_{A,B}(x_g)$ and the same chosen alternative i_g :

$$N_g = |\{(x_n, i_n) \mid H_{A,B}(x_g) = H_{A,B}(x_n), i_g = i_n\}|. \quad (5)$$

Jointly, the adopted sampling strategy and weighting scheme guarantee that the sum of all weights is equal to the number of observations in the full dataset.

3 RESULTS AND DISCUSSION

The efficacy of our method is demonstrated by means of a series of experiments based on the London passenger mode choice (LPMC) dataset (Hillel et al., 2018). The dataset consists of more than 81'000 trip records collected over three years. Four modes are distinguished: walk, cycle, ride public transport and drive. We divide the dataset into two parts: the first two years of data — 54'766 observations — are used for model estimation whilst the final year — 26'320 observations — is set aside for out-of-sample validation.

We use the data to train two multinomial logit models that we borrow from Hillel (2019). We refer to those as “MNL-S” and “MNL-L”. The former includes 10 continuous variables and 13 associated parameters, whereas the latter considers 11 continuous variables, 8 categorical variables encoded using binary indicators and 53 associated parameters. All model estimations are performed using the Biogeme package for Python (Bierlaire, 2018, 2020) on a 2.3 GHz 32-core cluster node with 192 GB of RAM.

Experiment A

We begin by estimating the MNL-S on samples generated by our proposed method. We apply the LSH-DR algorithm on the LPMC data 10'000 times, with $R = 4$ and w ranging from 0.02 to 0.2. The obtained samples range from 1'361 to 48'206 observations in size, that is, from 2% to

88% of the full dataset. The results are shown in Figure 1. We report the following quantities: (i) the execution time, which consists of the sampling and estimation times; (ii) the normalized out-of-sample log likelihood (OSLL), *i.e.*, the log likelihood yielded by the estimated model on the validation data, normalized by the number of observations; (iii) the mean absolute percentage error (MAPE) of the parameter estimates; and (iv) the value of time for the “drive” alternative, computed as the ratio between the estimates of the parameters associated with travel time and cost. For comparative purposes, we also report the results obtained on random samples. Figure 1 demonstrates that LSH-DR is capable of producing substantially better samples than random sampling, for a negligible increase in execution time: down to approximately 40% of the full dataset size, the samples generated by LSH-DR yield smaller MAPEs of the parameters and more accurate estimates of the value of time.

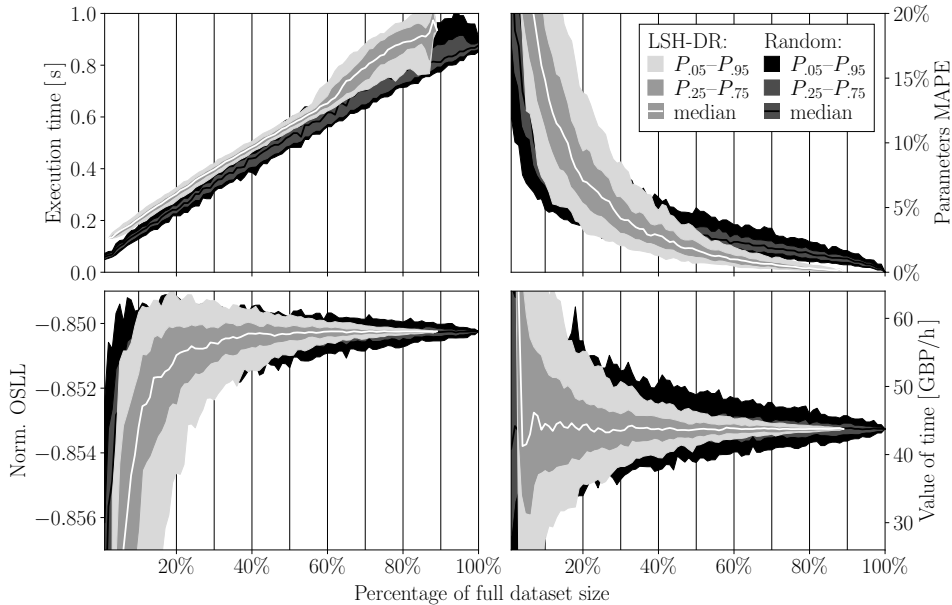


Figure 1: Estimation of the MNL-S on samples generated by LSH-DR.

Experiment B

In this experiment, we compare the performance of our method with three other dataset reduction techniques, namely: (i) random sampling; (ii) k -means clustering, similar to the approach taken in Schmid et al. (2022); and (iii) sampling of observations (SoO), as proposed by van Cranenburgh & Bliemer (2019). We proceed as follows: a certain percentage of the full dataset size is chosen and we retrieve from Experiment-B the 100 samples of size closest to that percentage; the three other dataset reduction techniques are then used to generate samples of those exact same sizes, which are finally used to train the MNL-S model. Figure 2 reports the sampling time, normalized OSLL, parameters MAPE and value of time for the “drive” alternative for 25%, 50% and 75% of the full dataset. The boxplot whiskers indicate the 5th and 95th percentiles. The size of the samples retrieved from Experiment B ranges from 13’480 to 13’984, from 27’146 to 27’615, and from 40’879 to 41’273, respectively.

Overall, Figure 2 illustrates that the samples producing the most accurate results are obtained via k -means. Still, despite its superiority, k -means is practically unusable because of its runtime: it takes from 8 up to 26 minutes to obtain a sample from a relatively small dataset. That is between 4’000 and 16’000 times longer than LSH-DR and up to 400’000 times longer than random sampling. As regards SoO, the method is shown to provide the worst results, while also displaying the largest runtimes. This is due to the fact that SoO is designed to maximize the efficiency of the parameter estimates rather than their precision or the model’s predictive accuracy.

Experiment C

Lastly, we estimate the MNL-L model on samples generated by LSH-DR. To this end, we apply the LSH-DR algorithm on the LPMC data 10’000 times, with $R = 4$ and w ranging from 0.1 to

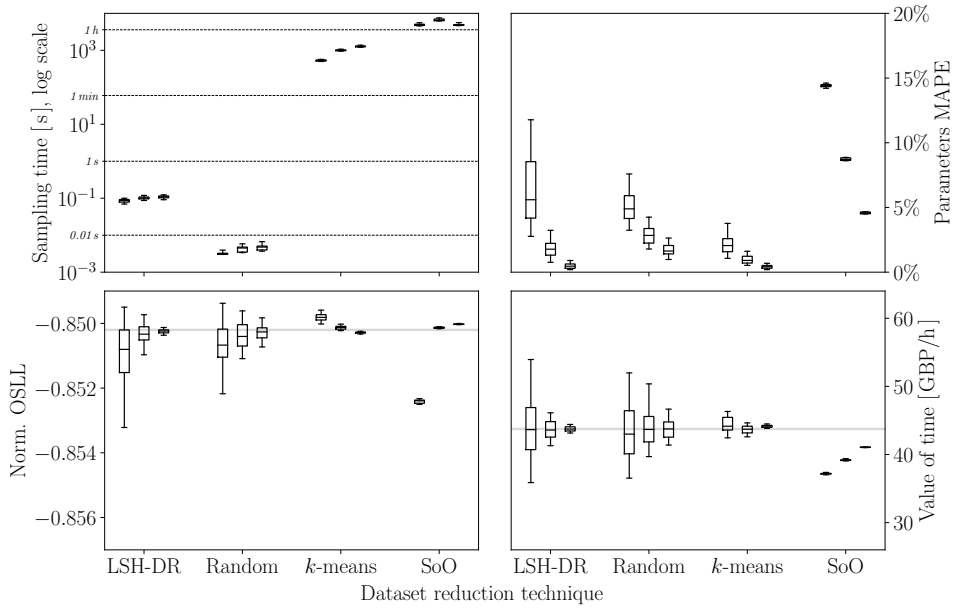


Figure 2: Comparison of dataset reduction techniques.

1. Note that the MNL-L model includes several discrete explanatory variables, but those are not treated differently by the LSH-DR method. The generated samples range from 3'584 to 51'574 observations in size, that is, from 7% to 94% of the full dataset. Figure 3 displays the achieved results. For the sake of comparison, the results obtained on random samples are also shown. Figure 3 demonstrates that our method may also be beneficial to larger models. Overall the improvement in comparison to random sampling may be less remarkable than with the smaller model, but it is worth noting that the MAPE remains within a reasonable level of accuracy even for samples down to 50% of the full dataset size.

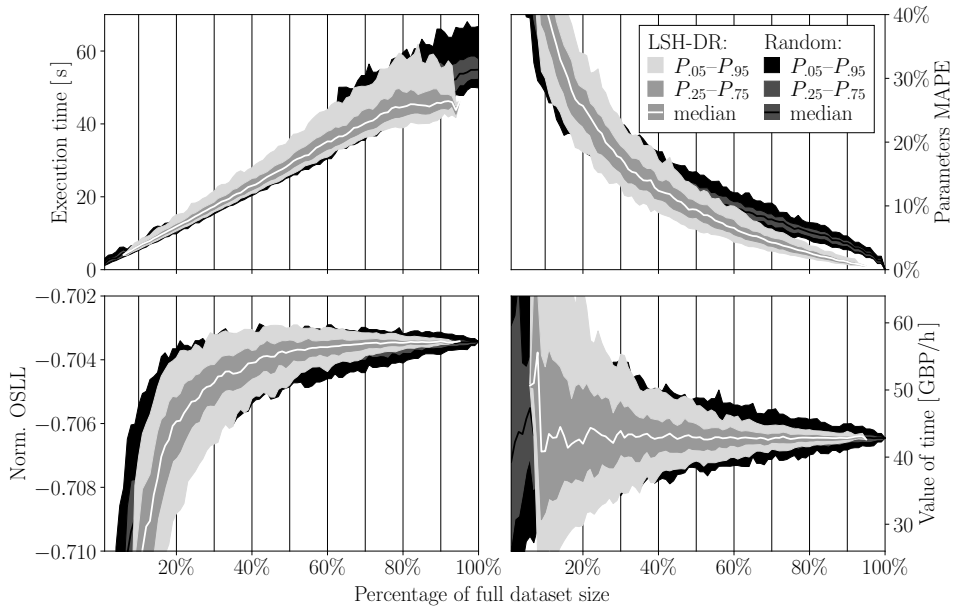


Figure 3: Estimation of the MNL-L on samples generated by LSH-DR.

4 CONCLUSIONS

In this paper, we propose a simple and fast dataset reduction technique that speeds up the estimation of discrete choice models. The gain in computational time naturally comes at the cost of

deteriorating the model estimation results; however, our method is specifically designed to mitigate this deterioration by preserving as much diversity as possible among the observations. As a result, the quality of the parameter estimates stays within reasonable ranges even for large reduction rates. The presented results additionally highlight the benefits of our method on the estimation of models of small and medium size.

Intended future work includes the development and testing of more elaborate sampling strategies for selecting observations from buckets. For instance, those could be designed to increase the probability of choosing the most representative observations within each bucket, or to rely on the content of each bucket to generate synthetic prototypical observations. Additional investigation could also consist in developing LSH functions that can accommodate the analyst’s knowledge of the dataset or the structure of the model of interest, for instance by giving more importance to some specific variables during the hashing. Finally, another promising direction of research consists in embedding the LSH-DR method within a stochastic gradient descent algorithm for model estimation, such as the one proposed in Lederrey et al. (2021). The use of carefully selected and weighted batches of data, rather than random ones, could result in significantly better approximates of the gradient and, as a result, speed up convergence.

REFERENCES

- Arnaiz-González, Á., Díez-Pastor, J.-F., Rodríguez, J. J., & García-Osorio, C. (2016). Instance selection of linear complexity for big data. *Knowledge-Based Systems*, 107, 83–95.
- Arteaga, C., Park, J., Beeramoole, P. B., & Paz, A. (2022). xlogit: An open-source Python package for GPU-accelerated estimation of mixed logit models. *Journal of Choice Modelling*, 42, 100339.
- Bierlaire, M. (2018). *PandasBiogeme: a short introduction* (Tech. Rep.). TRANSP-OR 181219. Transport and Mobility Laboratory, ENAC, EPFL.
- Bierlaire, M. (2020). *A short introduction to pandasBiogeme* (Tech. Rep.). TRANSP-OR 200605. Transport and Mobility Laboratory, ENAC, EPFL.
- Hillel, T. (2019). *Understanding travel mode choice: A new approach for city scale simulation* (Unpublished doctoral dissertation). University of Cambridge.
- Hillel, T., Elshafie, M. Z., & Jin, Y. (2018). Recreating passenger mode choice-sets for transport simulation: A case study of London, UK. *Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction*, 171(1), 29–42.
- Lederrey, G., Lurkin, V., Hillel, T., & Bierlaire, M. (2021). Estimation of discrete choice models with hybrid stochastic adaptive batch size algorithms. *Journal of choice modelling*, 38, 100226.
- Leskovec, J., Rajaraman, A., & Ullman, J. D. (2020). *Mining of massive data sets*. Cambridge university press.
- Molloy, J., Becker, F., Schmid, B., & Axhausen, K. W. (2021). mixl: An open-source R package for estimating complex choice models on large datasets. *Journal of choice modelling*, 39, 100284.
- Park, Y., Qing, J., Shen, X., & Mozafari, B. (2019). BlinkML: Efficient maximum likelihood estimation with probabilistic guarantees. In *Proceedings of the 2019 international conference on management of data* (pp. 1135–1152).
- Rodrigues, F. (2022). Scaling bayesian inference of mixed multinomial logit models to large datasets. *Transportation research part B: methodological*, 158, 1–17.
- Schmid, B., Becker, F., Molloy, J., Axhausen, K. W., Lüdering, J., Hagen, J., & Blome, A. (2022). Modeling train route decisions during track works. *Journal of Rail Transport Planning & Management*, 22, 100320.
- van Cranenburgh, S., & Bliemer, M. C. (2019). Information theoretic-based sampling of observations. *Journal of choice modelling*, 31, 181–197.
- van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., & Walker, J. (2021). Choice modelling in the age of machine learning-discussion paper. *Journal of Choice Modelling*, 100340.

Adaptive Traffic Signal Control: A Novel Modeling Approach

Lubing Li^{*1}, Ka Fai Ng¹, Jacob Chak Yan Lo¹, and Hong K. Lo¹

¹ Department of Civil and Environmental Engineering
The Hong Kong University of Science and Technology, Hong Kong, China

SHORT SUMMARY

We develop a novel cell-based two-stage stochastic program to address spatial, dynamic and stochastic features of traffic flow for adaptive signal control. Cell transmission model (CTM) is employed to capture the dynamic feature of traffic flow, with certain CTM cells designated as detector cells to capture real-time spatial queuing effects. We formulate a two-stage stochastic program to address uncertain demand for signal control. In stage 1, a base timing plan (BTP) is determined as the long-term default plan. In stage 2, cycle-based adaptive policies, i.e., green extension/cut-off based on the BTP, are implemented according to the detector cell states. We develop a specialised GA algorithm to search for the optimal BTP and adaptive policies. A case study of Tai Tam reservoir is conducted to elaborate the property of the proposed approach. The adaptive control plan can have 17% delay reduction compared to the optimal fixed-time plan.

Keywords: adaptive signal control, traffic flow theory, two-stage stochastic program

1. INTRODUCTION

To develop a demand-responsive traffic signal control system, traffic detectors play a vital role in collecting traffic information for making signal control decisions. For vehicle actuated signal control, detectors are deployed near the stop lines of all or some of the junction approaches. According to the vehicle detection information, the green times of certain stages are extended or cut off. The question of how long the green time ought to be extended or cut off is not addressed explicitly. The extension/cut-off durations are generally determined according to traffic engineer experiences in practice. Urban traffic control (UTC) signal control system generally produces fixed-time plans for each identified traffic pattern based on detector information. Appropriate timing plans are selected by the time of day, day of the week, or special event situations. Once selected, the fixed-time plan is implemented for that period. Albeit detectors provide seconds/minutes level sensing data, the traffic variations within the period are neglected. For adaptive signal control systems, such as SCOOT (Robertson and Bretherton, 1991), detectors are deployed at the entry to the network links to predict the time and shape of the flow profiles at the stop line. Optimal signal timing variables are calculated with the incremental change algorithm. The optimality of the signal timing plans may be hampered by the restricted computational time budget. Because of platoon dispersion and lane changing effects, the estimated flow profiles are likely biased toward the arrival patterns at the stop line. In this paper, we utilize the vehicle detection data from first principles and incorporate dynamic detector status information into a mathematical model to derive the optimal adaptive control policies.

Our study aims to address three characteristics of traffic flow, i.e., spatial, dynamic, and stochastic, for adaptive signal control. To capture spatial and dynamic features, we adopt the cell transmission model (CTM) (Daganzo, 1994) as the underlying traffic flow model. Certain cells in the CTM network are selected as the detector cells. Cell occupancy information is collected to

indicate queue presence status at the detector locations. To capture the stochastic traffic demand feature, we adopt a two-stage stochastic program method. In stage 1, a base timing plan, including cycle time, base green time, and offset, is determined as the long-term default plan or ‘fallback’ plan. In stage 2, upon the traffic demand realization scenarios, cycle-based adaptive control policies are determined to address residual queues. At the end of each signal cycle, adaptive policies are selected according to the status of the detector cells. Each detector status pattern corresponds to a certain adaptive policy. The selected adaptive policy will be implemented in the next cycle. We develop a specialized GA algorithm to determine the optimal base timing plan and adaptive policies. A case study at Tai Tam reservoir is provided to indicate the properties of the proposed adaptive signal control method.

2. METHODOLOGY

signal control in CTM

As shown in (1), the green start time $S_{i,j,z}$ of phase-1 ($i=1$) is equal to the initial offset O_j^{ini} at junction j plus the lost time L plus the z^{th} cycle start time $(z-1) \cdot C$. For other phases ($i>1$), the green start time $S_{i,j,z}$ is equal to the lost time L plus the green end time of the predecessor phase $E_{i-1,j,z}$, as shown in (2). The green end time $E_{i,j,z}$ is equal to the green start time $S_{i,j,z}$ plus the green duration $G_{i,j}$, as shown in (3). The signal effects can be captured in CTM by adjusting the inflow capacity of the signal cells (M_l) according to the state of signal light. In (4), when the time step t is between $S_{i,j,z}$ and $E_{i,j,z}$ (in green), the inflow capacity of the signal cell, i.e., $Q_{m,l}(t)$, is set to be the saturation flow rate s ; otherwise $Q_{m,l}(t)$ equals zero.

$$S_{i,j,z} = O_j^{ini} + L + (z-1) \cdot C, i=1, \forall j, z, \quad (1)$$

$$S_{i,j,z} = L + E_{i-1,j,z}, i > 1, \forall j, z, \quad (2)$$

$$E_{i,j,z} = S_{i,j,z} + G_{i,j}, \forall i, j, z, \quad (3)$$

$$Q_{m,l}(t) = \begin{cases} s, & S_{i,j,z} < t \leq E_{i,j,z}, \forall i, j, z, m \in M_1. \\ 0, & otherwise \end{cases} \quad (4)$$

detector cell settings

We designate certain CTM cells as detector cells to mimic real-time traffic detector deployment. Adaptive signal control policies are triggered based on detector states. We employ indicator functions in (5) to represent whether the state of a detector is active or inactive. Denote $d_{i,j}(t)$ as the state of detector cell at time t for phase i , junction j . Denote set $\Phi_t = \{d_{i,j}(t), \forall i \in I', \forall j \in J'\}$ as the set that includes all the detector cell states in the target network. I' is the phase set, for which phase detectors are deployed. J' is the junction set, for which junction detectors are deployed. In (5), if the occupancy of detector cell $c^{d_{ij}}$ is greater than or equal to the critical occupancy κ_f at time t , $d_{i,j}(t) = 1$; otherwise, $d_{i,j}(t) = 0$. (5) can be applied to represent queue presence detection. For instance, if the occupancy of detector cell is full or almost full, e.g., $\kappa_f > 0.9$, vehicle queue is regarded as reaching to the detector location.

$$d_{i,j}(t) = \begin{cases} 1, & occ_t(c^{d_{ij}}) \geq \kappa_f \\ 0, & occ_t(c^{d_{ij}}) < \kappa_f \end{cases} \quad (5)$$

two-stage signal control formulation

With the CTM model embedded, we formulate a cell-based two-stage stochastic program to optimize the traffic signal control plan (Li et al., 2018; Li et al., 2021). In stage 1, the expected total delay $E_{\Omega}[D(C, \mathbf{G}^b, \mathbf{O}^{ini})]$ is minimized with respect to the base timing plan, i.e., $\{C, \mathbf{G}^b, \mathbf{O}^{ini}\}$, shown in (6). It is subject to CTM dynamics constraints. (7)-(10) are constraints for signal control variables.

In stage 2, given the base timing plan $\{C, \mathbf{G}^b, \mathbf{O}^{ini}\}$ from stage 1, adaptive policies, $\Delta \mathbf{G}$, hence, the actual green time \mathbf{G}^a , are determined to minimize the total delay D_k upon the realization of traffic demand $\hat{\mathbf{v}}_k$ in scenario k , as shown in (11). $\Delta \mathbf{G}$ is the set that includes all the adaptive policies, i.e., $\Delta \mathbf{G} = \{\{\Delta G_{i,j}^c, \forall i \in I, \forall j \in J\}, \forall c\}$. $\Delta G_{i,j}^c$ is the adaptive policy c for phase i at junction j . Denote \mathcal{A} as the set that includes all the detector state combinations, i.e., $\mathcal{A} = \{\{d_{i,j}^c, \forall i \in I, \forall j \in J\}, \forall c\}$. $d_{i,j}^c$ is the state of detector (either 1 or 0, as shown in (5)) for adaptive policy c for phase i at junction j . Each adaptive policy c corresponds to a specific detector state combination, i.e., $\Delta \mathbf{G}(c) \rightarrow \mathcal{A}(c)$. For adaptive policy c , the summation of ΔG^c over all phases for junction j should be zero to maintain the common cycle time, as shown in (12). The adaptive policy ΔG^c is selected according to the state of detector cells Φ . If the detector state combinations for adaptive policy c , i.e., $\mathcal{A}(c)$, are the same as the state of detector cells, i.e., $\mathcal{A}(c) = \Phi$, indicator function $1_{\Phi}(\mathcal{A}(c))$ returns 1, as shown in (13); otherwise, $1_{\Phi}(\mathcal{A}(c)) = 0$. In (14), the actual green time $G_{i,j}^a$ for phase i junction j equals the base green time $G_{i,j}^b$ plus the selected adaptive policy $\Delta G_{i,j}^c$. (15) is the range of actual green time. (16) is the actual green end time constraint.

(Stage 1:)

$$\min_{C, \mathbf{G}^b, \mathbf{O}^{ini}} E_{\Omega}[D(C, \mathbf{G}^b, \mathbf{O}^{ini})] = \sum_{k \in \Omega} P_k(\hat{\mathbf{v}}_k) \cdot D_k(C, \mathbf{G}^b, \mathbf{O}^{ini}, \hat{\mathbf{v}}_k) \quad (6)$$

subject to CTM dynamics constraints and (1)-(4), and

$$\sum_i (G_{i,j}^b + L) = C, \quad \forall j, \quad (7)$$

$$G_{i,j}^{min} \leq G_{i,j}^b \leq G_{i,j}^{max}, \quad \forall i, j, \quad (8)$$

$$C^{min} \leq C \leq C^{max}, \quad (9)$$

$$0 \leq O_j^{ini} \leq O_j^{ini,max}, \quad \forall j, \quad (10)$$

where $D_k(C, \mathbf{G}^b, \mathbf{O}^{ini}, \hat{\mathbf{v}}_k)$ is the total delay value in demand scenario k in stage-2.

(Stage 2:)

$$\min_{\Delta \mathbf{G}, \mathbf{G}^a} D_k(\Delta \mathbf{G}, \mathbf{G}^a | \hat{\mathbf{v}}_k) = \sum_l \sum_t d_{l,k}(t) \quad (11)$$

subject to CTM dynamics constraints and (1)-(2), (4)-(5) and

$$\sum_i \Delta G_{i,j}^c = 0, \forall j, c \quad (12)$$

$$\mathbf{1}_\Phi(\mathbf{A}(c)) = \begin{cases} 1, & \mathbf{A}(c) = \Phi_t \\ 0, & \mathbf{A}(c) \neq \Phi_t \end{cases}, t = n \cdot C, n \in \mathbb{N}; \quad (13)$$

$$G_{i,j}^a = G_{i,j}^b + \sum_c \mathbf{1}_\Phi(\mathbf{A}(c)) \cdot \Delta G_{i,j}^c, \forall i, j \quad (14)$$

$$G_{i,j}^{\min} \leq G_{i,j}^a \leq G_{i,j}^{\max}, \forall i, j, \quad (15)$$

$$E_{i,j,z} = S_{i,j,z} + G_{i,j}^a, \forall i, j, z \quad (16)$$

The optimal total delay D_k in stage 2 is accrued into the objective function to obtain the expected total delay $E_\Omega[D(C, \mathbf{G}^b, \mathbf{O}^{ini})]$. As shown in (6), the expected total delay $E_\Omega[D(C, \mathbf{G}^b, \mathbf{O}^{ini})]$ equals the summation of the total delay over each demand scenario k ($k \in \Omega$). Ω is the space of the realizations of demand scenario. $P_k(\hat{\mathbf{v}}_k)$ is the probability of scenario k for which the realized demand is $\hat{\mathbf{v}}_k$. We assume a finite number of demand scenarios sampled in space Ω and the summation of probability P_k equals 1, i.e., $\sum_{k \in \Omega} P_k = 1$. The whole optimization problem is to determine the optimal solutions for the base timing plan in stage 1 and the set of adaptive policies in stage 2, which will minimize the expected total delay under stochastic demand condition.

solution algorithm

In this study, we develop a specialised solution approach based on genetic algorithm (GA) to solve the proposed adaptive signal control problem. We have introduced GA into solving signal control optimisation problem and elaborated the effectiveness of GA in generating quasi-optimal signal control solutions (Lo et al., 2001; Lo and Chow, 2004). In GA, signal control variables are transformed into 0-1 binary representations. The duration of each variable is coded as a series of chromosomes. An example of the gene structure of the cycle-based two-stage signal control is shown in Figure 1. The head part is the common cycle time for the junctions in the network, following with the control variables of each junction. Junction control variables consists of the base timing plan (including initial offset O_j^{ini} , base green time $G_{i,j}^b$) and adaptive policies $G_{i,j}^a$.

The actual green time \mathbf{G}^a is coded as the adaptive decision variables in the gene since the range of $\Delta \mathbf{G}$ includes negative domain which is not operable in GA. The population of genes are manipulated with three operators, i.e., reproduction, crossover, mutation, to search for optimal signal control solutions. The gene (signal plans) with the best delay performance will be selected as the optimal solution.

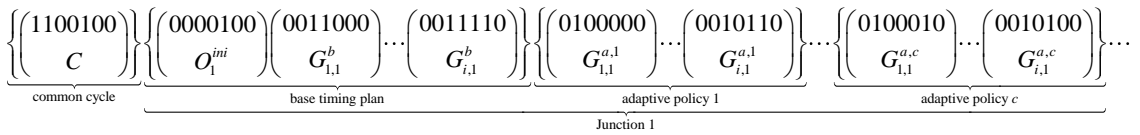


Figure 1. Gene structure of the cycle-based two-stage signal control

3. RESULTS AND DISCUSSION

We select Tai Tam reservoir as the test site to illustrate the performance of the proposed adaptive signal control method. Tai Tam reservoir road is a section of Tai Tam road located at eastern

Hong Kong Island, which is a bi-directional two-lane road connecting Stanley and Chai Wan, as shown in Figure 2. Since the construction of the reservoir was completed in 1918, the width of the road is too narrow to accommodate bi-directional traffic passing through the reservoir simultaneously. A two-phase signal is deployed on the reservoir. Only one direction of traffic is allowed going through per phase. The vehicle clearance time (all red time) of 45 seconds is set for every phase switch. Due to the long all red time, frequent phase switches may incur severe delay deterioration. The key question is how to effectively assign green time of the two phases so that the total delay is minimised under uneven traffic arrivals of the two directions. We apply the data set collected in three days (June, 2019) for signal optimisation. The data set of each day includes three-hour PM peak bi-direction traffic arrivals to the Tai Tam reservoir. We code the CTM network for Tai Tam reservoir as shown in Figure 2. The parameters of the CTM network are calibrated with the real data.

We assign northbound traffic (to Chai Wan) to phase 1 and southbound traffic (to Stanley) to phase 2. The minimum phase duration is 60 seconds. We first optimise an optimal fixed-time plan for the reservoir via the conventional GA (the gene structure without the adaptive policy chromosomes). The optimisation is performed on Intel i7-3370 computer with 32GB RAM. The parameters of the GA include 10 generations, 100 population size. The expected total delay over three demand scenarios is minimised with respect to signal control variables. The optimal fixed-time signal plan is listed in Table 1.

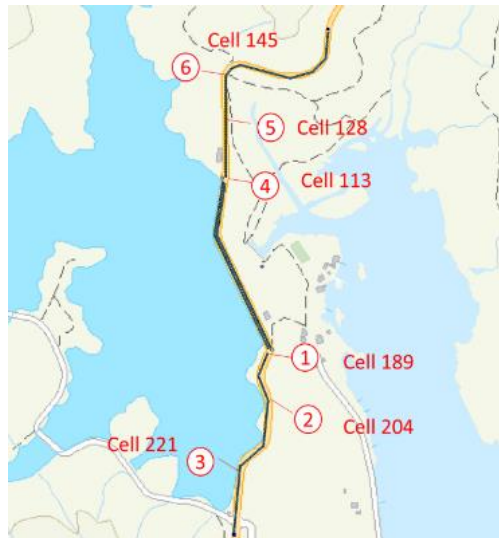


Figure 2. Tai Tam reservoir

Table 1. fixed-time signal plan

	Cycle time (s)	Offset (s)	Phase 1 (s)	Phase 2 (s)
duration	180	130	97	83

For the queue-based adaptive signal control. We deploy three detector cells on each side of the reservoir approaches as shown in Figure 2, (The detector cell ID are marked). The locations of the detector cells are 0 m, 100 m, 200 m, respectively, from the stop lines for both sides. If vehicle queue end is detected at the detector cells (cell occupancy > 90%), the detector state turns 1; 0 otherwise. We adopt the proposed specialised GA algorithm to optimise both the base timing plan in stage 1 and the adaptive control policies in stage 2. Since we deploy six detectors for Tai Tam case, there are $2^6 = 64$ detector state combinations, hence, 64 adaptive policies for adaptive signal control. Table 2 lists all the adaptive polices and the corresponding detector patterns. Detector

patterns are counted during the GA simulations. Detector pattern ‘100100’ (policy 37) is the most frequently activated pattern which is counted 50435 times in the simulations. ‘100100’ means that at the end of signal cycle, i.e., end of phase 2, vehicle queues are detected at detector 1 and detector 4 (close to the stop lines), but not at the other detectors. It is noted that only 12 detector patterns are activated and the rest of the detector patterns are never activated. Those non-critical adaptive policies can be eliminated from the gene structure. We further optimise the top 9 frequently activated adaptive policies and the base timing plan via GA. Whenever the detector states not included in the top 9 adaptive policies appear, the base plan will be implemented in the next cycle. The top 9 adaptive policies and the base plan are listed in Table 3.

The base green times are 89 seconds for phase 1 and 91 seconds for phase 2. For pattern ‘100100’ (policy 37), it indicates that there is overflow at the SB approach at the end of phase 2 and vehicle queue at the NB approach. The adaptive plan is to extend 4 sec green time for phase 1 and cut off 4 sec for phase 2 in the next cycle. For pattern ‘111000’ (policy 57), long queue is detected at the NB approach but no queue at the SB approach. The adaptive plan is to extend 27 sec green time for phase 1 to clear the long queue. For pattern ‘100110’ (policy 39), a long residual queue is detected at the end of phase 2. The adaptive plan is to cut off 10 sec green time from phase 1 and extend 10 sec green time for phase 2. The average delay results are listed in Table 4. The average delay of the optimal fixed-time plan is 105 sec/veh. The adaptive plan (86.5 sec/veh) can have a further 17.6% delay reduction compared to the optimal fixed-time plan. Even though some of the adaptive policies are eliminated from the GA optimisation and only top 9 adaptive policies are optimised, the adaptive plans can still have 17.0% delay improvement, which does not deteriorate too much compared to the full policy adaptive plans.

Table 2. adaptive policies and the corresponding detector patterns

adaptive policy		detector patterns	activated patterns	adaptive policy		detector patterns	activated patterns
<i>c</i>		‘123456’	counts	<i>c</i>		‘123456’	counts
1		‘000000’	1493	33		‘100000’	34308
2		‘000001’	0	34		‘100001’	0
3		‘000010’	0	35		‘100010’	0
4		‘000011’	0	36		‘100011’	0
5		‘000100’	1089	37		‘100100’	50435
6		‘000101’	0	38		‘100101’	0
7		‘000110’	59	39		‘100110’	1786
8		‘000111’	0	40		‘100111’	0
9		‘001000’	0	41		‘101000’	0
10		‘001001’	0	42		‘101001’	0
11		‘001010’	0	43		‘101010’	0
12		‘001011’	0	44		‘101011’	0
13		‘001100’	0	45		‘101100’	0
14		‘001101’	0	46		‘101101’	0
15		‘001110’	0	47		‘101110’	0
16		‘001111’	0	48		‘101111’	0
17		‘010000’	0	49		‘110000’	15172
18		‘010001’	0	50		‘110001’	0
19		‘010010’	0	51		‘110010’	0
20		‘010011’	0	52		‘110011’	0
21		‘010100’	0	53		‘110100’	27584
22		‘010101’	0	54		‘110101’	0
23		‘010110’	0	55		‘110110’	3099
24		‘010111’	0	56		‘110111’	0

25	'011000'	0	57	'111000'	22707
26	'011001'	0	58	'111001'	0
27	'011010'	0	59	'111010'	0
28	'011011'	0	60	'111011'	0
29	'011100'	0	61	'111100'	20344
30	'011101'	0	62	'111101'	0
31	'011110'	0	63	'111110'	9520
32	'011111'	0	64	'111111'	0

Table 3. optimal base timing plan and adaptive signal plans

adaptive policy c	detector patterns '123456'	Phase-1 (sec)	ΔG_1	Phase-2 (sec)	ΔG_2
37	'100100'	93	4	87	-4
33	'100000'	96	7	84	-7
53	'110100'	93	4	87	-4
57	'111000'	116	27	64	-27
61	'111100'	104	15	76	-15
49	'110000'	111	22	69	-22
63	'111110'	89	0	91	0
55	'110110'	87	-2	93	2
39	'100110'	79	-10	101	10
Base plan (0)	--	89	0	91	0

Table 4. signal plans delay performance

	fixed-time plan	adaptive plan (full policy)	adaptive plan (partial policy)
Delay (sec/veh)	105.0	86.5	87.1
reduction	--	17.6%	17.0%

4. CONCLUSIONS

In this study, we developed a novel approach to address dynamic, spatial and stochastic characteristics of traffic flow for adaptive signal control. Our approach was able to map adaptive signal control policies to dynamic traffic detector information and address the closed-loop signal control strategy with a mathematical model. Cell transmission model was adopted as the underlying dynamic traffic flow model. Certain CTM cells were designated as detector cells to indicate the queue presence information. We developed a two-stage stochastic program to address uncertain demand for signal control. In stage 1, the base timing plan is determined as a long-term default plan. In stage 2, various demand scenarios are loaded to CTM to simulate detector cell state patterns and the corresponding adaptive control policies. The case study at Tai Tam reservoir illustrated the performance of the approach. At the end of a signal cycle, we examined the detector states to identify the queuing status. The green extension/cut-off policies were re-adjusted in response to the unbalanced queues for the next cycle. The preliminary simulation results showed that a further 17% delay reduction can be achieved by the adaptive control policies. In the future study, we will further investigate a stage-based adaptive control strategy, i.e., the adaptive policy is adjusted at the end of each signal stage according to the detector states. We anticipate that the stage-based control strategy will exhibit more flexible control impacts to the traffic flow pattern variations.

ACKNOWLEDGEMENTS

The study is supported by General Research Fund #16204518 of the Research Grants Council, and Smart Traffic Fund STF22EG01 of the HKSAR Government.

REFERENCES

Daganzo, C.F. 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4), 269-287.

Li, L., Huang, W., Lo, H.K. 2018. Adaptive Coordinated Traffic Control for Stochastic Demand. *Transportation Research Part C: Emerging Technologies*, 88, 31-51.

Li, L., Huang, W., Chow, A. H., & Lo, H. K. 2021. Two-stage stochastic program for dynamic coordinated traffic control under demand uncertainty. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 12966-12976.

Lo, H.K., Chang, E., Chan, Y.C. 2001. Dynamic network traffic control. *Transportation Research Part A: Policy and Practice*, 35(8), 721-744.

Lo, H.K., Chow, A.H.F. 2004. Control Strategies for Oversaturated Traffic. *Journal of Transportation Engineering*, 130(4), 466-478.

Robertson, D. I., & Bretherton, R. D. 1991. Optimizing networks of traffic signals in real time-the SCOOT method. *IEEE Transactions on vehicular technology*, 40(1), 11-15.

Charging demand for the unserved — an agent-based model approach

F. Hipolito^{*1}, J. Rich², and Peter Bach Andersen³

¹Postdoc, Department of Management Engineering, Technical University of Denmark, DTU,
Denmark Department, University, Country

²Professor, Department of Management Engineering, Technical University of Denmark, DTU,
Denmark Department, University, Country

³Senior Researcher, Department of Wind and Energy Systems, Technical University of Denmark,
DTU, Denmark Department, University, Country

SHORT SUMMARY

Deploying charging infrastructure requires identifying the most effective placement and size of charging facilities. This is particularly challenging when electric vehicles (EVs) are gradually introduced, as it creates a dynamic target that must be met to ensure successful adoption of EVs. This paper introduces an agent-based simulation model that tracks movements of EVs in space and time. Our model is based on a choice model of charging behaviour, which is integrated with non-parametric queues and information-sharing of waiting time. Our simulation captures demand resulting from choice of charging and the unserved demand represented as charging intentions that were not met. It is demonstrated that unserved demand varies over time and across locations, and that it can be greatly reduced by our information-sharing strategy. The model is applied to Copenhagen where we examine changes in charging infrastructure requirements between 2021-2030 when going from EV shares of 2% to 30%.

Keywords: agent-based simulation, censored demand, electric vehicles, information-sharing

1 INTRODUCTION

The recent worldwide uptake of EVs has led to an increasing interest for the EV charging situation and the design of the corresponding charging infrastructure.

Many papers have considered the placement and sizing problem of chargers from an operational research perspective (Hengsong Wang et al., 2010; Kuby & Lim, 125). Included in these types of models are the maximum set covering models as considered in Wang & Lin (2009) and Wang & Wang (2010). A common challenge, however, when applying such models, is that it requires simplifications when stating the demand-side. As demonstrated in Luo & Qiu (2020) and Kavianipour et al. (2021), it often involves use of simple queuing systems, limited heterogeneity, and few or no interactions between users.

An alternative and increasingly popular approach, however, is the use of agent-based simulation models to explore heterogeneous inputs and the consequences it has for different types of users. Specifically, the impact of flexible arrivals, vehicle heterogeneity (e.g. battery range and charging speed) and different driving patterns. This is the approach taken here, as will be discussed in more detail below.

The challenge

When planning a charging infrastructure for future years, several non-trivial challenges exist. Before considering the solution approach, it is timely to dive into some of these challenges.

A general problem is the prediction problem of charging demand in future years. This is complicated because it depends on many heterogeneous inputs, which in itself are difficult to project. Some inputs, such as battery range and charging speed, are dependent on technological innovations and buying preferences of users. Other input's, such as car ownership and population growth, depend on the income development as well as the urban development. Because of the accumulated uncertainty from all of these inputs, the idea that a single unified solution exists renders itself useless. Rather, the solution to the problem should be seen as a distribution of solutions over a parameter

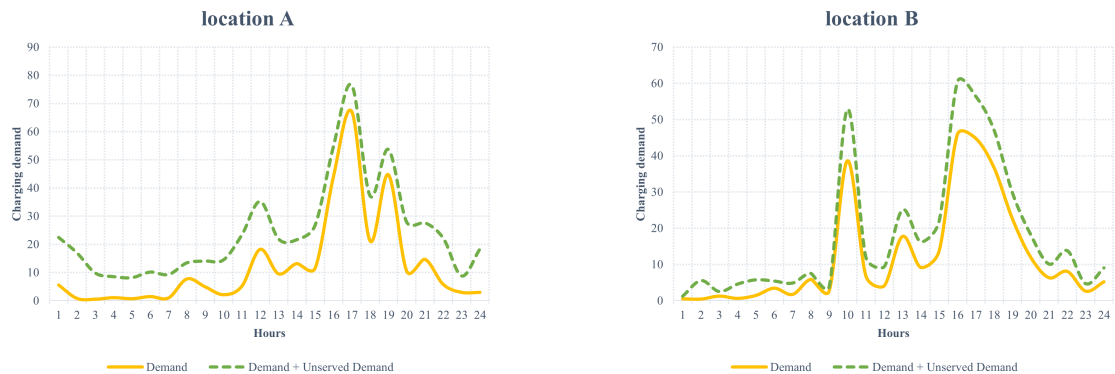


Figure 1: Unserved and served charging demand at different locations.

space, where some solutions are more likely than others. This observation, in our view, suggests that the agent-based simulation approach is the more suitable solution approach.

Another challenge, which exists for all capacity-constrained demand problems, is that supply deficiencies will cause users to suppress demand. This is particularly relevant in the context of charging because unserved charging demand will fluctuate in space and in time. An attempt is made to illustrate this in Figure 1 below.

The figure illustrates several things. Firstly, that the share of unserved demand varies in space and time. This suggests that from a charger location perspective, it is not only relevant to look at charging locations from the sole perspective of observed demand. Unserved demand should also be taken into account as it may influence the business case for operators and waiting time observed on the user side. Hence, a charging infrastructure that is designed with the unserved demand in mind, might look differently from a charging infrastructure solely based on what can be observed. It follows from this that the share of unserved demand can be considered as an important performance metric. Essentially, if measurable, the gap represents important information regarding the spatial and temporal match between demand and supply. The relevance of this is particularly useful when considering the gradual integration of EVs into the vehicle fleet over an extended period. This is because the balance between demand and supply presents itself as a moving target (in space and time) where insufficient supply may hinder EV adaptation in certain areas. As a result, it is highly important to monitor this balance over time and make sure that the inconvenience resulting from charging is minimal.

Sketch of a solution

Following the challenges as described in Section 1, it is relevant to delve into possible methodologies that allow consistent monitoring of unserved demand.

A requirement is that the modelling framework is agent-based, where utility functions and choice probabilities are monitored in space and time as a function of the different event triggers, allowing the systematic monitoring of unserved demand. Figure 2 attempts to illustrate the trajectories of a hypothetical agent’s diary over time, driven distance and SoC level. The solid black line represents the unperturbed diary path, i.e. no charging, whereas the coloured lines indicate some options to charge during the morning and afternoon activities. The colour coding indicates the ordering of the probabilities for each option, from yellow being the most likely to dark purple the least. The details regarding the evaluation of the utility of each option and decision process are addressed in §2, while the agent generation of the process is discussed in §2.

Realistic monitoring of the state-of-charge in the space-time domain as well as the waiting time, requires, on the one hand realistic arrival patterns among users, but also flexibility resulting from queuing spill-back at the level of chargers. Specifically, for users to realistically react to queuing dynamics and trade different charging options to circumvent waiting time, it must be assumed that users possess certain forward-looking properties. This, in turn, necessitates an information-sharing system where waiting time is shared and where people react to this information. Our approach is based on the most recent state-of-art as presented in Vandet & Rich (2021).

Moreover, the tracking of behaviour over a range of heterogeneous inputs requires these to be properly modelled and forecasted. One example is that we use flexible battery state-of-charge distributions, which is based on recent research in Hipolito et al. (2022). Another example is that

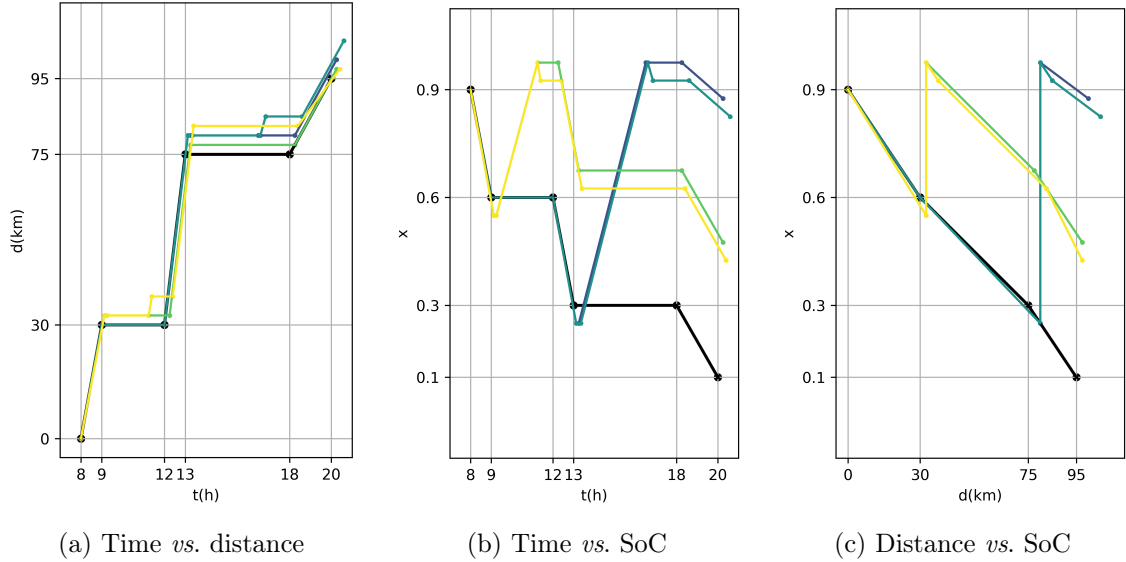


Figure 2: Phase space trajectories over time, driven distance and SoC level, for a hypothetical agent with 4 charging options. Black lines trace non-charging path.

we use probabilistic range projections as applied in Rich et al. (2022).

Literature and contribution

Recent progress on analysis of censored demand Gammelli et al. (2020); Huttel et al. (2022) has highlighted the importance of a proper quantification of the real demand for efficient use of resources in transportation systems. Given the potential large impact of censored data, several authors elect to discard results when censoring is observed Rudloff & Lackner (2014); O’Mahony & Shmoys (2015), or alternatively calibrate their models based on historical data when no censoring was observed Albiński et al. (2018); Jian et al. (2016); Freund et al. (2019). However, previous approaches have been agnostic to queueing dynamics and to the potentially long duration of charging sessions in slow charging infrastructure. Given the large duration of charging events, and the expected queueing at fast charging clusters, it becomes crucial to quantify and track the impact of censored demand. Moreover, by monitoring not only the queueing and charging processes, but also each agent’s decision path, we can track and record all instances in which demand is not served, and thus quantify what would otherwise be censored demand, whilst decoupling it from choice substitution. Hence, quantifying unserved demand is a better indicator of the limitations of the charging infrastructure, as performance can suffer significantly from high demand occurring in short periods of time, which are less likely to be identified when analysing performance solely on utilization.

Methodologically, this paper is based on an even-based micro simulation framework and thereby goes in the direction of Pruckner et al. (2017); Q. Yang et al. (2019), who use discrete event-simulators to represent demand and to analyse capacity utilization in charging systems. Moreover, inspired by Rich et al. (2022) and Vandet & Rich (2021) we use up-sampled trip diaries as input to the simulator. The benefits of doing so, in terms of attaining realistic trip and arrival time patterns, clearly outweigh the potential shortcomings, as discussed in more detail in Rich et al. (2022).

The contribution of the paper lies in the focus on unserved demand when dealing with the charging infrastructure problem. Add to this that, by monitoring this KPI over many different heterogeneous inputs, we are able to generate corresponding heterogeneous outputs and identify emergent behaviour in the system. The relaxation of simple Markovian queueing dynamics (van der Kam et al., 2019; Viswanathan et al., 2016; J. Yang et al., 2017) for non-parametric queues of the G/G/s type creates patterns that are closer to reality. Hence, we refrain from the single-solution approach and present a distribution of solutions which is a function of several probabilistic inputs.

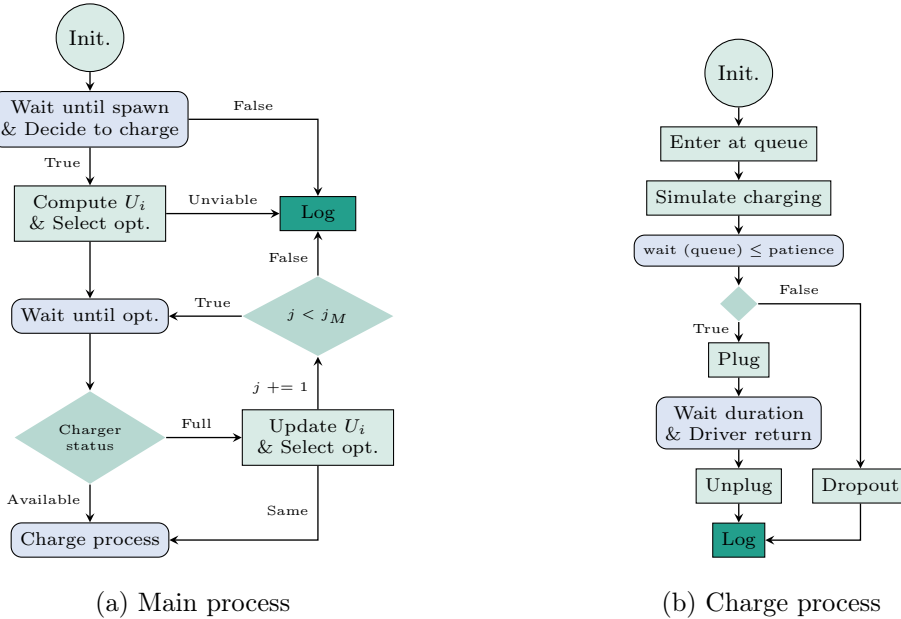


Figure 3: Diagrammatic representation of the iterative simulation process. Light blue background and rounded corners indicate the steps that include a waiting period.

2 METHODOLOGY

We present an agent-based simulation framework to predict the demand for EV charging over many heterogeneous inputs and for agents that interact through an information-sharing system (Vandet & Rich, 2021). The framework is designed with the intent of determining potential configurations of future charging infrastructure that satisfy the growing adoption of EVs, while minimizing disruption to present day travelling patterns and infrastructure deployment costs. To achieve such goal, we develop a synthetic population of EV users by combining trip dairies, and additional household information, based on the Danish National travel survey Baescu & Christiansen (2020). The origin-destination of all agents and their timing of trips (e.g. arrivals and departures) are calibrated to origin-destination (OD) traffic flow data and assigned to the OpenStreetMap road network using the shortest path algorithm. Similar applications of synthesized trip diary data have been presented in Rich et al. (2022); Kavianiipour et al. (2021).

Clearly, the fact that we estimate demand for EV charging in the future based on calibrated trip dairies of today can be criticised. The underlying premise, then, is that the trip demand patterns are invariant with respect to the replacement of conventional cars with EVs. As argued in Rich et al. (2022), if this were not true, it would imply that people would have to change their otherwise optimal behaviour (where battery restrictions are absent) to a different behaviour, which is currently not known. In our view, designing a system where people can maintain a similar behaviour as today, reduces the burden associated with transitioning and thus facilitates adoption of EVs. Therefore, it represents a reasonable compromise when designing a societal optimal infrastructure for the future.

The simulation algorithm is represented diagrammatically in Figs. 3a and 3b, with the latter depicting exclusively the charging process. The intricacies of the simulation algorithm, as well as the agent synthesis process, will be discussed in a forthcoming manuscript, where we explore all branching options that lead to unserved demand. By monitoring the entire system over a period greater than 24 h, it is possible to estimate a range of probabilistic system KPIs, e.g. such as charging station utilization and waiting time at chargers. As we are monitoring, not only charging choices but also intentions to charge in cases where this is not possible, we can also track unserved charging demand in the system. This, in turn, is an extremely powerful performance indicator for the system as it quantifies an otherwise censored part of the demand, thus providing an important metric to plan the development of an efficient of public charging infrastructure.

At the programming level, this model is framed as a discrete event-based simulator, based on the SimPy package for Python Scherfke et al. (2022), allow us to track and coordinate the agent’s movement in space and time.

Table 1: The price of electricity v_e for charging EV in DKK/kWh per charger type.

	normal	fast	private
P_{ch}	< 50	≥ 50	11
v_e	3.50	4.90	1.12

Decision model

The decision framework adopted here makes use of two models deployed in sequence. The first assesses the decision to charge Hipolito et al. (2022), while the second selects the desired opportunity. Both models are stochastic in nature, offering the agents the opportunity to make suboptimal choices at each step, introducing variance to the overall decision process, thus allowing the simulation path to branch away from the local minima. To select the an option, the utilities U_i are assessed, and the respective probabilities are determined by a classical multinomial logit model

$$p_i = \frac{e^{-\beta U_i}}{\sum_i e^{-\beta U_i}}, \quad (1)$$

where $\beta = 1$ is the global scaling factor. The choice set i , as well as the full form of all terms contributing to the utilities, will be described in detail in a forthcoming paper, below follows a brief summary. The selection process relies on a weighted random draw, where the weights are defined by p_i . The utilities are defined in generalized time form, where monetary costs are expressed as $U_i = c_i/\nu_\tau$, considering the general value-of-time (VoT) $\nu_\tau = 91$ DKK/h DTU (2022); Rich & Vandet (2019). The utilities include several contributions

$$U_i = [c_{i\epsilon} + \nu_i(\Delta t_i + w_i) + c_{ip} + c_{ih} + c_{ia}]/\nu_\tau. \quad (2)$$

including the cost of electricity c_ϵ , the value of time for waiting w_i , detours and charging, costs of parking c_{ip} and idle charger time c_{ih} , and finally the cost associated with interrupting an activity.

3 RESULTS

Here, we analyse a baseline scenario that offers a reference to validate the simulator against real-world data. We then proceed with an assessment of the viability of meeting growing demand in a dense urban municipality by expanding fast charging infrastructure.

Baseline scenario

To compare the simulator results against real-world utilization data, we make use of data for a charge station in the municipality of Frederiksberg covering a period of 390 days since the fall of 2021, reporting a mean charged energy of $\delta\epsilon = 17.2$ kWh at a mean power of $P = 7.42$ kW. The charging duration and total parking time are $\Delta t_{ch} = 3.0$ and $\Delta t = 6.2$ h, with utilization rate of $u = 0.52$ events per outlet per day.

To compare, we performed simulations covering a period of one week in the fall of 2022, considering a population of $N_r = 800$ resident and $N_v = 1200$ visitor EVs Danmarks Statistik (2023); Baescu & Christiansen (2020). In addition, the VoT is set as in DTU (2022), the costs of electricity are Tab. 1, and we follow municipality’s parking constraints, whereby parking is free, but limited to 3h between 9h00 and 20h00. Finally, the estimate for the mean charge duration is set at $\tau_c = 0.5, 6$ h. Our simulations indicate that EVs would charge 16.9 and 23.9 kWh on the normal and fast infrastructures, with a utilization of 0.45 and 5.36, respectively. Charging and total duration at the stations are 2.3 and 7.6 h in normal charging infrastructure, while at fast the durations are 42 and 58 min. Agents with home charging (~ 40 %) mostly charge at home, with fewer than 1 % opting for public infrastructure.

The results for the normal charging infrastructure are in line with observed utilization of real infrastructure, but we lack data to validate the results found in the fast charging infrastructure. Discrepancies such as the reduced charge duration in the simulation, 2.3 h, can be traced to two main factors. The first is a higher prevalence of EV that only support charging in single- phase at 3.6kW and load sharing at the charge station.

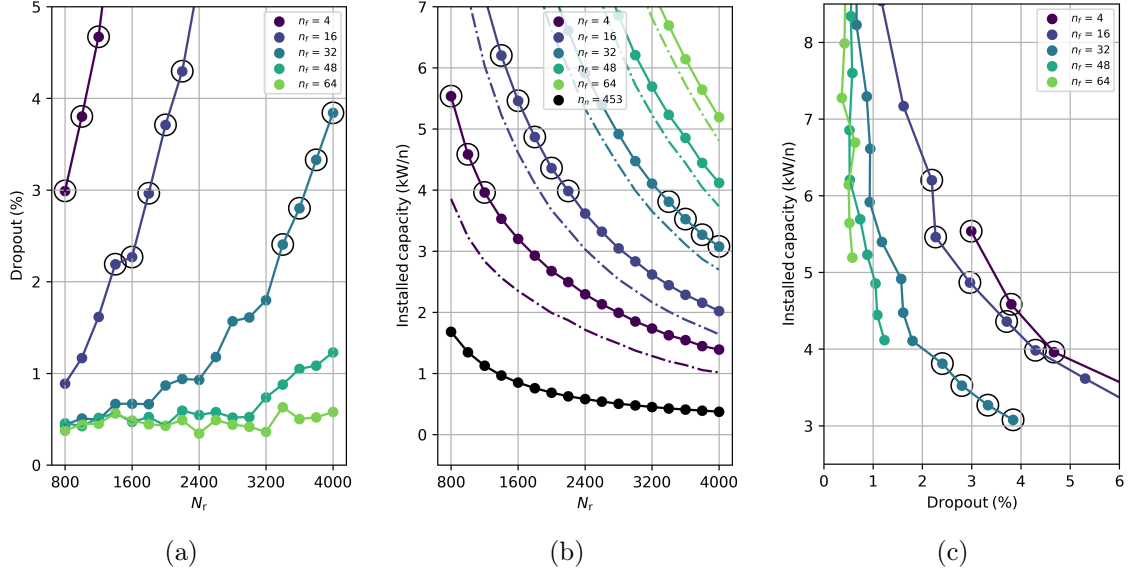


Figure 4: Dropout rate and installed capacity per EV. Lighter colours indicate growing number of fast outlets. Large circles indicate configurations with a dropout rate in $[2, 5]$ %.

The baseline scenario indicates that our simulations are consistent with reality, but indicate a 3% dropout rate from the charging queue. There is little data on quality of service, but anecdotal reports of occasional long waiting, as well as sporadic cases of full charger utilization in the real-world data, indicate that waiting times could easily exceed the lower bounds of the agent’s patience, which manifest in our simulations as dropouts.

Fast charging scenarios

Baseline results indicate that the fast infrastructure can serve up to 10 times more EV than normal charging, indicating that developing this type of infrastructure could reduce the urban impact of EV charging. To explore this hypothesis, we simulate the impact of introducing 3 new fast charging clusters within the municipality towards the West, South and East, in addition to the current cluster at its centre, adding equal number of outlets at each cluster. In Fig. 4a, we show the change in dropout rate for each infrastructure configuration. Results indicate that the baseline dropout rate ~ 3 % can be maintained at least up to $N_r \sim 4000$ with total of 32 fast charging outlets. Moreover, as shown in Fig. 4b and 4c, the efficiency of the infrastructure increases as we observe that more EV can be served while maintaining the baseline performance. This is corroborated by utilization data which increases to 1.99 and 7.66 events per normal and fast outlet per day. Several insights can be drawn from this analysis. First, in all scenarios, the required installed power per EV is larger than the European guideline of 1 kW per EV {Secr tariat g n ral du Conseil} (2022). This discrepancy can be understood in light of population density and the relatively low share of residents with access to private parking in this municipality. Second, upon combined analysis of Fig. 4a and 4b, it is possible to verify that for configurations with similar dropout rate, the required power per EV decreases with the increase in the number of EV, indicating that the infrastructure becomes more efficient as it grows in size.

The daily demand is shown in Fig. 5. While Fig. 5a indicate that some degree of geographical variation is observed, the unserved/censored demand Figs. 5b and 5c is mostly location-independent. At first glance, one would expect variation as depicted in Fig. 1, yet note that information-sharing and the short distances between fast charging infrastructure, facilitate the redistribution of demand, which in turn further increases the efficiency of the infrastructure. Moreover, the discrepancies found in south charger can be understood in light of the lower density of that part of the municipality, as well as the proximity to municipality boarder leads to increase slip-over demand being transferred to other stations in the neighbour municipalities. Accounting for an effectively reduced number of EV using this charger, we can verify that the demand is in line with results in the other 3 fast chargers.

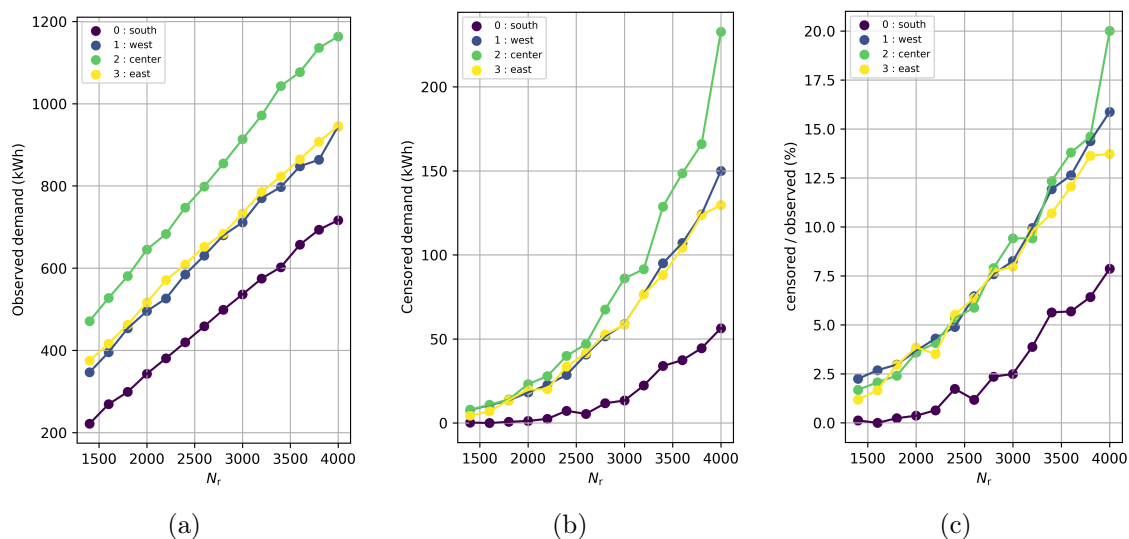


Figure 5: Daily observed (a) and censored (b) demand at four fast chargers with 4 outlets each. Ratio of censored against observed demand (c).

4 CONCLUDING REMARKS

We present an agent-based approach to simulate the EV demand, tracking it over space and time. By virtue of introducing branching options in the simulation algorithm, we can decouple the demand into the observed and censored parts. The censored part offers a quantification of the unserved demand, that can be used to improve rollout strategies of charging infrastructure, that satisfy demand, while minimizing impact on society.

Moreover, the results strongly suggest that the charging demand, both observed and censored, can be evenly distributed, thanks to the introduction of an information-sharing strategy, that leads agents to avoid overcrowded stations.

Future research perspectives

In addition to an expansion of the present analysis, a forthcoming publication will extend the analysis discussed here beyond the spatial distribution of demand at the fast charging infrastructure. It will include, among other goals, an analysis of the intraday variation of demand, to characterize the distinct rhythms of utilization of each type of infrastructure. In addition, the simulation region will be expanded beyond the current municipality, to include the whole city of Copenhagen, which will offer a better platform to quantify the role of information-sharing and additional strategies to balance demand.

ACKNOWLEDGEMENTS

The authors acknowledge T. Unterluggauer, C. A. Vandet and F. C. Pereira for many helpful comments. This work was supported by the FUSE project (EUDP grant number: 64020-1092).

REFERENCES

- Albiński, S., Fontaine, P., & Minner, S. (2018, August). Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations. *Transportation Research Part E: Logistics and Transportation Review*, 116, 59–69. doi: 10.1016/j.tre.2018.05.011
- Baescu, O., & Christiansen, H. (2020). *The Danish National Travel Survey Annual Statistical Report TU0619v2* (Tech. Rep.). Kgs. Lyngby, Denmark: DTU Management. Retrieved from https://orbit.dtu.dk/files/221136854/TU_Denmark_2019.pdf doi: 10.11581/dtu:00000034

- Danmarks Statistik. (2023). *BIL53: Nyregistrerede motorkøretøjer efter område, køretøjstype, brugerforhold og drivmiddel*. Retrieved 2023-02-28, from <https://www.statistikbanken.dk/BIL53>
- DTU. (2022). *Transport unit prices* (Tech. Rep.). Author. Retrieved from <https://www.cta.man.dtu.dk/modelbibliotek/teresa/transportoekonomiske-enhedspriser>
- Freund, D., Henderson, S. G., & Shmoys, D. B. (2019). Bike Sharing. In M. Hu (Ed.), *Sharing Economy* (Vol. 6, pp. 435–459). Cham: Springer International Publishing. doi: 10.1007/978-3-030-01863-4_18
- Gammelli, D., Peled, I., Rodrigues, F., Pacino, D., Kurtaran, H. A., & Pereira, F. C. (2020, November). Estimating latent demand of shared mobility through censored Gaussian Processes. *Transportation Research Part C: Emerging Technologies*, *120*, 102775. doi: 10.1016/j.trc.2020.102775
- Hengsong Wang, Qi Huang, Changhua Zhang, & Aihua Xia. (2010, December). A novel approach for the layout of electric vehicle charging station. In *The 2010 International Conference on Apperceiving Computing and Intelligence Analysis Proceeding* (pp. 64–70). Chengdu, China: IEEE. doi: 10.1109/ICACIA.2010.5709852
- Hipolito, F., Vandet, C. A., & Rich, J. (2022). Charging, steady-state soc and energy storage distributions for ev fleets. *Applied Energy*, *317*, 119065. doi: 10.1016/j.apenergy.2022.119065
- Huttel, F. B., Peled, I., Rodrigues, F., & Pereira, F. C. (2022, November). Modeling Censored Mobility Demand Through Censored Quantile Regression Neural Networks. *IEEE Transactions on Intelligent Transportation Systems*, *23*(11), 21753–21765. doi: 10.1109/TITS.2022.3190194
- Jian, N., Freund, D., Wiberg, H. M., & Henderson, S. G. (2016, December). Simulation optimization for a large-scale bike-sharing system. In *2016 Winter Simulation Conference (WSC)* (pp. 602–613). Washington, DC, USA: IEEE. doi: 10.1109/WSC.2016.7822125
- Kavianipour, M., Fakhrmoosavi, F., Singh, H., Ghamami, M., Zockaie, A., Ouyang, Y., & Jackson, R. (2021, April). Electric vehicle fast charging infrastructure planning in urban networks considering daily travel and charging behavior. *Transportation Research Part D: Transport and Environment*, *93*, 102769. doi: 10.1016/j.trd.2021.102769
- Kuby, M., & Lim, S. (125). The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences*, *39*(2). doi: 10.1016/j.seps.2004.03.001
- Luo, X., & Qiu, R. (2020, April). Electric Vehicle Charging Station Location towards Sustainable Cities. *International Journal of Environmental Research and Public Health*, *17*(8), 2785. doi: 10.3390/ijerph17082785
- O’Mahony, E., & Shmoys, D. (2015, February). Data Analysis and Optimization for (Citi)Bike Sharing. *Proceedings of the AAAI Conference on Artificial Intelligence*, *29*(1). doi: 10.1609/aaai.v29i1.9245
- Pruckner, M., German, R., & Eckhoff, D. (2017). Spatial and temporal charging infrastructure planning using discrete event simulation. In *Proceedings of the 2017 acm sigsim conference on principles of advanced discrete simulation* (p. 249–257). New York, NY, USA: Association for Computing Machinery. Retrieved from <https://doi.org/10.1145/3064911.3064919> doi: 10.1145/3064911.3064919
- Rich, J., & Vandet, C. A. (2019). Is the value of travel time savings increasing? analysis throughout a financial crisis. *Transportation Research Part A: Policy and Practice*, *124*, 145–168. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0965856418307079> doi: <https://doi.org/10.1016/j.tra.2019.03.012>
- Rich, J., Vandet, C. A., & Pilegaard, N. (2022). Cost–benefit of a state-road charging system: The case of denmark. *Transportation Research Part D: Transport and Environment*, *109*, 103330. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1361920922001584> doi: <https://doi.org/10.1016/j.trd.2022.103330>

- Rudloff, C., & Lackner, B. (2014, January). Modeling Demand for Bikesharing Systems: Neighboring Stations as Source for Demand and Reason for Structural Breaks. *Transportation Research Record: Journal of the Transportation Research Board*, 2430(1), 1–11. doi: 10.3141/2430-01
- Scherfke, S., Lünsdorf, O., Müller, K., & Vignaux, T. (2022). *SimPy: v4.0.1*. Retrieved from simpy.readthedocs.io
- {Secrétariat général du Conseil}. (2022, May). *Proposition de RÈGLEMENT DU PARLEMENT EUROPÉEN ET DU CONSEIL sur le déploiement d'une infrastructure pour carburants alternatifs et abrogeant la directive 2014/94/UE du Parlement européen et du Conseil - Orientation générale* (No. ST 8902/1/22 REV1 ST 10877/21 ADD1-ADD7).
- van der Kam, M., Peters, A., van Sark, W., & Alkemade, F. (2019). Agent-Based Modelling of Charging Behaviour of Electric Vehicle Drivers. *Journal of Artificial Societies and Social Simulation*, 22(4), 7. doi: 10.18564/jasss.4133
- Vandet, C. A., & Rich, J. (2021). Optimal placement and sizing of charging infrastructure for evs under information-sharing. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3991823
- Viswanathan, V., Zehe, D., Ivanchev, J., Pelzer, D., Knoll, A., & Aydt, H. (2016, January). Simulation-assisted exploration of charging infrastructure requirements for electric vehicles in urban environments. *Journal of Computational Science*, 12, 1–10. doi: 10.1016/j.jocs.2015.10.012
- Wang, Y.-W., & Lin, C.-C. (2009, September). Locating road-vehicle refueling stations. *Transportation Research Part E: Logistics and Transportation Review*, 45(5), 821–829. doi: 10.1016/j.tre.2009.03.002
- Wang, Y.-W., & Wang, C.-R. (2010, September). Locating passenger vehicle refueling stations. *Transportation Research Part E: Logistics and Transportation Review*, 46(5), 791–801. doi: 10.1016/j.tre.2009.12.001
- Yang, J., Dong, J., & Hu, L. (2017, April). A data-driven optimization-based approach for siting and sizing of electric taxi charging stations. *Transportation Research Part C: Emerging Technologies*, 77, 462–477. doi: 10.1016/j.trc.2017.02.014
- Yang, Q., Sun, S., Deng, S., Zhao, Q., & Zhou, M. (2019). Optimal sizing of pev fast charging stations with markovian demand characterization. *IEEE Transactions on Smart Grid*, 10(4), 4457–4466. doi: 10.1109/TSG.2018.2860783

Varying critical time to collision: a perspective of driver space

Yiru Jiao^{*1}, Simeon C. Calvert¹, Sander van Cranenburgh², and Hans van Lint¹

¹Department of Transport & Planning, Delft University of Technology, The Netherlands

²Transport and Logistics Group, Department of Engineering Systems and Services, Delft University of Technology, The Netherlands

SHORT SUMMARY

Collision warnings play a crucial role in preventing crashes based on estimating the critical time to potential accidents. Existing research and applications mainly focus on longitudinal vehicle interaction and headway keeping on highways. However, urban driving with frequent lateral interaction may have different critical time to collision. This study considers both longitudinal and lateral interaction through a two-dimensional spacing measure, driver space, to estimate the average critical time for drivers to respond to a potential collision. With the average spacing between vehicles at different levels of discomfort and in different relative speeds, we estimate the critical time via linear regression. Our experiments on two trajectory datasets find that drivers are more alert to collision dangers on highways compared to urban intersections, and drivers respond to potential collisions more quickly during lateral interaction than longitudinal. These findings emphasise the need of tailored collision warning systems for further improving road safety.

Keywords: Critical time to collision, driver space, driving safety, trajectory data.

1 INTRODUCTION

Autonomous driving and advanced driving assistance systems (ADAS) have been a rapidly evolving research area, aiming to enhance road safety and efficiency. As part of that, forward collision warning and collision avoidance systems play a critical role in preventing crashes. These systems distinguish emergencies based on the critical time to collision, which refers to the amount of time a vehicle has to respond to a potential collision.

Critical time to collision has been typically considered in car following scenarios for longitudinal interaction between vehicles (Bella & Russo, 2011; Tawfeek & El-Basyouny, 2018). As an example, Time to Collision (TTC) is one of the most effective and broadly used indicators for warning rear-end collisions (Lu et al., 2021). It is calculated based on the relative position and relative velocity between two approaching vehicles, assuming that their movement continue without change. Then a critical threshold (denoted as TTC^*) is set to distinguish un(safe) situations. ADAS usually set a fixed threshold, but some studies found that TTC^* can vary among drivers (Kusano et al., 2015) and in different traffic environments (Arun et al., 2021) due to different human perception.

As the number of ADAS-equipped vehicles increases in urban areas, there is a growing need to consider critical time to collision in a wider range of driving scenarios. In urban traffic, the interaction between vehicles involves more angled movement such as lane changing and turning (Zhao et al., 2020). This entails potential collisions beyond rear-end crashes, such as head-on and side-swipe accidents (Theofilatos et al., 2012).

Does human perception of collision danger differ on highways and in urban traffic, with or without lateral interaction? This study addresses the question from the perspective of driver space. Driver space measures two-dimensional spacing between vehicles. Our previous study (Jiao et al., 2022) presented a method to quantify the 2d spacing in a probabilistic manner, which is based on the accumulative presence of vehicles. The method can locate a set of boundaries of spacing in various driving scenarios, with which we can estimate the average critical time to collision given the relative speed between vehicles.

In the following sections, we will firstly introduce our methods in Section 2. Then we apply the methods to two trajectory datasets collected by drones in the U.S. One is over an expressway and the other is over an intersection. The results and discussion will be presented in Section 3. Section 4 will conclude the study. Our findings are expected to aid the development of collision warning systems.

2 METHODOLOGY

This section will at first briefly explain driver space and its quantification. Then we will introduce how to obtain comfortable and uncomfortable spacing from the quantification. Finally, we will present how this study estimates critical time to collision.

Driver space and its quantification

The driver space of a vehicle refers to a set of boundaries where the driver experiences different levels of discomfort. When the driver space is intruded (i.e., when two vehicles are close enough), discomfort is raised. This discomfort motivates the drivers to maintain a proper distance from each other.

We use formula (1) to quantify the varying levels of discomfort caused by driver space intrusion into a value between 0 and 1. The higher the $p(x, y)$, the more discomfort is caused. For two vehicles i and j , x and y are transformed coordinates of j in a system where i is at the origin and the y -axis points to the direction of their relative velocity.

$$p(x, y|\boldsymbol{\theta}) = \exp\left(-\left|\frac{x}{r_x}\right|^{\beta_x} - \left|\frac{y}{r_y}\right|^{\beta_y}\right), \quad (1)$$

where each of $\boldsymbol{\theta} = (r_x, r_y, \beta_x, \beta_y)^\top$ has two components:

$$\begin{cases} \theta = \frac{1 + \text{sgn}(x)}{2}\theta^+ + \frac{1 - \text{sgn}(x)}{2}\theta^- & \text{for } \theta = r_x, \beta_x, \\ \theta = \frac{1 + \text{sgn}(y)}{2}\theta^+ + \frac{1 - \text{sgn}(y)}{2}\theta^- & \text{for } \theta = r_y, \beta_y. \end{cases} \quad (2)$$

Formula (1) is adapted from the density function of the generalised Gaussian distribution. We use it to parameterise the comfort-discomfort transition during approaching. Among the parameters in formulae (2), $\boldsymbol{r} = \{r_x^+, r_x^-, r_y^+, r_y^-\}$ determine driver space boundaries in different directions, where $p = e^{-1}$; and $\boldsymbol{\beta} = \{\beta_x^+, \beta_x^-, \beta_y^+, \beta_y^-\}$ determine how fast comfort changes to discomfort across the boundaries in different directions.

Driver space, i.e., two-dimensional vehicle spacing, can vary in different scenarios. Correspondingly, given samples of vehicles in various scenarios, $\boldsymbol{\theta}$ can also be different. The inference of $\boldsymbol{\theta}$ is achieved by estimating the density of the accumulative presence of vehicle pairs. In each pair, one of them is considered as an ego vehicle and the other is a surrounding vehicle. By aggregating all pairs in the same scenario, an average ego vehicle is abstracted, and driver space is shaped by the surrounding vehicles. In our previous research, we developed a method that makes consistent inference of driver space. The readers are referred to Jiao et al. (2022) for technical details.

Spacing between approaching vehicles

Driver space reflects the average preference of drivers' spacing behaviour. This includes spacing when vehicles are approaching each other and leaving away. As collision is mainly a result of getting overly close, we only take approaching spacing into account.

As stated above, our quantification of driver space transforms vehicle coordinates. The transformation aligns the y-axis in the target coordinate system to the direction of the relative velocity between two interacting vehicles. This ensures the independence of x and y . For two vehicles approaching each other, the transformed y will be larger than 0. In formula (2), \boldsymbol{r} and $\boldsymbol{\beta}$ define where the intrusion discomfort increases most quickly and how quickly the increase is. Therefore, among the parameters, r_y^+ and β_y^+ characterise the spacing between vehicles that are approaching each other.

With the inferred r_y^+ and β_y^+ in various scenarios, we can correspondingly estimate the spacing between approaching vehicles under different extent of discomfort. Given a certain extent of discomfort indicated by p , the approaching spacing $s(p)$ can be computed by solving the inverse of equation (1). As x and y are independent from each other, $s(p)$ is solved as

$$s(p) = \hat{r}_y^+ (-\ln(p))^{1/\hat{\beta}_y^+}. \quad (3)$$

For vehicle samples in different situations and at different relative speeds, approaching spacing can be computed correspondingly, and then we can estimate the critical time to collision.

Critical time to collision

Our quantification of driver space is based on the assumed negative correlation between the presence of vehicles and the extent of discomfort. During the approaching between vehicles, drivers can feel increasing discomfort, and this increase is assumed to be fastest when $p = e^{-1}$. When p is close to 0, drivers are comfortable with other vehicles' presence. When p is close to 1, the vehicles are so close that a collision is imminent.

In this study, we consider two kinds of critical time. One is the time from comfort to collision, the other is the time from discomfort to collision. The former represents the time from when a driver starts to feel discomfort due to approaching to a potential collision, and the latter represents the time from when a driver experiences clear discomfort and seeks for change to a potential collision.

We firstly consider spacing from comfort to collision as $s_c = s(p = 0.1)$ and spacing from discomfort to collision as $s_d = s(p = e^{-1})$. Generally denote a series of spacing under different relative speeds \boldsymbol{v} as \boldsymbol{s} . For each relative speed condition v , there is a corresponding s . Given the linear physical relationship between speed and distance, we assume that \boldsymbol{s} and \boldsymbol{v} are linearly correlated as equation

(4) and the coefficient t^* is the critical time.

$$s = s_0 + vt^* + \epsilon, \quad (4)$$

where s_0 is the spacing when vehicles are relative static, and ϵ is the random error that satisfies $E(\epsilon) = 0$.

We then use least squares fitting to estimate \hat{s}_0 and \hat{t}^* . The unbiased estimation of them are solved as equations (5).

$$\begin{cases} \hat{t}^* = \frac{\sum_{i=1}^n (v_i - \bar{v})(s_i - \bar{s})}{\sum_{i=1}^n (v_i - \bar{v})^2}, \\ \hat{s}_0 = \bar{s} - \bar{v}\hat{t}^*. \end{cases} \quad (5)$$

In addition, the variance of \hat{t}^* can also be estimated as shown in equation (6)

$$V(\hat{t}^*) = \frac{V(\epsilon)}{\sum_{i=1}^n (v_i - \bar{v})^2} = \frac{V(s)}{\sum_{i=1}^n (v_i - \bar{v})^2}. \quad (6)$$

In this way, based on s_c and s_d , respectively, we can estimate the critical time from comfort to collision, denoted as \hat{t}_c^* , and the critical time from discomfort to collision, denoted as \hat{t}_d^* .

3 RESULTS AND DISCUSSION

Datasets

We apply the proposed approach to two trajectory datasets. Both of them are collected in the U.S., of which the videos of vehicle movement were recorded by drones, and then computer vision algorithms were used to process the videos into numerical coordinates. As shown in Figure 1, one of them is a weaving segment of an expressway marked with *A* (Zheng et al., 2022), where 9,956 vehicles in 1.44 hours were recorded; and the other is an intersection denoted as *GL* (Zhan et al., 2019), where 10,510 vehicles in 4.34 hours were recorded. This study considers vehicle-vehicle interaction only, so pedestrians and cyclists in the intersection *GL* are excluded.

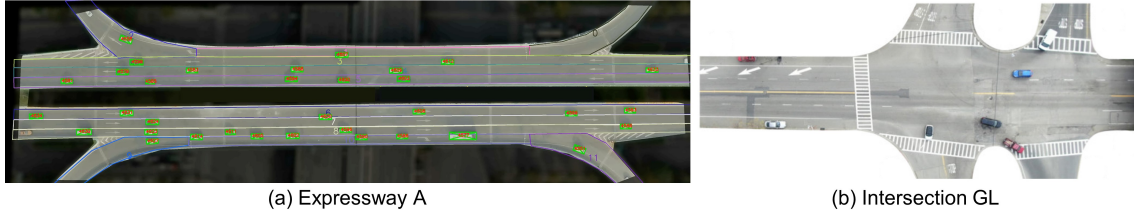


Figure 1. Two road segments that are analysed. (a) and (b) are respectively reused from Figure 8 in Zheng et al. (2022) and Fig.2 in Zhan et al. (2019).

We consider lateral interaction at the intersection *GL*. If the angle between the moving directions of the two vehicles in a pair is smaller than 15 degrees or larger than 165 degrees, they are considered to interact only in the longitudinal direction (i.e., car following); otherwise, their interaction is considered to also involve the lateral direction (e.g., lane-changing and turning). In this way, our analysis considers three situations: interaction in the expressway, longitudinal interaction at the intersection, and lateral interaction at the intersection. They are referred to as *Expressway A*, *Non-lateral GL*, and *With-lateral GL* in the following.

Inferred driver space

We first sample vehicle pairs present at the same frame, and then infer the driver spaces that they accumulatively shape in different scenarios. Figure 2 shows several inference results. Driver spaces in different situations and under 2, 4, 6, 8, 10 m/s of the relative speed are drawn. In each subplot, the yellow dots indicate surrounding vehicle positions, and the driver spaces that they shape are plotted as contours at different levels of intrusion discomfort.

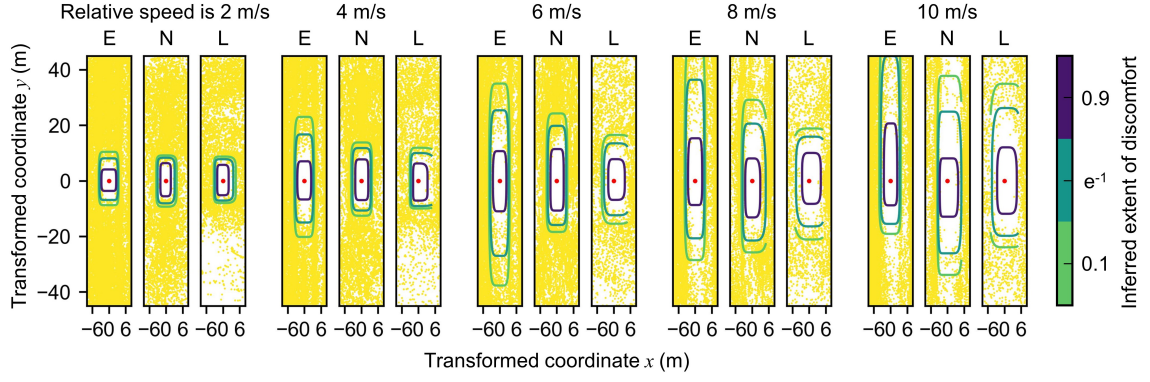


Figure 2. Inferred driver space in different scenarios. E: Expressway A; N: Non-lateral GL; L: With-lateral GL.

Figure 2 provides an intuitive impression of how drivers maintain distance in various scenarios. Driver space expands as the relative speed between the vehicles increases. While at the same relative speed, the average driver space is largest on the expressway, followed by that formed by vehicles at the intersection during non-lateral interaction, and the driver space formed during lateral interaction is the smallest.

The difference in driver space in different scenarios imply that drivers perceive and react differently during interaction. For instance, when two vehicles are moving in the same direction on the expressway at a similar relative speed, they maintain a larger spacing compared to when they are driving at the intersection with non-lateral interaction. In the former case, drivers tend to be more cautious and distant from each other. Similarly, if the two vehicles were to encounter each other during a lateral interaction (i.e., they are in angled directions but the relative speed value is still similar), their spacing would be even smaller. Such differences reflect drivers' different perception about the approaching between each other, which results in different reaction.

Estimated critical time to collision

In the three situations of *Expressway A*, *Non-Lateral GL*, and *With-lateral GL*, we compute the spacing at various relative speeds and then estimate the corresponding critical time to collision. The results are presented in Figure 3. Each column in this figure represents a specific situation. The upper plot in each column displays the computed spacing, and the bottom plot displays the estimated time and its variance.

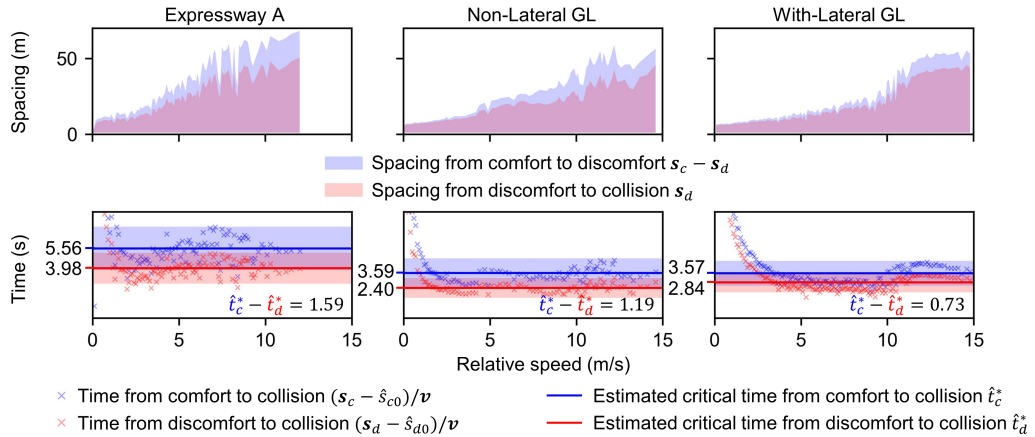


Figure 3. Average spacing and estimated critical time to collision in different situations.

The red lines indicate the estimated critical time from discomfort to collision \hat{t}_d^* . Our results show that this value is 3.98s on the expressway A, but reduces to 2.40s at the intersection GL during non-lateral interaction and to 2.84s during lateral interaction. This estimated time represents the duration that a driver realises the danger before a potential collision. A higher value means that

the driver perceive the potential danger earlier. Therefore, the results suggest that drivers tend to perceive potential collisions more quickly when they drive on the expressway than when they drive at the intersection. This makes sense as the traffic environment on highways is generally simpler than in urban areas, so that a potential collision is more predictable.

The difference between the blue lines and the red lines (i.e., $\hat{t}_c^* - \hat{t}_d^*$) indicates the estimated time from when a driver begins to feel discomfort to when the driver feels clear discomfort that motivates behaviour change. This time difference reflects how quickly the driver responds to a potential collision. Our results show that drivers respond most quickly when they drive at the intersection during lateral interaction, with a time difference of 0.73s. However, during non-lateral interaction, the time difference is 1.19s at the intersection and 1.59s on the expressway. This could be due to that lateral interaction typically takes place in a closer distance, which requires faster response of drivers. Corroborating support can be seen from the spacing plots, where the spacing from comfort to discomfort is larger for longitudinal interaction than for lateral interaction.

4 CONCLUSIONS

This study offers a driver space perspective to analyse the critical time to collision and the issuing of collision warning. Firstly the averaged two-dimensional spacing of drivers in different situations is quantified based on the accumulative presence of vehicles. Then we use least squares to estimate the critical time to collision by fitting the linear relationship between the approaching spacing and the corresponding relative speeds. We analysed three situations in this study, including highway interaction, longitudinal interaction at an intersection, and lateral interaction at the same intersection. Our results show that drivers may take different amount of time to perceive and react to a potential collision in different situations. Specifically, drivers perceive the danger of a potential collision sooner on highways than at urban intersections, which implies that they are more sensitive to potential collision when driving on highways. Meanwhile, drivers take quicker reaction during lateral interaction than longitudinal interaction. These results suggest that collision warning systems need to be tailored to specific driving situations, with a greater warning sensitivity on highways and a proper time design to avoid distracting drivers' proactive reaction. By accounting for the variation of driving situations and interactions, drivers can get more reliable reminders and the road safety can be improved for all road users.

In the next step, we will apply the approach to more intersection situations. Figure 3 shows two thresholds for the situation of *With-Lateral GL*. The lower threshold is seen when the relative speed is smaller than around 10 m/s, and the higher one appears when the relative speed is larger. This implies that, during lateral interaction, drivers may have even slower awareness of potential collisions if their speed difference is smaller. Such implication poses a concern as the delayed awareness may result in inadequate reaction time to avoid potential collisions. Therefore, we still need to examine whether this phenomenon is common across different intersections and to explore the underlying causes.

This study contributes to the growing research field on ADAS and autonomous driving by extending the scope of critical time to collision to a wider range of driving scenarios, beyond car-following that is typically focused. However, it is important to note that the critical time to collision is just one aspect of safe driving. Other factors, such as vehicle-to-vehicle communication, pedestrian and bicycle safety, and the ability to detect and respond to unexpected situations on the road, also need to be considered in developing more advanced collision avoidance systems.

ACKNOWLEDGEMENTS

This work is supported by the TU Delft AI Labs programme.

REFERENCES

Arun, A., Haque, M. M., Washington, S., Sayed, T., & Mannering, F. (2021). A systematic review

- of traffic conflict-based safety measures with a focus on application context. *Analytic Methods in Accident Research*, 32, 100185. doi: 10.1016/j.amar.2021.100185
- Bella, F., & Russo, R. (2011). A collision warning system for rear-end collision: a driving simulator study. *Procedia - Social and Behavioral Sciences*, 20, 676–686. doi: 10.1016/j.sbspro.2011.08.075
- Jiao, Y., Calvert, S. C., van Cranenburgh, S., & van Lint, H. (2022). Probabilistic representation for driver space and its inference from urban trajectory data. *Available at SSRN*. doi: 10.2139/ssrn.4187513
- Kusano, K. D., Chen, R., Montgomery, J., & Gabler, H. C. (2015). Population distributions of time to collision at brake application during car following from naturalistic driving data. *Journal of Safety Research*, 54, 95.e29–104. doi: 10.1016/j.jsr.2015.06.011
- Lu, C., He, X., van Lint, H., Tu, H., Happee, R., & Wang, M. (2021). Performance evaluation of surrogate measures of safety with naturalistic driving data. *Accident Analysis & Prevention*, 162, 106403. doi: 10.1016/j.aap.2021.106403
- Tawfeek, M. H., & El-Basyouny, K. (2018). A perceptual forward collision warning model using naturalistic driving data. *Canadian Journal of Civil Engineering*, 45(10), 899–907. doi: 10.1139/cjce-2017-0592
- Theofilatos, A., Graham, D., & Yannis, G. (2012). Factors affecting accident severity inside and outside urban areas in greece. *Traffic Injury Prevention*, 13(5), 458–467. doi: 10.1080/15389588.2012.661110
- Zhan, W., Sun, L., Wang, D., Shi, H., Clause, A., Naumann, M., ... Tomizuka, M. (2019). INTERACTION Dataset: An INTERnational, Adversarial and Cooperative moTION Dataset in Interactive Driving Scenarios with Semantic Maps. *arXiv*. doi: 10.48550/arXiv.1910.03088
- Zhao, J., Knoop, V. L., & Wang, M. (2020). Two-dimensional vehicular movement modelling at intersections based on optimal control. *Transportation Research Part B: Methodological*, 138, 1–22. doi: 10.1016/j.trb.2020.04.001
- Zheng, O., Abdel-Aty, M., Yue, L., Abdelraouf, A., Wang, Z., & Mahmoud, N. (2022). Citysim: A drone-based vehicle trajectory dataset for safety oriented research and digital twins. *arXiv*. doi: 10.48550/arXiv.2208.11036

Understanding the cycle traffic impacts of Cycle Superhighways in London

Xiaowei Zhu¹, Daniel J. Graham^{1,2,*}, and Anupriya²

¹Department of Mathematics, Imperial College London, London, SW7 2AZ, UK.

²Transport Strategy Centre, Imperial College London, London, SW7 2AZ, UK.

SHORT SUMMARY

Cycle Superhighways (CS) are the cycle routes that run between central London and outer London. They were introduced in 2008 as a way to encourage cycling and improve safety. This paper investigates the causal cycling demand and safety impacts arising from the introduction of CS. The analysis uses road traffic and accident data from the Department for Transport in the UK. Propensity score matching and panel outcome regression models are employed and compared to estimate the effects of CS for two different infrastructure types - segregated and non-segregated. Our results suggest that, on average, the intervention had a positive effect on cycle flow volume and cycle accidents, but no statistically significant effect on the cycle accident rate. Nevertheless, we find that segregated CS show a statistically significant decrease in cycle accident rate.

Keywords: cycleway investments, demand, safety, causal analysis, heterogeneous impact.

1 INTRODUCTION

Cycling has long been regarded as a healthy, economic, and environmental-friendly way of fulfilling one's day-to-day travel needs. With the aim to increase cycling in London, Cycle Superhighways (CS) were introduced across London in 2008. Figure 1 presents an initial route map of the CS. CS are cycle pathways extending from the outer parts of London to its centre¹, that were developed to enable safer, quicker, and more direct travel within the city. Several variants of CS have been also introduced across North America, Australia and Europe to serve the longer distance cycle commutes in metropolitan centres (Pucher & Buehler, 2017). The overarching aim of this paper is to contribute to the growing empirical evidence on the impact of such cycling infrastructure investments on cycle traffic and cyclist safety.

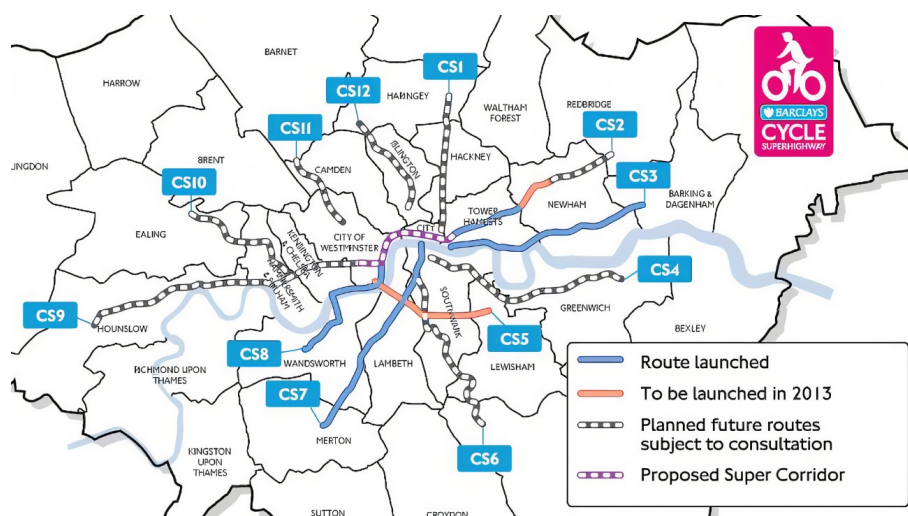


Figure 1: Route plan map of the Cycle Superhighways in London.

CS incorporated a variety of measures to improve cyclist safety including (Transport for London, 2011): (1) realigned traffic and bus lanes to create more space for cyclists on busy stretches of the

¹<https://tfl.gov.uk/modes/cycling/routes-and-maps/cycleways>

routes, (2) re-designed junctions to make them safer for cyclists (say, by removing left-turn slip roads), (3) blind-spot visibility mirrors at signalised junctions in order to improve the visibility of cyclists to heavy goods vehicle drivers, (4) new advanced stop lines and extensions to existing ones (to a minimum of 5 meters) in order to help cyclists move away from traffic signals before other traffic, and, (5) segregated cycle lanes at particularly busy sections of the routes, including Stockwell Gyratory and Wandsworth Bridge roundabout. However, given the associated infrastructure costs, the initial implementation of CS drew widespread criticism. Opponents claimed that the safety impacts of CS were overstated and referred to CS as nothing but blue paint². It is, therefore, imperative to understand the traffic impacts of CS, particularly those related to cyclist safety. In this paper, we investigate the causal effect of CS on cycle flow volume, number of cycle accidents, and cycle accident rate.

Several ex-post evaluations have been carried out in the past to understand the traffic impacts of cycle lanes, especially in regards to collisions (see DiGioia et al., 2017, for a detailed review). These studies mostly compare crashes before and after the deployment of cycle lanes to quantify the effects of the intervention. We argue the impact estimated in these studies may suffer from confounding biases which occur primarily from the non-random nature of such infrastructure investments. In other words, there may exist confounding factors that determine both the likelihood of the intervention and the resulting demand and safety impacts. For instance, CS are more likely to be chosen for roads with large cycle flow volumes, however, there is an inherent scale effect: more cycling usually implies higher cycle-related accidents. Additional biases may emerge from temporal trends in the data. Thus, the estimates derived from a simple before-after comparison of demand and safety indicators may not reflect the true intervention effect.

In this study, we adopt causal inference approaches; (i) propensity score matching and (ii) panel outcome regression with fixed effects; that allow an unbiased estimation of the causal effect by effectively adjusting for such confounding factors. Our analysis uses road traffic and accident data from the UK Department for Transport. The closest precedent to our analysis is the study by Li et al. (2017) that quantified the causal impact of CS in London on cycling volume and collision rate at the network (that is, aggregate) level. However, we exploit the granularity of the data in hand to estimate the impact of CS for different infrastructure types - segregated versus non-segregated CS. We thus contribute with novel insights on how segregated CS segments perform with respect to the non-segregated ones.

2 METHODOLOGY

This section has two main subsections. The first subsection introduces the causal inference framework, which is followed by a description of the methods used in this paper: propensity score matching and panel outcome regression. Both of the methods are applied to compare the reliability of estimation. In the second subsection, we summarise the relevant details of the data that we use to estimate the impact of CS.

Causal Inference Framework

We use Rubin’s Causal Model Rubin (1974) to develop a causal inference framework as follows: Let $Z_i = (Y_i, T_i, X_i)$ represents the observed data, where $i = 1, 2, \dots, N$. N is the total population. Here, Y_i is the outcome of interest for an individual unit i . T_i denotes the binary treatment indicator. If $T_i = 1$, the unit i receives the treatment, otherwise, $T_i = 0$. X_i is the covariates, describing the characteristics of unit i . The potential outcome is defined as $Y_i(T)$ for each unit i . $Y_i(1)$ and $Y_i(0)$ represent the potential outcomes for unit i under treatment and control respectively. Then, the treatment effect for each unit i can be defined as:

$$\tau_i = Y_i(1) - Y_i(0) \tag{1}$$

However, we can only observe one of potential outcome of $Y_i(1)$ or $Y_i(0)$. As a result, we can not directly estimate each unit treatment effect τ_i . Instead, we can estimate the average treatment effect (ATE).

$$\tau_{ATE} = E[Y_i(1) - Y_i(0)] \tag{2}$$

We can estimate the ATE if the following three assumptions hold.

²<https://ecf.com/news-and-events/news/evolution-cycle-superhighways-london>

- **Conditional Independence Assumption:** This assumption means that given the observed covariates X , the potential outcomes are independent of the treatment, i.e. $Y(0), Y(1) \perp\!\!\!\perp T|X$.
- **Common Support -** This condition requires that each unit i has a positive probability of both receiving the treatment or not. There is no probability that one unit is always treated or untreated: $0 < P(T = 1|X) < 1$.
- **Stable Unit Treatment Value Assumption(SUTVA) -** The SUTVA requires that the observed outcomes under a given treatment allocation must be equivalent to potential outcomes under that allocation. i.e. $Y_i = I_1(T_i)Y_i(1) + (1 - I_1(T_i))Y_i(0), \forall i = 1, 2, \dots, n$, where $I_1(T_i)$ is the indicator function for receiving the treatment.

All the three assumptions above are known as strong ignorability. Under this condition, the ATE (2) can be estimated from the observational data. This can be demonstrated as follows:

$$\begin{aligned}
\tau_{ATE} &= E[Y_i(1) - Y_i(0)] \\
&= E_X[E(Y_i(1)|X_i = x) - E(Y_i(0)|X_i = x)] \\
&= E_X[E(Y_i(1)|X_i = x, T_i = 1) - E(Y_i(0)|X_i = x, T_i = 0)] \\
&= E_X[E(Y_i|X_i = x, T_i = 1) - E(Y_i|X_i = x, T_i = 0)]
\end{aligned} \tag{3}$$

Propensity Score Matching

Propensity Score Matching (PSM) is a statistical method to estimate an intervention or treatment. The concept of PSM was first introduced by Rosenbaum and Rubin (1983) Rosenbaum & Rubin (1983) and further developed by Heckman et al. (1997) Heckman et al. (1997). Suppose that we want to compare the treatment effect between a control group and a treatment group. For example, in this study, the treatment is the construction of Cycle Superhighways(CS) or not. The treatment group contains those road segments with the installation of CS, while the control group includes other road segments. A naive method to estimate the treatment effect is to directly compare the difference between the two original groups. However, usually, the treatment is not assigned randomly on each individual. There exist confounding variables which affect both the treatment and the outcome.

To avoid this selection bias, we can do a matching between the control and treatment groups. PSM is one of the matching methods. It first uses models like logit or probit models to estimate the probability that each individual receives the treatment, which is called the propensity score (PS). Here we use logit model, which is defined as:

$$P(T = 1|X) = \frac{1}{1 + \exp(-(\alpha + X\beta))} \tag{4}$$

where α is the intercept and β is the coefficient vector. Then, each individual in the treatment group is matched to the individual in the control group with similar propensity score. There are four main matching algorithms: nearest neighbour matching, caliper and radius matching, stratification and Interval matching, kernel and local linear matching. Finally, the treatment effect is estimated by the difference between the two matched groups. If the strong ignorability assumption is satisfied and the matching algorithm is the nearest neighbour matching, then the ATE can be estimated by the following equation:

$$\hat{\tau}_{ATE} = \frac{1}{N_T} \sum_{i=1}^{N_T} (Y_i(1) - \hat{Y}_i(0)) \tag{5}$$

where N_T represents the total units of treatment group, $Y_i(1)$ is the outcome of i th unit in the treatment group, $\hat{Y}_i(0)$ is the closest unit in the control group in terms of propensity score distance that is matched to the i th treatment unit. It should be noted that different matching algorithm has different estimation equation. The main advantage of PSM is that it reduces the multiple dimension of matching to a single dimension, i.e. propensity score. More details of PSM can be found in Caliendo and Kopeinig (2008) Caliendo & Kopeinig (2008)

Panel outcome regression with fixed effects

One drawback of the propensity score matching method is that its performance highly relies on the choice of confounding factors. If there exist significant confounding factors but unobserved, the result may be unreliable. Thus, in order to control for these unobserved confounding factors,

we also implement panel outcome regression to compare with PSM. There are four main kinds of panel outcome regression models: pooled model, first differences, random effects and fixed effects. Here we introduce panel outcome regression with fixed effects. For more details of other models, see Wooldridge(2010) Wooldridge (2010). Suppose that the data generating process is

$$y_{it} = X_{it}^T \beta + W_i^T \gamma + \epsilon_{it} \quad (6)$$

where X_{it}^T is a $K \times 1$ vector of observed time-variant covariates and W_i^T is an $J \times 1$ vector of unobserved time-invariant covariates. ϵ_{it} is the error term. $E[\epsilon_{it}|X_{it}, W_i] = 0$, $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T$. The fixed effects model assumes that each individual has a unique attribute that is constant through time. The panel model is

$$y_{it} = \alpha_i + X_{it}^T \beta + \epsilon_{it} \quad (7)$$

One possible way to estimate is to use the within estimator. The formula is as follows:

$$y_{it} - \bar{y}_i = (X_{it} - \bar{X}_i)^T \beta + (\epsilon_{it} - \bar{\epsilon}_i)$$

where \bar{y}_i , \bar{X}_i , $\bar{\epsilon}_i$ are the respective mean over time. The advantage of the fixed effect model is that it can effectively deal with the unobserved time-invariant confounding factors. However, it fails to control for the time-varying confounding factors.

DATA

In this section, we describe the relevant datasets and variables we used in this study. There are four main datasets which are highly related to this study: accident data, road data, cycleway data, socioeconomic data.

- Accident data

The accident data is from STATS19, published by Department for Transport³. This dataset gives a detailed description of road accidents in Great Britain, including the date and location of accident, vehicle type, casualty details and severity. In this study, we only focus on the cycle-related accident data from 2000 to 2020 in Greater London.

- Road data

The road data is from road traffic statistics, published by Department for Transport⁴. This dataset gives number of vehicles that travel past the count point on an average day of the year. Here, we use the annual average daily traffic volume (AADT) and annual average daily bicycle volume (AADB) from 2000 to 2020 in Greater London.

- Cycleway data

The cycleway data is from public TfL data⁵. The dataset records the position and type of cycleway in London. In this paper, we mainly focus on the Cycle Superhighway.

- Socioeconomic data

The socioeconomic data is from Office for National Statistics (ONS)⁶ provides the data related to economy, population and society at national, regional and local levels in United Kingdom. Here, we use the population density, employee numbers, index of multiple deprivation (IMD) at the level of Lower Layer Super Output Areas (LSOA).

In this study, the observation unit is LSOA. We use the count points in road data as basis and link other dataset to the road data. To be more specific, for each record of the accident data, we calculate its nearest count point in road data. If the nearest distance is less than a pre-defined threshold (here we use 0.4 kilometer), then we can allocate the record to its nearest count point. Similarly, for each record of the cycleway data and socioeconomic data, we can allocate it to its nearest count point. And if the nearest distance of one record of cycleway data is greater than

³<https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data>.

⁴<https://roadtraffic.dft.gov.uk/regions/6>.

⁵<https://cycling.data.tfl.gov.uk/>.

⁶<https://www.ons.gov.uk/>.

a pre-defined threshold (e.g 1.5 kilometer), we can assume that there is no construction of Cycle Superhighway.

Although 12 CS routes had been planned, only part of them were put into use. In this paper, we studied 6 CS routes. The detailed description of these 6 CS routes is in Table 1.

CS No.	Open time	Length	Route	Type
CS1	2016 April	9.5km	The city to Tottenham	Partly segregated two-way cycle tracks. Most shared with bus Segregated cycle tracks are rarely seen.
CS2	2011 July	6.8km	Stratford to Aldgate	(Note: An upgradation was added in 2016) Mostly segregated two-way cycle tracks.
CS3	2010 July	12.3km	Barking to Tower Hill	Completely segregated two-way cycle tracks.
CS5	2015 Autumn	1.4km	Oval to Pimlico	Segregated cycle tracks are rarely seen. Mostly shared with buses.
CS7	2010 July	13.7km	Merton to the City	Segregated cycle tracks are rarely seen.
CS8	2011 July	8.2km	Wandsworth to Westminster	

Table 1: The characteristics of each Cycle Superhighways

The distribution of these 6 CS routes can be seen in Figure 2.

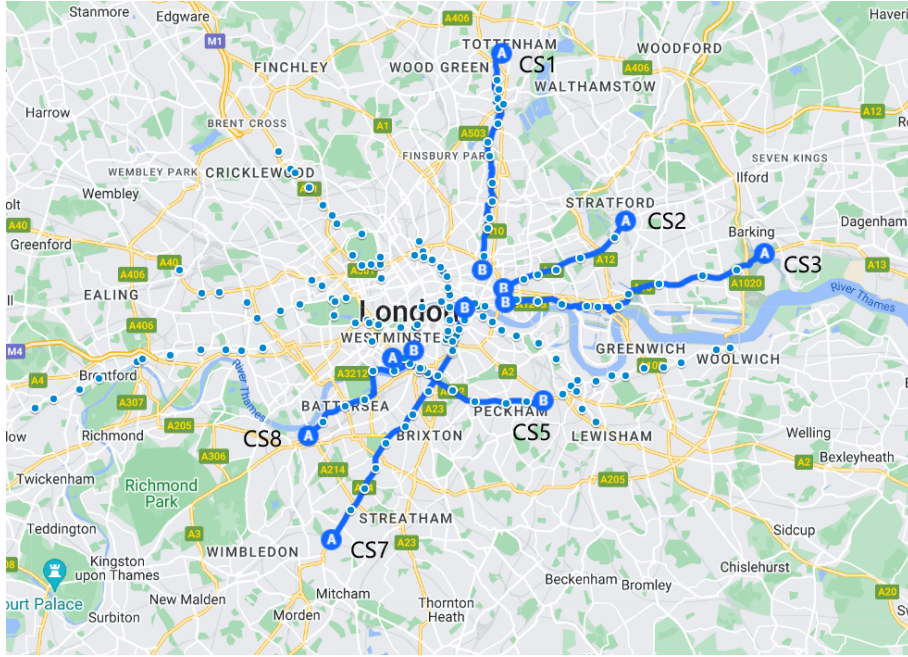


Figure 2: The distribution of 6 CS routes

Here, we choose 80 road segments with construction of CS as the treatment group, and select 434 road segments as the control group. Their distribution can be found in Figure 3. The red points in Figure 3 represent the CS segments (treatment group) while the blue points represent the control segments.

It should be noted that the installation of Cycle Superhighways is not randomly assigned. There exist some confounding variables which could not only affect the construction of Cycle Superhighways, but also influence the cycle accident rate. Here, we consider the following covariates: the traffic flow volume, the bicycle flow volume, the number of previous accidents, population density, employees, IMD, bus density. The choice of covariates is based on empirical findings. The description of the covariates is in Table 2.

In this study, the treatment variable is binary, representing presence of a Cycle Superhighway or not. If there is no CS around a road segment in a pre-defined distance (e.g 1.5 kilometers), we assume that there is no construction of CS in this road segment and the treatment variable is 0, otherwise, it is 1. We define the pre-intervention period as the three years before the CS is open, the post-intervention period as the three years after the open time. The outcome variables we are interested, include average cycle accident rate, cycle flow volume, the number of cycle accidents over the post-intervention period. The average cycle accident rate is defined as the mean of yearly cycle accidents divided by AADB during the post-intervention period. The cycle flow volume is reflected by the average AADB during the post-intervention period. The number of cycle accidents is the total number of cycle related accidents that happened during the post-intervention period.

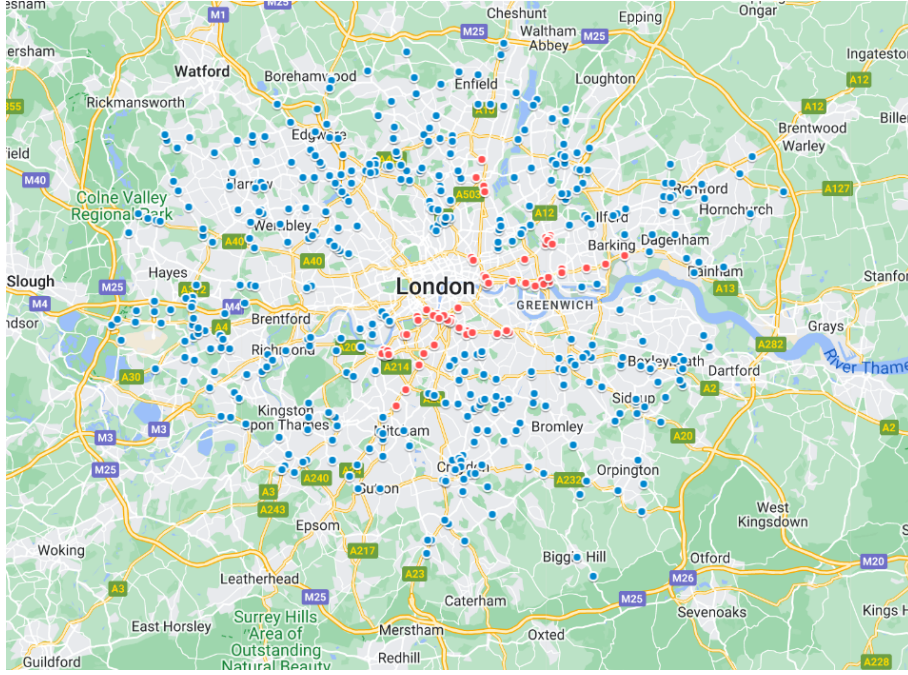


Figure 3: The distribution of treatment and control samples

variable	description	mean	std	min	max
AADB_pre	Annual average daily bicycle volume in the pre-intervention period	387.66	728.87	0.00	7167.33
AADT	Annual average daily traffic volume in the pre-intervention period	29666.31	24101.18	1825.66	202912.33
accident_pre	Total accident nums in the pre-intervention period	2.59	4.27	0.00	28.00
Population Density	population density in LSOA	75.77	51.06	1.16	437.04
Employees	Number of employees in LSOA	1468.48	5565.53	0.00	52000.00
IMD	The index of multiple deprivation	25.37	13.44	2.68	62.57
bus density	Number of bus stop within 0.5 km	13.38	6.47	0.00	35.00

Table 2: Descriptive statistics of the covariates

3 RESULTS AND DISCUSSION

In this section, we will estimate the effect of Cycle Superhighway (CS) on cycle accident rate, cycle volume and the number of cycle accidents. We will implement both PSM and panel outcome regression with fixed effects and compare the relative results. "MatchIt" package (?) in R is applied to perform the PSM. There are various kinds of matching methods in this package. Here, we use full matching because it performs quite well for the balance and overlap test. Next we will check the covariate balancing. The summary is in Table 3. Table 3 shows that all the variables are

	Std. Mean difference	M.Threshold	Variance Ratio
distance	0.0123	Balanced, <0.1	0.9936
AADB_pre	-0.0389	Balanced, <0.1	1.1077
AADT	0.0481	Balanced, <0.1	0.4973
accident_pre	0.0454	Balanced, <0.1	1.2428
Population.Density	-0.2404	Not Balanced, >0.1	1.3790
Employees	0.0606	Balanced, <0.1	0.3864
IMD	0.0386	Balanced, <0.1	1.1947
bus.density	-0.0576	Balanced, <0.1	0.7224

Table 3: Summary of balance for matched data

well-balanced except population density. The standardized mean difference of population density

is -0.2404. The absolute value is still not too big and population density may be considered as less important compared to other road characteristic variables. As a result, we can assume that the covariates achieve balance after matching.

Then, we check the overlap test by comparing the distribution plot of propensity score. The plot is in Figure 4. Figure 4 shows that before matching, the distributions of propensity score between

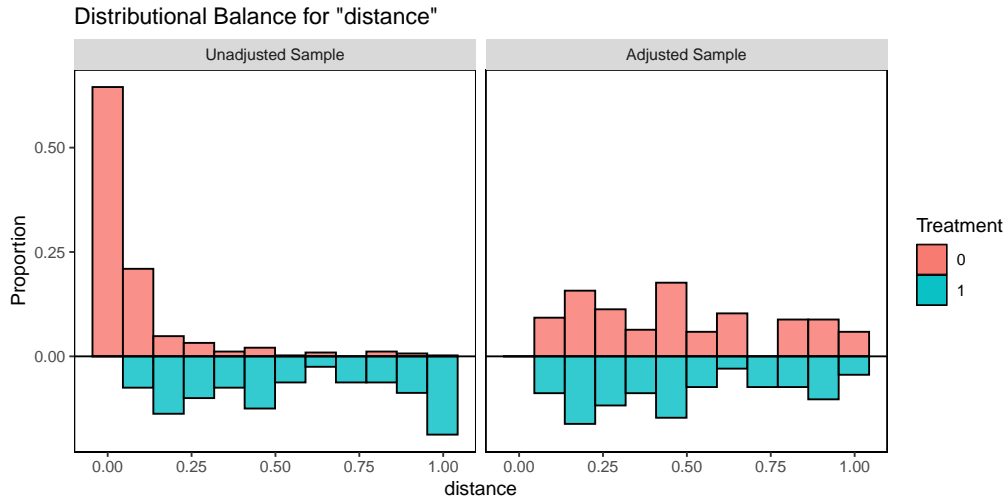


Figure 4: Overlap test

treatment and control groups are quite different. However, after matching they are similar to each other. Thus, there is support for the overlap assumption being verified.

Effect of CS on cycle accident rate

We can estimate the treatment effect ATE using the standard linear regression with matching weights. The coefficients and standard errors can be estimated by the `lmtest` (?) and `sandwich` packages (?) in R. First, we consider the effects of CS on cycle accident rate. The result is in Table 4.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	62.0057	32.2701	1.9215	0.0561
CS_or_not	11.0589	23.6364	0.4679	0.6404
AADB_pre	-0.0401	0.0130	-3.0847	0.0023
AADT	0.0009	0.0003	2.7053	0.0074
accident_pre	1.8630	1.0624	1.7536	0.0811
Population.Density	0.0455	0.1462	0.3113	0.7559
Employees	-0.0037	0.0015	-2.5259	0.0123
IMD	-1.5655	0.7903	-1.9808	0.0490
bus.density	0.5267	0.4907	1.0733	0.2845

Table 4: *t* test of coefficients (CS on cycle accident rate)

From Table 4, we can see that the *p*-value of CS_or_not is 0.6404, which indicates that the CS has no significant effect on the cycle accident rate. Then, we will use panel outcome regression with fixed effect to estimate the effect of CS on cycle accident rate. This can be achieved by using the `plm` package (?) in R. The result is in Table 5.

We can see that in Table 5, the *p*-value for CS_or_not is 0.51 and the R^2 is 0.00119. This means that the effect of CS on cycle accident rate is also not significant and the model does not fit well. The result of panel outcome regression is similar to the one of PSM. Both of them suggest that there is no significant effect of CS on cycle accident rate.

	Estimate	Std. Error	t-value	Pr(> t)
CS_or_not	1.50e+02	2.25e+02	0.67	0.51
AADT	9.74e-03	1.48e-02	0.66	0.51
R^2	0.00119			

Table 5: Summary of panel OR(CS on accident rate)

Effect of CS on the cycle volume

In this section, we will examine the effect of CS on the cycle volume. The result for PSM is in Table 6. From Table 6, the p -value for CS_or_not is 0.0029 and the estimated coefficient is 273.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.82e+02	1.78e+02	-1.02	0.3069
CS_or_not	2.73e+02	9.02e+01	3.02	0.0029
AADB_pre	1.11e+00	8.12e-02	13.68	<2e-16
AADT	-5.89e-04	7.69e-04	-0.77	0.4446
accident_pre	-1.74e+01	2.35e+01	-0.74	0.4593
Population.Density	3.22e-01	6.53e-01	0.49	0.6229
Employees	5.02e-03	2.72e-03	1.85	0.0666
IMD	4.00e+00	3.53e+00	1.13	0.2584
bus.density	-2.77e+00	6.43e+00	-0.43	0.6668

Table 6: t test of coefficients (CS on the cycle volume)

This indicates that the introduction of CS significantly increased the number of cycle volume. In average, it could increase 273 in the number of AADB.

The result of panel outcome regression with fixed effects is in Table 7. From Table 7, the p -value

	Estimate	Std. Error	t-value	Pr(> t)
CS_or_not	3.32e+02	9.38e+01	3.54	0.00065
AADT	6.52e-04	1.22e-02	0.05	0.95741
accident_num	-2.56e+01	1.89e+01	-1.36	0.17835
R^2	0.137			

Table 7: Summary of panel OR(CS on the cycle volume)

of CS_or_not is 0.00065 and the estimated coefficient is 332. This also indicates that the CS has a significant effect on the cycle volume. The result is consistent with the previous PSM result and the coefficient 332 is close to 273, which is the coefficient using PSM. As a result, we can conclude that CS significantly increased the cycle flow volume.

Effect of CS on the number of cycle accidents

We will next estimate the effect of CS on the number of cycle accidents. The result of using PSM is in Table 8. In Table 8, the p -value of CS_or_not is 0.0113 and the estimated coefficient is 1.64. This suggests that the CS significantly increased the number of cycle accidents.

The result of panel outcome regression is in Table 9. The p -value of CS_or_not is nearly 0 and estimated coefficient is 2.46. The result is also similar to the previous PSM result. As a result, both PSM and panel OR method suggest that the CS significantly increased the number of cycle accidents.

So far, we have found that overall, the CS significantly increase the cycle flow volume and the number of cycle accidents with no significant effect on the cycle accident rate. However, as shown in Table 2, each CS has different segregated condition. Among them, CS5 is the only fully segregated, while CS2, CS7, CS8 rarely have any segregation installation. Considering the heterogeneity in different CS routes, we will perform the causal analysis on each CS route and study the effect of segregation.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-4.69e-01	1.49e+00	-0.31	0.7533
CS_or_not	1.64e+00	6.42e-01	2.56	0.0113
AADB_pre	1.42e-03	5.46e-04	2.60	0.0101
AADT	-1.91e-05	8.51e-06	-2.25	0.0256
accident_pre	7.98e-01	8.29e-02	9.62	<2e-16
Population.Density	2.20e-03	4.72e-03	0.47	0.6417
Employees	9.16e-05	3.19e-05	2.87	0.0046
IMD	1.26e-02	2.22e-02	0.57	0.5690
bus.density	8.38e-02	4.52e-02	1.85	0.0654

Table 8: t test of coefficients (CS on cycle accidents)

	Estimate	Std. Error	t-value	Pr(> t)
CS_or_not	2.46e+00	2.22e-01	11.06	<2e-16
AADB	-7.15e-04	2.60e-04	-2.75	0.0062
AADT	-2.82e-05	1.39e-05	-2.03	0.0427
R^2	0.181			

Table 9: Summary of panel OR(CS on the cycle accidents)

The safety effect of different CS routes

In this subsection, we evaluate the impact of each CS route. The result using PSM is in Table 10. The result using panel outcome regression with fixed effects is in Table 11.

	Estimate	Std. Error	t-value	Pr(> t)
CS1	1.5860	3.4992	0.4532	0.6541
CS2	3.9906	2.8908	1.3804	0.1843
CS3	3.5691	4.7206	0.7561	0.4539
CS5	-5.6549	1.9912	-2.8400	0.0078
CS7	7.7870	5.2608	1.4802	0.1492
CS8	-2.1071	2.5961	-0.8116	0.4188

Table 10: Summary of the effect of each CS using PSM

	Estimate	Std. Error	t-value	Pr(> t)
CS1	-53.9521	335.71	-0.1607	0.8725
CS2	257.69	252.90	1.0190	0.3098
CS3	101.13	115.12	0.8785	0.3806
CS5	-16.7454	181.34	-0.0923	0.9265
CS7	353.40	274.82	1.2859	0.2003
CS8	141.57	227.91	0.6212	0.5353

Table 11: Summary of the effect of each CS using panel OR

In Table 11, the result using panel outcome regression still does not show any significance. From Table 10, we can see that although other CS routes still do not show any significant effect, the p -value of CS5 is less than 0.05 and the estimated coefficient is -5.6549 . This implies that CS5 significantly decreases the cycle accident rate. In Table 1, we can see that CS5 is the only CS route that is fully segregated. This indicates that the existence of segregation may be a crucial factor that influences the cycle accident rate.

The effect of Segregation on CS road

In this subsection, we will examine the effect of segregation on the CS road. Segregated cycle lane spared a space of the road for cycle use only. It is reported that there are lots of benefits of segregation. For example, the shift to segregated cycle lane can increase the carrying capacity of congested streets Aldred et al. (2017). Also, studies from Denmark have shown that segregated cycle lane reduces cyclists deaths by 35%⁷. Segregated cycle lanes are far more likely than those non-segregated ones to encourage people to cycle, especially women.

The characteristic of each road segment can be found at Google map. It not only provides the recent street view, but also provides the past few years' photos. With the help of Google map, we can denote each road segment as segregated or not. We also perform the PSM to inspect the effect of segregation on cycle accident rate. The result is in Table 12. From Table 12, we can see that

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-28.6089	131.5069	-0.2175	0.8284
CS_Seg	-58.8391	19.3721	-3.0373	0.0033
accident_rate_pre	0.6367	0.1043	6.1033	4.934e-08
AADT	0.0024	0.0021	1.1437	0.2565
Road.type	96.5759	76.8731	1.2563	0.2131
Population.Density	-0.2186	0.4600	-0.4752	0.6360
Employees	0.0116	0.0025	4.5338	2.294e-05
IMD	3.9610	0.8972	4.4145	3.543e-05
bus.density	-7.6020	1.9579	-3.8826	0.00022

Table 12: t test of coefficients(CS_seg vs CS_non_seg)

the p -value of "CS_Seg" is nearly 0 and the estimated coefficient is -58.839 . This suggests that the segregated Cycle Superhighways are significantly safer than those without segregation.

4 CONCLUSIONS

London Cycle Superhighways are a significant part of the "cycling revolution". In this paper, we studied the safety effect of Cycle Superhighways. 80 CS segments and 434 control segments were chosen to analyze. The covariates included annual average daily traffic (AADT), annual average daily bicycle volume (AADB), previous cycle accidents, population density, employees, index of multiple deprivation (IMD), and bus density. We implemented the propensity score matching and panel outcome regression with fixed effects to estimate the safety effect. Both of the methods showed that the installation of Cycle Superhighways has no significant effect on the cycle accident rate. However, they both indicated that the Cycle Superhighways significantly increased the cycle flow volume and the number of cycle accidents.

Then, we studied the heterogeneity of different Cycle Superhighways and found that CS5 performs best among these CS routes. CS5 is reported to be the only fully segregated CS. Thus, we next examined the effect of segregation among CSs. Using propensity score matching, it turned out that the segregated CS significantly decreased the cycle accident rate compared to those non-segregated CSs. As a result, in order to improve the safety of Cycle Superhighways, more segregation implementation should be encouraged.

For further study, one can consider more choices of the covariates. Limited by the data availability, we only considered AADT, AADB, previous cycle accidents, population density, employees, IMD, and bus density. One can also include some other plausible variables, e.g. traffic speed, intersection density.

Also, in this study, we constructed the model based on each "count point". This was because the dataset we got only had the latitude and longitude information. Thus, we just simply allocated each accident to its nearest count point. This might lead to some bias. A better choice is to do the reverse geocoding. That is, to map each coordinate to the corresponding road and construct the causal model base on each road. Then, each accident takes place exactly on its respective road.

Another possible improvement is to quantify the percentile of segregation for each road segment. In this paper, we only classified each road segment as segregated or not. To better inspect the

⁷<https://www.trafficchoices.co.uk/traffic-schemes/segregated-cycle-lanes.shtml>.

effect of segregation, one can quantify the percentile of segregation. In this case, the treatment variable is not binary any more. Instead, it becomes continuous. We would need to construct the propensity score matching with continuous treatment variable.

REFERENCES

- Aldred, R., Elliott, B., Woodcock, J., & Goodman, A. (2017). Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age. *Transport reviews*, *37*(1), 29–55.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31–72.
- DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of safety research*, *61*, 105–119.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605–654.
- Li, H., Graham, D. J., & Liu, P. (2017). Safety effects of the london cycle superhighways on cycle collisions. *Accident Analysis & Prevention*, *99*, 90–101.
- Pucher, J., & Buehler, R. (2017). Cycling towards a more sustainable transport future. *Transport reviews*, *37*(6), 689–694.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.
- Transport for London. (2011). *Barclays cycle superhighways faqs*. (Tech. Rep.).
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

How do electric bikes affect the route choice of cyclists? A case study of Greater Helsinki

Khashayar Khavarian¹, Shaghayegh Vosough*², and Claudio Roncoli³

¹Ph.D. Candidate, Department of Civil Engineering, Sharif University of Technology, Iran

²Postdoctoral Researcher, Department of Built Environment, Aalto University, Finland

³Associate Professor, Department of Built Environment, Aalto University, Finland

SHORT SUMMARY

Cycling as a clean, green, and environmentally friendly mode of transportation plays a crucial role in society by fostering physical activity and a healthy lifestyle, reducing traffic congestion, and improving mobility. To create more efficient strategies for promoting cycling, there is a need to gain a better understanding of the influential factors on cyclists' route choice behaviour. Electric bikes (e-bikes) are an emerging technology that appeared to assist cycling by using battery-powered motors. Researchers consider e-bikes as an emerging technology with its most certain effect being easing up cycling. Hence, investigating individual route choice behaviour with respect to their bike type can unveil new insights for cycling promotion. To this end, we used data collected via a stated preference (SP) survey in Finland not only to investigate the factors affecting cyclists' route choice behaviour but also to compare the behaviour of e-bikers with regular bike users (r-bikers) in order to identify the changes that may happen by easing the pedalling fatigue due to the pedal-assist feature of e-bikes. Our results indicate that low interaction with traffic, fewer intersections, and separated bike facilities are the main factors unchanged to promote cycling among r-bikers and e-bikers. Furthermore, we compare the outputs of simple Logit models (SLMs) and random parameter Logit models (RPLMs) for r-bikers' and e-bikers' route choices to address the impact of error correlation among observations in SP data. Our findings imply that the SP data is well-designed to capture the preferences of the individuals accurately, so the observations are not severely correlated, i.e., the IID assumption is held. This suggests that using SLMs can lead to similar outputs with RPLMs, without increasing the complexity of the estimation process. **Keywords:** Cycling, E-bike, Route choice, Discrete choice modelling.

1 INTRODUCTION

Motivation

Promoting active modes of travel provides many advantages, from decreasing air pollution to declining obesity cases and related diseases (Salabun et al., 2019; Anderson et al., 2022). One popular active mode is cycling, which can be used for almost all purposes, including with kids and for the elderly. With advancements in technology, there are solutions available that are believed to decrease the physical strain of cycling. Pedal-assisted electric bikes (in short *e-bikes*) can help the cyclist while pedalling, especially on routes with hills. It is proven to be effective on obese people evidencing that the physical and mental status of overweight people in Australia has improved after 12 weeks of e-bike cycling (Anderson et al., 2022).

It is worth noting that easing up the cycling pedalling is not enough to promote an active mode to the rest of the currently passive transport users unless planned properly. Although bicycle usage is promoted in many European countries, different patterns regarding cycling have been observed. For instance, France, Italy, and Germany have witnessed a more than 10% increase in cycling demand, in 2020, compared to 2019, while Finland and Ireland lost more than 10% of their weekday cyclist in the same period (Counter, 2021). One method to promote cycling in cities is believed to be planning the infrastructures for absorbing cyclists. Therefore, planners must know what measures affect cycling more (Broach et al., 2012; Huber et al., 2021).

A key cycling decision associated with transportation infrastructure is route choice. Researchers usually employ route-choice modelling for active modes (e.g., bicycles) to assess the infrastructure characteristics' impacts on the mode users (Segadilha & Sanches, 2014; Bernardi et al., 2018).

Studies using these models conclude that length, maximum steepness, and the type of road significantly affect cyclists' route choice. It can be concluded that safety and pedalling fatigue are the main concerns reducing the cycling demand when the infrastructure, traffic laws, and affordability of bikes are in place (Hull & O'Holleran, 2014).

Background

The literature on bicycle route choice is pretty rich and comprehensive reviews of influential factors on the cyclists' route choice behaviour can be found in previous studies (e.g., see Hull & O'Holleran (2014), Tarkkala (2022), Huber et al. (2021), and Tarkkala et al. (2023)). Studies dedicated to route choice of e-bikers or the change in the attitude of cyclists using e-bikes are limited but provide interesting insights. Chavis & Martinez (2021) found that e-bikes increase the length that cyclists ride, while they also reported shorter travel times for e-bikers than regular bike users (in short r-bikers), which means a significant increase in speed is observed. Moreover, with the increase in e-bike numbers, major roads were more frequently selected by cyclists than minor roads.

Rérat (2021) surveyed more than 2000 e-bikers and almost 11000 r-bikers, revealing an increased usage of e-bikes by females (50% of e-bikers vs 40% of r-bikers), as well as an increase in average age in the e-bikers. Regarding the season of cycling, it was observed that the e-bikers were almost abandoning their bikes in winter, switching to public transport or other motorised modes, probably due to the fact that e-bikes are used more frequently for longer trips than r-bikes.

Dane et al. (2020) provided mixed Logit models for r-bikers and e-bikers that show different factors affecting their route choices. However, their study results in favour of longer trips for both e-biker and r-bikers which is not in line with previous literature on the bike route choice models. They have used the interaction of different variables with length to account for differences in length for different groups of people which may have caused the positive sign of the length variable, however, no clear effect was found. They stated that the positive sign may be caused by the alternative generation algorithm they employed to generate the shortest paths. However, their findings regarding differences in route choice of r-bikers and e-bikers refer to variables that cannot be affected by specific policies (e.g., they found that daylight and weekday play major roles in r-bikes and e-bikes usage).

Research Contributions

According to the background section, little attention has been paid to the e-bike's effect on the route choice of cyclists. Although general studies are available, they have not assessed the change due to electrification and did not compare their results with route choice models with r-bikes. Simply put, studies regarding changes due to e-bike usage have not comprehensively resulted in the main variables responsible for r-bike and e-bike promotion. Hence, this research addresses this gap in the body of literature that, to the best of the authors' knowledge, has not been explored before.

Accordingly, in this research, the effects of e-bikes on the route choice decision of cyclists are investigated. We evaluate e-bikes' effects on different aspects of cycling and route choice using separate discrete choice models developed for cyclists with r-bikes and e-bikes. We evaluate the effects of various factors on cyclists' route choice while the pedalling is eased up by e-bikes. This may cause some factors to have a decreased importance in the route choice. In fact, a contribution of this research is to test the hypothesis that e-bikes change the important factors affecting cyclists' route choice decisions that may be used in infrastructure planning.

The above-mentioned contribution is obtained by comparing bikers' and e-bikers' route choice models, estimated using one source of data. The data used for this research is obtained through a stated preference (SP) survey that provides us with the chance to analyze the findings regarding route choice model specifications. The model specification may interfere with factors' effects on route choice. Two different types of models, i.e., simple Logit model (SLM) and random parameter Logit model (RPLM), are estimated for each type of bike. Comparing the models including different variables' significance depicts the impacts of model specification on our main findings. Using an SLM requires the errors to be independently and identically distributed (IID). On the other hand, the error term in the RPLM is not bounded to these assumptions. As in this research, each choice situation presented to respondents is considered an observation, the error terms of discrete choice models may not be IID (Axhausen et al., 2006). Moreover, similar types of models are used in (Meister et al., 2022) and for many other studies that are looking for the model specification effects in their results (Brownstone et al., 2000).

Thus, the contribution of this research to the literature is threefold:

1. investigating factors affecting route choice of e-bikers;
2. comparing r-bikers and e-bikers to identify the main affecting factors of cycling promotion with less pedalling fatigue; and
3. analyzing the impacts of model specification on research findings.

The remainder of this paper presents the data and method we employ to investigate the route choice behaviour of cyclists in Section 2; the outputs of the estimated models and the results' interpretations in Section 3; Finally, conclusions are drawn in Section 4.

2 METHODOLOGY

Data Collection

To analyze the route choice of bikers concerning the technology of their bikes, and bicycle route choice data, this study uses the SP data collected in Greater Helsinki, Finland. The study area is a collaborative region of 14 municipalities and the Siuntio Municipality which has around 1.53 million population with 1.20 million living in the capital area. More details about the data can be found in Tarkkala (2022). The data is gathered using a survey assessing the following general factors: the presence or type of a bike facility, the road type, the vehicle traffic, the presence of controlled intersections along the route, the route gradients, and its length. The survey was offered online for one month during September 2021 and 1029 respondents filled out the questionnaire. Figure 1 depicts one of the hypothetical choice situations used in the survey.

5. Which route would you choose? *

Choose the desired alternative by pressing on it.

A route which most of the way follows main streets on a separated cycle path. Other factors are

- 2 light-controlled intersections,
- 1/5 of the trip has moderate uphill,
- length 4 km.



A route which most of the way follows arterial roads on an adjacent cycle path. Other factors are

- substantial traffic volume,
- 2 light-controlled intersections,
- no hills,
- length 3 km.



Figure 1: An example of a choice situation used in the survey

The characteristics of the sample population including trip purposes, age groups, their experience in riding a bike, and the time of year they bike, are depicted in charts of Figure 2. The e-bikers share of the respondents is almost 9.6% which is similar to the reported share from the market, i.e., 9% (Kuva, 2020). Moreover, the share of female respondents from the filled questionnaire is 49.3% which is a fair share regarding the target society composition.

Method and Models

One of the common approaches to identify factors affecting route choice decisions is implementing the discrete choice models. In this research, as said before, two different types of discrete choice models are implemented: SLM and RPLM, which enables to investigate the model specification impacts by comparing the models' results.

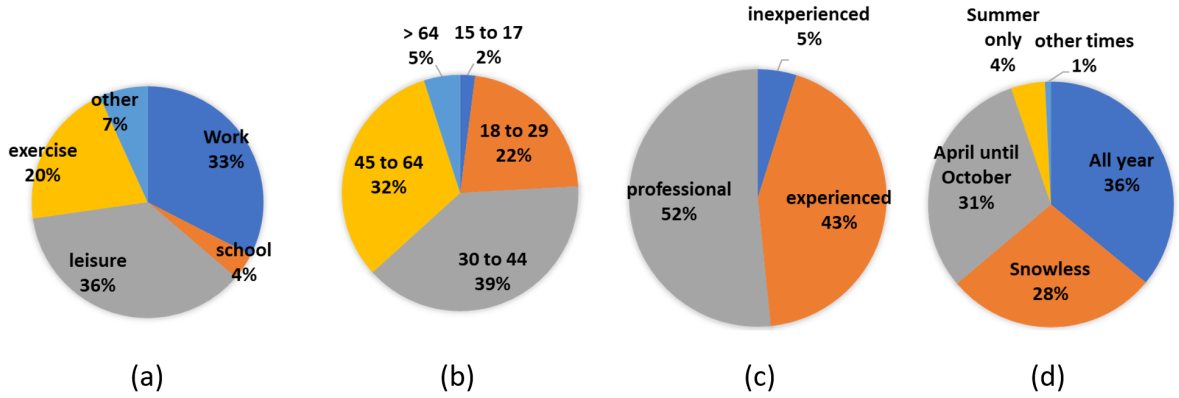


Figure 2: Understudy population characteristics; (a) Trip purpose, (b) Age, (c) Cycling experience, and (d) Time of the year cycling

SLM estimates the probability of choosing each route based on a linear combination of factors forming a utility value, as shown in Eq. (1). Since it is impossible to capture completely the utility value; a utility function composed of two parts is employed: the deterministic part, V_{in} , and the random/error part, ϵ_{in} , where i and n refer to alternative and individual, respectively. The deterministic part of the utility is a linear combination of effective factors in which β_{ik} is the coefficient related to k^{th} variable representing individual or alternative characteristics, X_{ink} . Then, the probability of each alternative selection is derived using Eq. (2).

$$U_{in} = V_{in} + \epsilon_{in} = \beta_i + \beta_{i1}X_{in1} + \beta_{i2}X_{in2} + \dots + \epsilon_{in} \quad (1)$$

$$P_{in} = \frac{\exp(V_{in})}{\sum_j \exp(V_{jn})} \quad (2)$$

The RPLM relaxes the IID assumption by introducing a random term, which eventually changes the error term to $\delta_{i1}X_{in1} + \epsilon_{in}$, as shown in Eq. (3). Different assumptions regarding the distribution of δ_{i1} is possible, with the most common to be a normal distribution.

$$U_{in} = V_{in} + \epsilon_{in} = \beta_i + \beta_{i1}X_{in1} + \delta_{i1}X_{in1} + \beta_{i2}X_{in2} + \dots + \epsilon_{in} \quad (3)$$

This change in error term distribution would be addressed through simulation and the expected probability of each alternative selection is derived by approximately estimating the result of Eq. (4) (Train, 2009).

$$P_{in} = \int_{-\infty}^{\infty} \frac{\exp(V_{in})}{\sum_j \exp(V_{jn})} f(\delta_{i1}) d\delta_{i1} \quad (4)$$

We estimate both models for r-bikers and e-bikers separately. The maximum likelihood method, which is used for model calibration, estimates the covariance matrix of coefficients as well. The results obtained through all four models are then compared to show the differences between r-bikers and e-bikers as well as the significance of the impact of error correlation among observations.

3 RESULTS AND DISCUSSION

Two sets of models (SLM and RPLM) are calibrated using Stata 17 (StataCorp, 2021) for r-bike and e-bike route choices, and their results are presented in Table 1. For each type of model, two models are presented for e-bikes' route choice. The first one includes the same variables as the r-bike model while the second one is without the variables found to be insignificant in the first model. In RPLMs, the random parameter is the coefficient of the route's length variable with the normal distribution.

Then, Logit models of r-bikers and e-bikers are compared to identify the prominent factors affecting the route choice of individuals, while the pedalling fatigue of cycling is removed due to the electrification of bikes. Another comparison is made between SLMs and RPLMs for regular bikes and e-bikes to evaluate the effect of error correlation on the effective route choice factors.

Table 1: SLMs and RPLMs for e-bikers' and r-bikers' route choice

Row	Variables	SLM			RPLM		
		R-Bikers	E-bikers (1)	E-bikers (2)	R-Bike	E-bike (1)	E-bike (2)
1	Route consists of main streets	0.102** (2.22)	-0.01 (-0.07)	-	0.411*** (6.80)	0.103 (0.60)	-
2	Route consists of arterial streets	0.253*** (4.04)	0.024 (0.14)	-	0.366*** (6.16)	-0.084 (-0.48)	-
3	Route is mixed with vehicular traffic	-0.846*** (-12.36)	-0.731*** (-3.78)	-0.873*** (-6.35)	-1.089*** (-12.31)	-0.848*** (-3.30)	-1.010*** (-4.21)
4	Route is on a bike lane	-0.221*** (-3.15)	0.234 (1.18)	0.293** (2.15)	-0.161* (-1.80)	0.447* (1.64)	0.481*** (1.80)
5	Route is on a separated adjacent path	-0.203*** (-3.75)	0.024 (0.15)	-	-	-	-
6	Moderate traffic near bike facility	-	-	-	-0.501*** (-7.52)	-0.500** (-2.37)	-0.485** (-2.28)
7	Heavy traffic near bike facility	-0.749*** (-10.85)	-0.663*** (-3.38)	-0.791*** (-5.27)	-1.678*** (-17.55)	-1.542*** (-5.46)	-1.639*** (-7.15)
8	Substantial traffic near bike facility	-0.699*** (-8.98)	-0.703*** (-3.18)	-0.787*** (-4.79)	-1.195*** (-15.40)	-1.161*** (-5.21)	-1.139*** (-5.33)
9	Route has controlled intersections	-0.322*** (-15.14)	-0.332*** (-5.44)	-0.316*** (-5.56)	-0.308*** (-11.09)	-0.415*** (-5.25)	-0.434*** (-5.42)
10	Route has hills	-0.661*** (-17.45)	-0.439*** (-3.85)	-0.433*** (-3.81)	-1.103*** (-15.17)	-0.802*** (-4.14)	-0.832*** (-5.38)
11	2 nd variable and being female	-0.246*** (-2.91)	-0.479* (-1.80)	-0.587*** (-2.94)	-0.561*** (-6.00)	-0.520* (-1.65)	-0.485** (-2.16)
12	3 rd variable and being female	-0.430*** (-5.08)	-0.217 (-0.77)	-	-	-	-
13	4 th variable and being female	0.153* (1.77)	0.110 (0.38)	-	-	-	-
14	7 th variable and being female	-0.236** (-2.53)	-0.458 (-1.46)	-	-0.477*** (-4.26)	-0.500 (-1.30)	-
15	8 th variable and being female	-0.200** (-2.04)	-0.230 (-0.77)	-	-	-	-
16	9 th variable and being female	0.120*** (4.76)	0.199** (2.45)	0.148** (2.34)	0.132*** (3.53)	0.241** (2.05)	0.236** (2.06)
17	10 th variable and being female	-	-	-	-0.140 (-1.47)	-0.164 (-0.56)	-
18	3 rd variable and being older than 65	-0.611*** (-2.96)	-0.451 (-0.97)	-	-	-	-
19	4 th variable and being older than 65	-0.490** (-2.37)	0.497 (1.17)	-	-	-	-
20	7 th variable and being older than 65	-	-	-	-0.717** (-2.33)	0.250 (0.45)	-
21	9 th variable and being older than 65	0.129** (2.16)	0.017 (0.14)	-	-	-	-
22	10 th variable and being older than 65	-	-	-	-0.356 (-1.40)	0.226 (0.52)	-
23	Route length	-0.311*** (-22.48)	-0.269*** (-6.28)	-0.267*** (-6.39)	-1.000*** (-25.65)	-0.929*** (-7.48)	-0.942*** (-7.56)
24	Standard deviation of the route length	-	-	-	0.599*** (2.48)	0.700*** (0.02)	0.701*** (-2.29)
25	Constant	2.850*** (31.33)	2.359*** (8.63)	2.379*** (9.03)	-0.101** (-2.48)	0.119 (0.02)	-0.270** (-2.29)

*, 90%, **, 95%, ***, 99%

R-bikers vs. E-bikers

Both sets of models verify the previous findings in the literature regarding the route choice behaviour of r-bikers. These findings consist of the negative influence of length and steepness on the selection probability of a route. Moreover, less interaction with traffic through low adjacent traffic and the provision of completely separated bike facilities are the main factors that remained effective in r-bikers' and e-bikers' route choices. Some new insights are also observed for r-bikers. Female cyclists are avoiding vehicular traffic and prefer controlled intersections in their routes more than men. A similar attitude is observed for r-bikers older than 65 years.

On the other hand, e-bikes provide ease and a sense of confidence for cyclists that changes vehicle avoidance preferences. The changes are clearly observable in previously cautious and maybe vulnerable cyclists like females and old people. Since the e-bike route choice models can make no distinction based on gender and age among e-bikers. For instance, we observed that r-bikers older than 65 years find the traffic disturbing more than other r-bikers while e-bikers older than 65 years do not get bothered by heavy traffic. A similar attitude towards traffic situations is observed in female cyclists. Some other differences e-bikes make in cyclist route choice behaviour can be concluded as:

- The male cyclists' preference towards main streets mitigates due to e-bikes while the corresponding coefficient for females stays the same (negative) as r-bikers.
- The 65 years old and older cyclists riding e-bikes are not affected by the presence of hills anymore.
- Although the length of the trip and hills are significant factors for both r-bikers and e-bikers, yet, as expected, the impacts of these variables are much milder for e-bikers.
- E-bikers prefer bike lanes, which is not the case for r-bikers.
- Female r-bikers have significant preferences for traffic avoidance, compared to men, while all e-bikers do not like heavy and mixed traffic almost similarly.

Furthermore, the results of the random parameter of the RPLMs indicate that the taste variation is significantly present in both r-bikers and e-bikers' route choices, due to the large value of the standard deviation coefficient. However, the confidence interval for e-bikers $([-2.31, 0.43])$ shows more dispersion than the r-bikers $([-2.17, 0.17])$, implying that the e-bike increases the variation of people's opinions toward the length of cycling.

SLM vs. RPLM

In general, the outputs of the two types of models are quite aligned, especially for e-bikers, and RPLMs results verify the findings from SLMs. In comparison to the RPLMs, more variables are found significant in the SLM for the r-bikers' route choice model (e.g., variables in rows 5, 12, 13, 15, 18, 19, and 21). Besides, there are a few significant variables in the RPLMs that are not found significant in the SLM for r-bikers (e.g., variables in rows 6 and 20). These differences lead to the following conclusions:

- The effect of substantial traffic on route choice is significant for female r-bikers with SLM, whereas it is not significant using the RPLM.
- Controlled intersections are favourable for r-bikers older than 65 based on SLM which is not significant using RPLM.
- Female and elderly r-bikers, based on SLM, are reluctant to cycle in routes with mixed traffic, while RPLM does not confirm the significance of these interactive variables.
- Interestingly, no significant difference is observed between the two types of models for e-bikers.

It should be noted that no significant difference is found between the coefficients' signs of the two types of models. These findings demonstrate that the SP survey is designed properly to capture the preferences of the individuals so that the errors in the responses are not severely correlated. This is why there are no substantial differences between the two models' outputs.

4 CONCLUSIONS

The differences in route choice behaviour between r-bikers and e-bikers are investigated in this research. Two sets of models, SLM and RPLM, are estimated based on data gathered through a stated preference survey. Both sets of models verify the previous findings in the literature regarding the negative influence of length and steepness on the selection probability of a route by r-bikers. Riding an e-bike, on the other hand, reduces the importance of the length of the trip and steepness, and e-bikers care less about the type of facility, i.e., major or minor streets, that they are cycling along.

From the transportation planners' point of view, providing dedicated routes with no interruptions from vehicular traffic can be introduced as the main effective factor in bike promotion. We observed that women like to cycle in a completely dedicated path with signalized intersections that minimizes the probability of colliding with other vehicles. Therefore, there is a trade-off between vehicles and bike volumes.

Regarding the model specification, we realized that there is no substantial difference between SLMs and RPLMs for e-bikers, in our case. This implies that the errors in the responses are not severely correlated and can be assumed to possess the IID character. If the IID assumption holds, it is considered to be a desirable property of the SP data, meaning that despite the hypothetical situations and panel effect, respondents' preferences do not affect the error terms. Hence, the SLMs without increasing the complexity in the estimation process can lead to similar outputs with RPLMs.

A major limitation of this research (shared with previous literature as well) is that respondents are already cyclists; hence, the results cannot be simply used for addressing non-bikers about their preferences and obstacles towards biking. However, focusing on cyclists is needed in this research due to the fact that we were looking for the differences created by e-bike implementation.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Helsinki Region Transport Authority (HSL) for their support in collecting and providing the data used in this study. We also thank Konsta Tarkkala and Jens West for their constructive comments and help with the data. The work of Shaghayegh Vosough and Claudio Roncoli was partially funded by the FinEst Twins Center of Excellence (H2020 Grant 856602) and the Academy of Finland project AiforLEssAuto (no. 347200).

REFERENCES

- Anderson, C. C., Clarkson, D. E., Howie, V. A., Withyman, C. J., & Vandelanotte, C. (2022). Health and well-being benefits of e-bike commuting for inactive, overweight people living in regional australia. *Health promotion journal of Australia: official journal of Australian Association of Health Promotion Professionals*, 33(S1), 349–357.
- Axhausen, K. W., Hess, S., König, A., Abay, G., Bates, J., & Bierlaire, M. (2006). State of the art estimates of the Swiss value of travel time savings. *Arbeitsberichte Verkehrs-und Raumplanung*, 383.
- Bernardi, S., Geurs, K., & Puello, L. L. P. (2018, oct). Modelling route choice of Dutch cyclists using smartphone data. *Journal of Transport and Land Use*, 11(1), 883–900. Retrieved from <https://www.jtlu.org/index.php/jtlu/article/view/1143> doi: 10.5198/JTLU.2018.1143
- Broach, J., Dill, J., & Gilebe, J. (2012, dec). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46(10), 1730–1740. doi: 10.1016/J.TRA.2012.07.005
- Brownstone, D., Bunch, D. S., & Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B: Methodological*, 34(5), 315–338.

- Chavis, C., & Martinez, V. (2021). *E-Bikes' Effect on Mode and Route Choice: A Case Study of Richmond, VA Bike Share* (Tech. Rep.). Urban Mobility Equity Center, Morgan State University.
- Counter, E. (2021). *Bike count dashboard: tracking the growth of cycling by country*. Retrieved from <https://www.eco-counter.com/cycling-data-tracker/>
- Dane, G., Feng, T., Luub, F., & Arentze, T. (2020). Route choice decisions of E-bike users: Analysis of GPS tracking data in the Netherlands. *Lecture Notes in Geoinformation and Cartography*, 109–124. Retrieved from https://link.springer.com/chapter/10.1007/978-3-030-14745-7_7 doi: 10.1007/978-3-030-14745-7_7/COVER
- Huber, S., Lißner, S., Lindemann, P., Muthmann, K., Schnabel, A., & Friedl, J. (2021, jun). Modelling bicycle route choice in German cities using open data, MNL and the bikeSim web-app. *2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2021*. doi: 10.1109/MT-ITS49943.2021.9529273
- Hull, A., & O'Holleran, C. (2014, jan). Bicycle infrastructure: can good design encourage cycling? *Urban, Planning and Transport Research*, 2(1), 369–406. Retrieved from <https://www.tandfonline.com/doi/abs/10.1080/21650020.2014.955210> doi: 10.1080/21650020.2014.955210
- Kuva, H. (2020). *Smaller volumes but same e-bike trend in Finland*. Retrieved 2023-01-20, from https://digimagazine.bike-eu.com/market_reports/finland
- Meister, A., Axhausen, K. W., Felder, M., & Schmid, B. (2022). Route choice modelling for cyclists on dense urban networks. *Available at SSRN 4267767*.
- Rérat, P. (2021). The rise of the e-bike: Towards an extension of the practice of cycling? *Mobilities*. Retrieved from <https://www.tandfonline.com/doi/abs/10.1080/17450101.2021.1897236> doi: 10.1080/17450101.2021.1897236
- Sałałun, W., Palczewski, K., & Wątróbski, J. (2019). Multicriteria approach to sustainable transport evaluation under incomplete knowledge: Electric bikes case study. *Sustainability*, 11(12), 3314.
- Segadilha, A. B. P., & Sanches, S. d. P. (2014, dec). Identification of Factors that Influence Cyclists' Route Choice. *Procedia - Social and Behavioral Sciences*, 160, 372–380. doi: 10.1016/J.SBSPRO.2014.12.149
- StataCorp. (2021). *Stata Statistical Software: Release 17*.
- Tarkkala, K. (2022). *Application of cyclists' route choice preferences in travel demand modelling*. Finland.
- Tarkkala, K., Vosough, S., West, J., & Roncoli, C. (2023). Implementing cyclists' route choice models into travel demand modelling: A case study of greater helsinki. *Under review*.
- Train, K. E. (2009). *Discrete Choice Methods with Simulation* (2nd ed.). Cambridge: Cambridge University Press. Retrieved from <https://www.cambridge.org/core/books/discrete-choice-methods-with-simulation/49CABD00F3DDDA088A8FBFAAD7E9546> doi: DOI:10.1017/CBO9780511805271

Explainable predictions for real-time employee workload management in railway control rooms

Léon Sobrie^{*1}, Marijn Verschelde², and Bart Roets³

¹Ph.D. Candidate, Ghent University, Belgium

²Full Professor, IÉSEG School of Management, France

³Railway Scientist & Ph.D., Infrabel & Ghent University, Belgium

SHORT SUMMARY

Industry 5.0 targets a resilient, sustainable and human-centric European industry. A key initiative to reach this target is adopting a human-centric approach to digital technologies, which places the well-being of the worker at the center. As workload peaks/lows contribute to lower employee well-being, predictive employee workload analytics can empower management to undertake proactive prevention. For this purpose, we develop a real-time machine learning framework to predict and explain future workload. Our feature importance analysis demonstrates the value of human-machine interactions and partner workload exposure. The proposed 2-stage framework, inspired by deep Tobit models, is developed and implemented in an environment with a variable and imbalanced workload: the digital control rooms for railway traffic management of Infrabel, Belgium’s railway infrastructure company. The related application is tailored towards the managers, for whom it provides real-time and explainable insights.

Keywords: Analytics, explainable, operators, railways, real-time, workload

1 INTRODUCTION

The European Commission (EC) launched Industry 5.0 to complement Industry 4.0 by focusing on research and innovation that serves the transition to a resilient, sustainable and human-centric European industry (European Commission, 2022). A key initiative proposed by the EC is to adopt a human-centric approach toward digital technologies. We aim to contribute to this initiative by focusing on human workload in a highly digitized environment: the railway control rooms of Infrabel, Belgium’s railway infrastructure company. In this setting, millions of actions are taken on a monthly basis by traffic operators to control railway traffic in real-time.

Operations literature acknowledges monitoring of adequate employee workload as one of the objectives of the control room management (see e.g. Valls et al. (2009)). Adequate workload has to be considered in comparison to both overload and underload. High levels of workload (i) are connected to lower daily well-being (Ilies et al., 2010), (ii) lead to task preference for easier tasks hurting performance (Kc et al., 2020), (iii) are a causal factor for human fatigue (Li et al., 2020), (iv) induce quality degradation due to cognitive multitasking (Xu et al., 2022) and (v) are likely to trigger the health-impairing mental/physical conditions of individuals, which are often related to safety performance (Derdowski & Mathisen, 2023). Low levels of workload (i) result in extra performance-seeking risks (Xu et al., 2022) and (ii) lead to boredom and lack of attention (Young, 2021). Therefore, creating an environment with a balanced workload amongst and within employees contributes to their well-being and satisfaction (Inegbedion et al., 2020). Moreover, workload is a *multi-attribute concept* (Comstock Jr & Arnegard, 1992) which entails communication, resource management, automation, scheduling, monitoring and tracking. Hence, analytics for workload management should incorporate different attributes to provide granular and explainable insights.

Our contribution is fourfold. First, we contribute to the operations research literature, by showing the applicability of machine learning to provide real-time employee-centric predictive analytics. We answer the calls from literature to connect operations research with human resource management (Roels & Staats, 2021) and to bridge the gap between practice and research by constructing a

model that is embedded in a problem observed in practice (Ranyard et al., 2015)). Related literature on employee-centric predictions for French air traffic controllers (TC) postulates a neural network for workload prediction combined with a tree-based search model for optimal airspace partitions Gianazza (2010). We distinguish ourselves from the literature by our focus on granular explainability and our multi-attribute consideration of workload. Second, we contribute to the explainable artificial intelligence literature (Coussement & Benoit, 2021) by illustrating the usefulness for adequate explainability of sample selection as based on the unobserved predicted values, via our advocated two-stage methodology, inspired by Heckman (1979) and Zhang et al. (2021). Third, we contribute to the management science (MS) literature on human-machine interaction (see e.g., Brynjolfsson & McAfee (2014)) and team exposure (see e.g., Akşin et al. (2021)). Our Shapley analysis pinpoints the importance of including automation and team interplay in workload predictions. Last, we contribute to the literature on applications of smart data analytics (see e.g. Baesens et al. (2016)) through our real-time implementation. Not many systems that utilize machine and/or deep learning have been adopted in a real-world setting (Kraus et al., 2020). One of the exceptions utilizes regression trees to build real-time analytics on passenger flows in the control room of Heathrow (Guo et al., 2020).

The paper is structured as follows. In section 2, we elaborate on the 2-stage methodology. In section 3, we dive into the results of the accuracy and explainability of the proposed model, and show its implementation for real-time management in traffic control centers (TCCs). The last section formulates the main takeaways.

2 METHODOLOGY

Our methodological framework is developed to provide explainable and accurate predictive analytics on employee workload aggregated in 15-minute intervals. The idea is to dissect workload into different operational categories (see section 3). Methodologically, we are confronted with a classical sample selection issue (see Heckman (1979)). This is because we have no a priori reason to believe that the mechanisms affecting the presence of workload are the same as the mechanisms that affect the workload magnitude, when present. Recently, Zhang et al. (2021) have introduced the sample selection issue into the literature on deep learning via a two-stage ‘*deep Tobit model*’. Our proposed two-stage approach comprises an LSTM encoder-decoder model in stage 1 to select the workload categories and an XGBoost model in stage 2 to predict the amount of workload within the selected categories.

Stage 1: LSTM encoder-decoder for binary classification

We utilize a Long Short-Term Memory (LSTM) encoder-decoder model (Hochreiter & Schmidhuber, 1997; Cho et al., 2014) to predict the presence of workload. This approach leverages an input-to-output sequence data structure that incorporates information from previous intervals to predict the occurrence of different workload categories for future intervals. Moreover, the prediction model has a memory cell to store the past. This aligns with the call from Corman & Quaglietta (2015) to close the loop in real-time railway traffic control. The training set contains N (input sequence, workload sequence) pairs with K features per input sequence. Each pair combines T input feature sequences $x = (x_1, \dots, x_t, \dots, x_T)$, with $x_{t \in T} \in R^K$, together with M workload sequences $y = (y_1, \dots, y_m, \dots, y_M)$, with $y_{m \in M}$ a binary variable, representing the presence of workload in a category. Via the LSTM-based encoder-decoder model, we learn the conditional expectation $E(y_m | y_{m-1}, \dots, y_1, x_{T-1}, \dots, x_1)$. by estimating \hat{y} via minimizing squared error over the N training pairs:

$$\min_{\theta} \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (y_{m,n} - \hat{y}_{m,n}(\theta))^2, \quad (1)$$

with θ the model parameters. Operationalization implies the use of an encoder to model (x_T, \dots, x_1) into c . The latter serves as an input for the decoder model to learn the conditional expectation $E(y_m | y_{m-1}, \dots, y_1, c)$. By uncovering this conditional expectation, the encoder-decoder model learns to predict y_{m+1} , based on $y_{\leq m}$ and c .

Stage 2: XGBoost for quantifying selected workload categories

The extreme gradient boosting utilizes the gradient boosting algorithm which grows trees sequentially by minimizing a regularized objective function (Chen et al., 2015):

$$\sum_i l(\hat{y}_i, y_i) + \sum_j \Omega(f_j), \text{ where } \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2, \quad (2)$$

with l a differentiable convex loss function related to the difference between the prediction \hat{y} and the actual output y , Ω a regularization function, $i=1, \dots, N$ representing the input-output pair, $j = 1, \dots, J$ representing the tree, T the number of leaves in a tree, w the leaf weight, γ a user-defined pruning parameter, and λ a shrinkage parameter.

We leverage Shapley values to provide explainable insights for the proposed XGBoost model. Shapley values (Shapley, 1953) incorporate potential synergies between the features, by averaging the marginal contributions over all possible subsets and by taking the order of the marginal contribution into account (Lundberg & Lee, 2017).

3 RESULTS AND DISCUSSION

Empirical setting

All railway traffic in Belgium is managed in the digital control rooms of Infrabel, with every zone being monitored by one traffic and safety controller at any given time. The basis for the predictive model is a real-time data structure containing all actions taken in the control rooms, including over 5,000,000 tasks executed per month. In this paper, we focus on traffic controllers for the month of February 2022, aggregated per control room, workstation and 15-minute time interval.

We build on previous research (Topcu et al., 2019) that divides workload into 6 operational categories as presented in Table 1. Further, Table 1 contains the occurrence of zeros for each category. We notice that there is no workload in more than 40% of the considered 15-minute intervals, except for MOVE and AUT. This strengthens our empirical choice to propose a 2-stage methodology to first filter for which categories the operator will have workload and, thereafter, predict the workload magnitude in an explainable way. MOVE is excluded in the first stage as the occurrence of zero workload for this category is below 1%, making a binary filter redundant.

Table 1: Operational workload categories

Operational workload categories		Zero occurrence
MOVE	Monitoring of railway traffic by opening signals	0.7%
ADAPT	Reducing train delays by changing tracks and station platforms	40.7%
AUT	Changing the automation - Automatic Route Setting	16.3%
SAFETY	Safety interventions	78.2%
PHONE	Phone calls between the control room operator and train drivers	60.6%
JUSTIF	Justification of train delays	48.7%

Similarly to the operational categories, the considered features for our predictive model are grouped by the control room, workstation and 15-minute interval. The features for near-future workload prediction for the different categories consist of (i) the experience and training level of the operator, (ii) automation usage, (iii) trains monitored, (iv) delays, (v) current workload for each category, (vi) partner controller features, and (vii) control room and temporal fixed effects. The features are tested for potential multicollinearity to ensure meaningful explainability.

Accuracy of LSTM encoder-decoder for binary classification

The binary LSTM encoder-decoder uses 4 input sequences to predict 4 output sequences. We reach the lowest classification error for dropout=0.2, optimizer=Adam, learning_rate=0.0001, batch_size=25, and number_of_nodes=50. The Area Under the Receiver Operator Curve (AUC) values, provided in Table 2, shows the classification capability for the different workload categories.

Table 2: AUC values

	SEQ1	SEQ2	SEQ3	SEQ4
ADAPT	0.73	0.74	0.74	0.74
AUT	0.84	0.86	0.86	0.86
SAFETY	0.82	0.82	0.82	0.82
PHONE	0.78	0.78	0.78	0.77
JUSTIF	0.71	0.72	0.72	0.71

The AUC is above 0.7 for all categories, with over 0.84 for AUT and 0.82 for SAFETY. The stability of the AUC over the four output sequences demonstrates the classification power of stage 1. Figure 1 presents the confusion matrices for the different categories with the tuned thresholds.

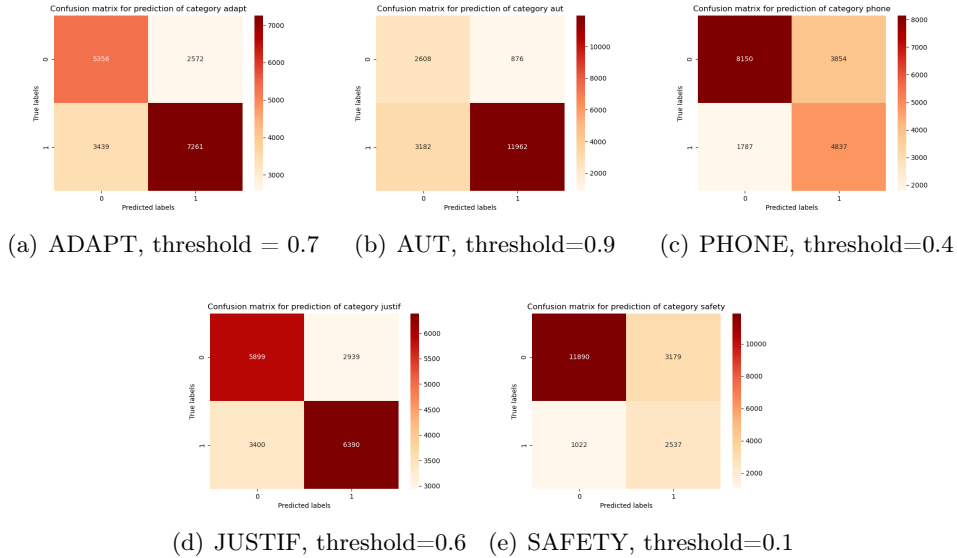


Figure 1: Confusion matrices

Explainability of extreme gradient boosting

The explainability analysis focuses on the first 15-minute interval in the future, as this is the most relevant horizon for real-time management. The XGBoost model reaches the lowest error for overall workload prediction with learning_rate=0.1, max_depth=3, number_of_estimators=100, percentage_sampled_columns_per_tree=0.4, and subsample_percentage_training_data=1. Table 3 presents the root mean square error (RMSE) and Spearman correlation.

Table 3: RMSE and Spearman correlation

	RMSE (in seconds)	Spearman correlation
MOVE	17.54	0.73
ADAPT	43.54	0.42
AUT	19.09	0.56
SAFETY	9.08	0.41
PHONE	66.00	0.42
JUSTIF	55.55	0.28

Figure 2 presents the most important features for predicting the overall workload. Automation usage is the top feature demonstrating the value of including human-machine interaction in explainable employee-centric analytics. In particular, TCs automate route setting in non-complex, low-workload situations to further reduce their workload (Balfe et al., 2015). Further, features on current workload have predictive power, next to experience and workstation fixed effects.

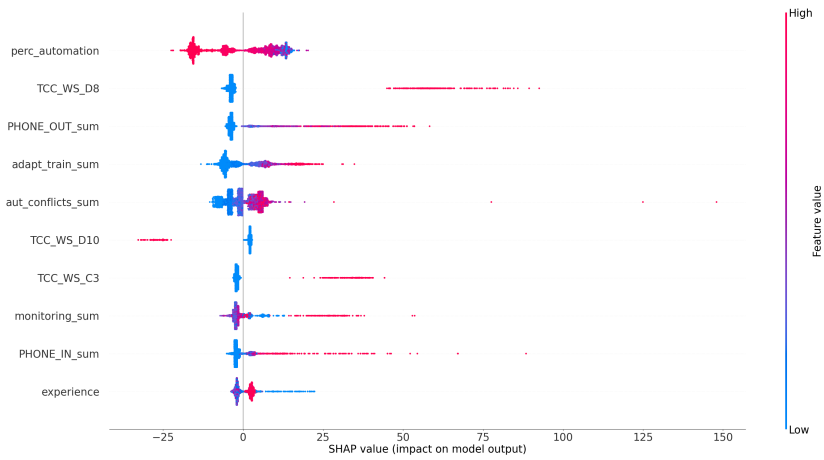


Figure 2: Global Shapley plot

Figure 3 demonstrates the importance of including the partner features for predicting the workload in the operational categories. This finding is in line with the previous MS research (Tan & Netessine, 2019) on the importance of the relationship between observed workload and partner characteristics.

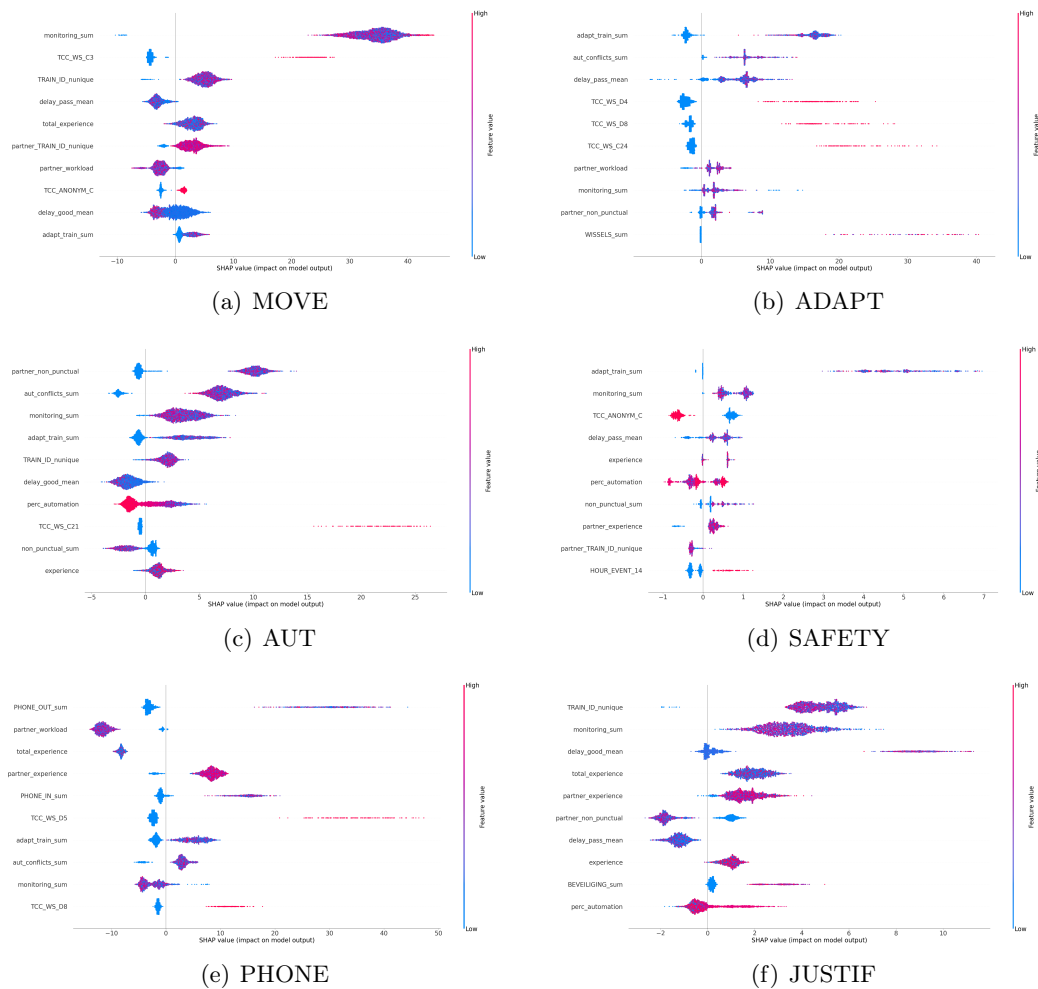


Figure 3: Global Shapley plots per workload category

Implementation

We implement the proposed prediction framework for control room managers in R Shiny. Figure 4 contains two panels with visualizations. Panel 4(a) presents the control room layout to more easily depict the over-/underloaded operators. Panel 4(b) provides explainable real-time insights into the expected near-future workload by breaking down the contribution of each feature. The real-time implementation showcases the usefulness of human-centric decision support in safety-critical settings.

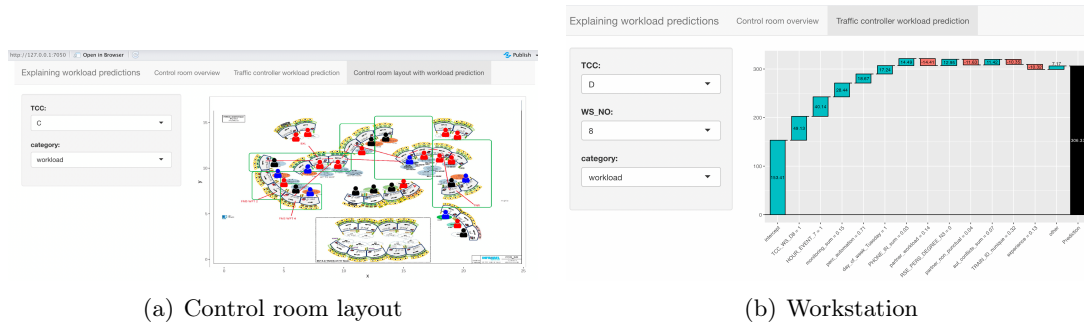


Figure 4: R Shiny implementation

4 CONCLUSIONS

The impact of workload and its imbalance between and within employees should not be underestimated, as the repercussions on well-being and satisfaction are noticeable. Highly digitized environments, such as railway control rooms, provide an opportunity to introduce data-driven decision support. We leverage this opportunity and propose explainable predictive analytics, customized for employee workload prediction. This is facilitated by our real-time data structure, covering all actions taken in the control rooms. Empirically, we show the binary classification capability of our proposed LSTM encoder-decoder model for the different workload categories. Next, we deploy Shapley values on the output of the XGBoost model, which demonstrates the importance of automation usage for near-future workload prediction. In addition, we unravel the impact of partner workload on the different workload categories, highlighting the value of including team dynamics in predictive analytics. Furthermore, an R Shiny application for the traffic supervisor deploys the proposed framework to provide near-future workload predictions in real time.

ACKNOWLEDGEMENTS

We would like to thank Infrabel for funding this research, preparing and implementing the datasets, and providing operational feedback. Special thanks to Jasmijn De Clercq, Pierre Gramme, Kristof Van der Strieckt, and the Fellows of the On Track Lab. The views expressed in this paper are those of the authors and do not necessarily reflect the opinions of Infrabel.

We would like to thank the participants of the EURO 2022 conference, the INFORMS 2022 Annual Meeting, and the 2022 On Track Lab workshop, for their comments.

Marijn Verschelde benefited from financial support from the French National Research Agency through the project BALANCE.

REFERENCES

- Akşın, Z., Deo, S., Jónasson, J. O., & Ramdas, K. (2021). Learning from many: Partner exposure and team familiarity in fluid teams. *Management Science*, *67*(2), 854–874.
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). Transformational issues of big data and analytics in networked business. *MIS quarterly*, *40*(4), 807–818.
- Balfe, N., Sharples, S., & Wilson, J. R. (2015). Impact of automation: Measurement of performance, workload and behaviour in a complex control environment. *Applied ergonomics*, *47*, 52–64.
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... others (2015). Xgboost: extreme gradient boosting. *R package version 0.4-2*, *1*(4), 1–4.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Comstock Jr, J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research* (Tech. Rep.).
- Corman, F., & Quaglietta, E. (2015). Closing the loop in real-time railway control: Framework design and impacts on operations. *Transportation Research Part C: Emerging Technologies*, *54*, 15–39.
- Coussement, K., & Benoit, D. F. (2021). Interpretable data science for decision making. *Decision Support Systems*, *150*, 113664.
- Derdowski, L. A., & Mathisen, G. E. (2023). Psychosocial factors and safety in high-risk industries: A systematic literature review. *Safety Science*, *157*, 105948.
- European Commission. (2022). *Industry 5.0 roundtable : Brussels 27 april 2022 : meeting report*. Publications Office of the European Union. doi: doi/10.2777/982391
- Gianazza, D. (2010). Forecasting workload and airspace configuration with neural networks and tree search methods. *Artificial intelligence*, *174*(7-8), 530–549.
- Guo, X., Grushka-Cockayne, Y., & De Reyck, B. (2020). London heathrow airport uses real-time analytics for improving operations. *INFORMS Journal on Applied Analytics*, *50*(5), 325–339.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.
- Ilies, R., Dimotakis, N., & De Pater, I. E. (2010). Psychological and physiological reactions to high workloads: implications for well-being. *Personnel Psychology*, *63*(2), 407–436.
- Inegbedion, H., Inegbedion, E., Peter, A., & Harry, L. (2020). Perception of workload balance and employee job satisfaction in work organisations. *Heliyon*, *6*(1), e03160.
- Kc, D. S., Staats, B. R., Kouchaki, M., & Gino, F. (2020). Task selection and workload: A focus on completing easy tasks hurts performance. *Management Science*, *66*(10), 4397–4416.
- Kraus, M., Feuerriegel, S., & Oztekin, A. (2020). Deep learning in business analytics and operations research: Models, applications and managerial implications. *European Journal of Operational Research*, *281*(3), 628–641.
- Li, F., Chen, C.-H., Zheng, P., Feng, S., Xu, G., & Khoo, L. P. (2020). An explorative context-aware machine learning approach to reducing human fatigue risk of traffic control operators. *Safety science*, *125*, 104655.

- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.
- Ranyard, J. C., Fildes, R., & Hu, T.-I. (2015). Reassessing the scope of or practice: The influences of problem structuring methods and the analytics movement. *European Journal of Operational Research*, 245(1), 1–13.
- Roels, G., & Staats, B. R. (2021). OM Forum-People-Centric Operations: Achievements and future research directions. *Manufacturing & Service Operations Management*, 23(4), 745–757.
- Shapley, L. (1953). A value for n-person games. In *Contributions to theory games (am-28)*. Princeton Univ. Press.
- Tan, T. F., & Netessine, S. (2019). When you work with a superman, will you also fly? An empirical study of the impact of coworkers on performance. *Management Science*, 65(8), 3495–3517.
- Topcu, T. G., Triantis, K., & Roets, B. (2019). Estimation of the workload boundary in socio-technical infrastructure management systems: The case of Belgian railroads. *European Journal of Operational Research*, 278(1), 314–329.
- Valls, V., Pérez, Á., & Quintanilla, S. (2009). Skilled workforce scheduling in service centres. *European Journal of Operational Research*, 193(3), 791–804.
- Xu, Y., Tan, T. F., & Netessine, S. (2022). The impact of workload on operational risk: Evidence from a commercial bank. *Management Science*, 68(4), 2668–2693.
- Young, M. S. (2021). In search of the redline: Perspectives on mental workload and the ‘underload problem’. In *International symposium on human mental workload: Models and applications* (pp. 3–10).
- Zhang, J., Li, Z., Song, X., & Ning, H. (2021). Deep tobit networks: A novel machine learning approach to microeconometrics. *Neural Networks*, 144, 279–296.

Identifying main drivers for students and staff members' mode choice or to work/study from home: A case study in Australia

Camila Balbontin*^{1,2}, John D. Nelson², David A. Hensher², Matthew Beck²

¹ Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Chile
Instituto de Sistemas Complejos de Ingeniería, Chile

² Institute of Transport and Logistics Studies, Business School, University of Sydney, Australia

SHORT SUMMARY

Universities are major trip attractors and generators in large cities, and they have a significant influence in the transport network particularly in high-density areas. The trips to and from university campus are made by staff, students and visitors, with an important daily rotation of people (e.g., students that leave early, arrive later, etc.). In this study, we aim to improve our understanding of the trips made to the University of Sydney campuses, one of the largest universities in Australia, how individuals (namely, staff and students) choose to study/work from home and their modes of transport used since the start of COVID-19. We have collected two sets of data in 2022 from a survey which was answered by both staff and students at the University of Sydney. A mixed logit model is estimated to understand the motivations and main drivers to work/study from home or to choose different modes of transport when attending campus.

Keywords: University travel choices; sustainable modes of transport; work/study from home; staff and student behaviour; choice modelling

1. INTRODUCTION

Since the start of COVID-19 in early 2020, the world has seen major changes in daily life. Different strategies from relevant authorities have led to different outcomes and impacts across numerous activities. Businesses, schools and universities have had to quickly adapt to this new normality by offering their employees and students the possibility to work and study from home. This quick adaptation has proven to have some benefits from the point of view of employers and employees who have stated they prefer to work from home (WFH) more often in the future (Beck & Hensher, 2021). The pandemic and the varied government strategies have had major impacts on commuting behaviour; at the start of the pandemic we saw a significant reduction in congestion around the city; however, since the start of 2022 we have seen levels of congestion that are close to pre-pandemic levels and, in some areas, even worse. Transport studies carried out in different cities around the world have shown that the use of public transport has decreased significantly, and, in their majority, these trips seem to have moved to car use (Balbontin et al., 2021; Hensher et al., 2022). This is a big setback for public transport where confidence has been slow to build back (Beck et al., 2022) and, more generally, for sustainable transport, which represents an important concern for transport authorities and government.

Universities are major trip attractors and generators in large cities, and they have a significant influence in the transport network particularly in high-density areas. The trips to and from university campus are made by staff, students and visitors, with an important daily rotation (e.g., students that leave early, arrive later, etc.). Despite the significant influence that universities have in the transport network, there is limited information on how trips to and from university campus are made, which modes of transport are preferred and why, and the effects that COVID-19 has had and will likely have in the medium to long-term in travel behaviour to campus. For the last three years, the University of Sydney, which is one of the largest universities in Australia with approximately 83,000 staff and students¹, has been offering a hybrid teaching model for a larger number of its courses, where students have face-to-face and online classes. However, it is moving to face-to-face classes exclusively towards mid-2023. The hybrid model is also being implemented for staff members, who are allowed to work from home for the whole or part of the week, when their role allows for it. However, there is still limited knowledge about the preferences towards working from home and towards studying from home (SFH), and how these have affected staff and students travel behaviour, principally in tertiary education.

In this study, we aim to improve our understanding of the trips made to the University of Sydney campuses, how often individuals (namely, staff and students) choose to study/work from home, and their modes of transport used since the start of COVID-19. We have collected two sets of data in 2022 from a survey which was completed by both staff and students. A mixed logit model is estimated to understand the motivations and main drivers to work/study from home or to choose different modes of transport when attending campus. The next section presents the data, followed by the methodology and model results and discussion. This short paper is finalised with the main conclusions.

¹ To put this in perspective, this figure would make the University of Sydney the 23rd largest city in Australia (behind Mackay whose population is 85,000 and above Rockhampton with a population of 80,200). It would be the 5th largest city in New South Wales behind Sydney, Newcastle, Wollongong, and Albury-Wodonga. It is larger than Wagga Wagga (57,000) and almost double the size of Orange (42,000).

2. DATA

The data used in this study was collected using an online survey in two waves. The first one was collected in May-June 2022 and the second one October-November 2022; a period where there existed no restrictions on movement but the University still offered education within the hybrid-format. The results discussed here only include students and staff members of the University of Sydney who lived in Sydney at the time of completing the survey. Table 1 represents some general descriptive statistics of the sample for both waves, separated by staff and students. As expected, the income and age of staff is significantly higher than students. Results show that students tend to live in households with more members – but slightly less children, and more cars/drivers licences available. Results in Table 1 show that in Wave 1, the total number of days/week studied from home last week (2.58) are higher than pre-COVID levels (1.86); and for staff members the increase is even higher, with an average of 2.3 days worked from home last week relative to 0.62 pre COVID-19. In Wave 2, the proportion of days worked / studied home is virtually unchanged from Wave 1 for students (previously 2.58, now 2.60) and has decreased slightly for staff (previously 2.31, now 2.15).

Table 1: General descriptive statistics students and staff - mean (standard deviation) - Waves 1 and 2

	Student		Staff	
	Wave 1	Wave 2	Wave 1	Wave 2
Age (years)	24.23 (7.31)	23.51 (6.10)	43.29 (11.43)	43.28 (11.49)
Gender female (1,0)	75%	65%	70%	70%
Personal annual income (000AUD\$)	22.08 (25.77)	24.00 (33.74)	125.17 (61.84)	125.29 (58.97)
Number of cars available in household	1.57 (1.23)	1.38 (1.68)	1.35 (0.84)	1.36 (0.90)
Number of people living in same household	3.48 (2.40)	3.82 (11.63)	2.75 (1.28)	2.85 (1.32)
Number of children in household	0.52 (0.80)	0.44 (1.16)	0.57 (0.86)	0.65 (0.93)
Number of drivers' license in household	2.33 (1.34)	2.34 (8.84)	1.90 (0.85)	1.93 (0.88)
Total number of weekly days worked/studied last week	5.48 (1.41)	5.59 (1.40)	4.97 (0.94)	4.97 (0.91)
Total number of weekly days worked/studied from home last week	2.58 (1.90)	2.60 (1.82)	2.31 (1.61)	2.15 (1.59)
Total number of weekly days worked/studied from campus last week	2.06 (1.69)	2.13 (1.61)	2.44 (1.63)	2.66 (1.56)
Total number of weekly days worked/studied partly from home and campus last week	0.84 (1.31)	0.86 (1.38)	0.21 (0.65)	0.16 (0.58)
Total number of weekly days worked/studied before COVID-19	5.82 (1.29)	6.03 (1.31)	5.11 (0.88)	5.14 (0.94)
Total number of weekly days worked/studied from home before COVID-19	1.86 (1.65)	2.46 (1.91)	0.63 (1.16)	0.68 (1.26)
Total number of weekly days worked/studied from campus before COVID-19	3.20 (1.77)	2.70 (1.77)	4.30 (1.29)	4.35 (1.36)
Total number of weekly days worked/studied partly from home and campus before COVID-19	0.76 (1.42)	0.87 (1.45)	0.17 (0.67)	0.12 (0.52)
Considered moving residence on the basis of a change in your activity travel pattern as a result of COVID-19 (1,0)	29%	26%	27%	23%
Total number of respondents	133	1,171	496	364

Figure 1 shows work/study behaviour for each day of the week, and Figure 2 shows the mode chosen to go to the campus. Results show that students are more likely to study partly from campus and from home than staff members (with Wave 2 mirroring Wave 1), and they are also more likely to do some study during the weekends than staff members. Regarding the modes used, staff

members are currently much more likely to use their car to go to campus, and students are more likely to use public transport and active modes. Staff appear to have increased their car use compared to pre-COVID. Compared to students, car use amongst staff is picking up quite rapidly compared to pre-COVID, while use of active modes remains static. These results reveal the behavioural differences between staff members and students.

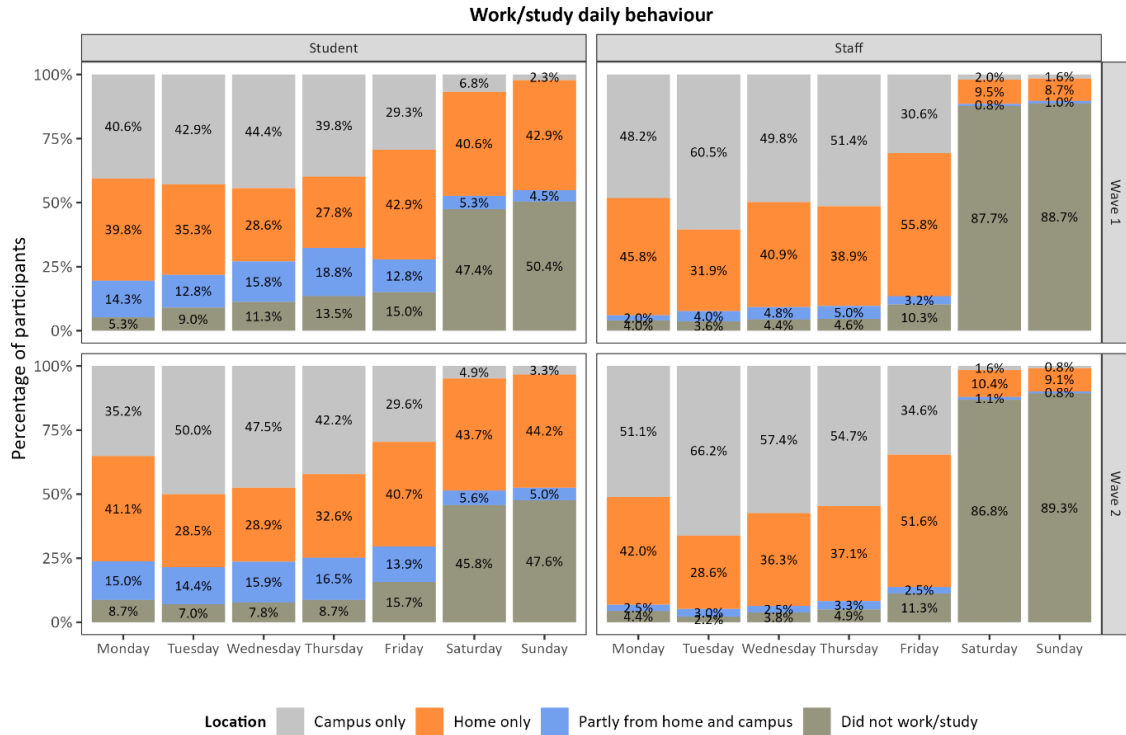


Figure 1: Work/study daily behaviour

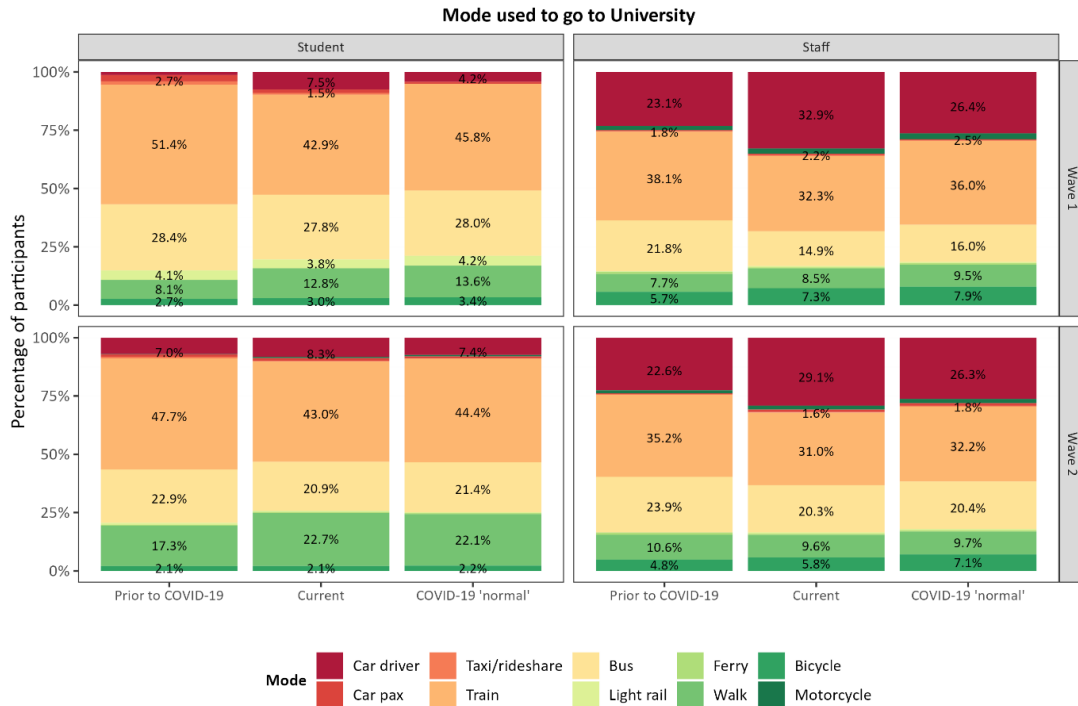


Figure 2: Mode used to go to University

In terms of activities which influence the decision to come into the university, Figure 3 shows that in both Waves 1 and 2, students are motivated to come to campus to participate in face-to-face classes or attend a class or feel that they learn more effectively when on campus relative to the start of the pandemic. They are also interested in building networks and meeting new people. Staff come to campus to participate in face-to-face discussion and also for a change from WFH with a very similar pattern reported in both waves (Figure 3).

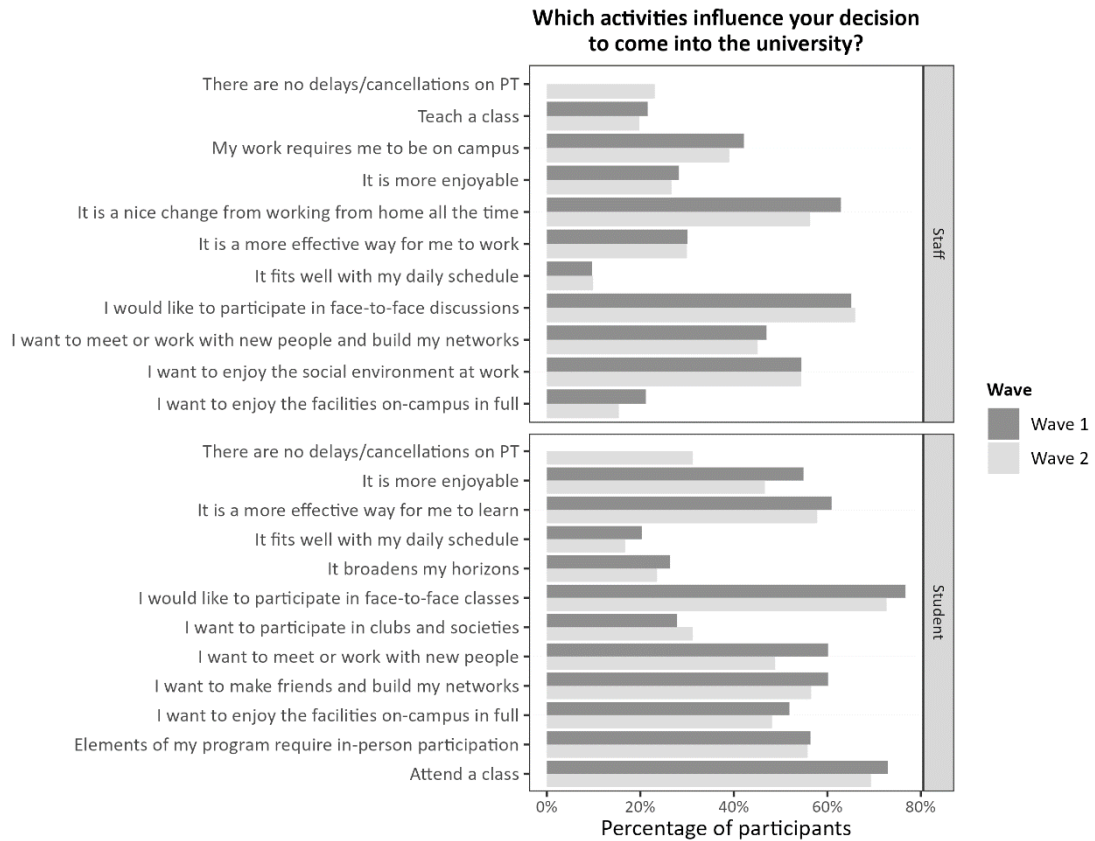


Figure 3: Which activities influence your decision to come into the university?

3. METHODOLOGY

A mixed logit model was estimated to identify the main drivers for students and staff members to decide where to work/study each day of the week. The modelling framework is presented in Figure 4. The decisions were represented by twelve alternatives, which are given in Table 2.

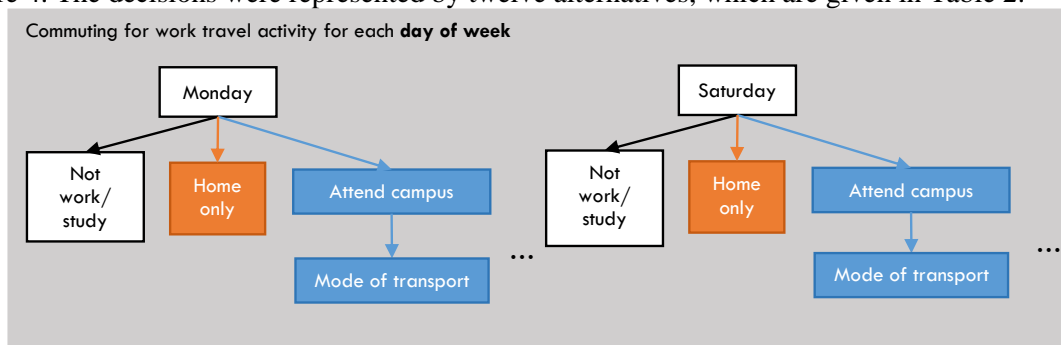


Figure 4: Modelling framework

Table 2: Alternative numbers per day of week

Monday - Sunday	
Alternative	Description
1	Not work/study
2	Work/study from home only
3	Attending campus - car driver
4	Attending campus - car passenger
5	Attending campus - taxi/rideshare
6	Attending campus - train
7	Attending campus - bus
8	Attending campus - light rail
9	Attending campus - ferry
10	Attending campus - walk
11	Attending campus - bicycle
12	Attending campus - motorcycle

The utility function of the WFH/SFH alternative, is expressed as follows:

$$U_{WFH/SFH} = \beta_{WFH/SFH} + \sum_j \beta_j \cdot Z_{qj} + \sum_i \beta_i \cdot H_{qi} + \varepsilon_{WFH/SFH} \quad (1)$$

where Z_{qj} represents characteristic j of respondent q (e.g., age, gender, income); H_{ni} represents attribute i of the home or work of respondent q (e.g., distance to campus, activities that influence decision to attend campus); ε represents the error term; and β are the estimated parameters associated with each attribute. The survey included different activities that could influence a participant's decision to attend campus. They were grouped into different categories, as shown in Table 3. The utility function for the alternatives to attend campus commuting by mode m is given by:

$$U_{Campus_m} = \beta_{Campus_m} + \sum_j \beta_j \cdot Z_{qj} + \sum_i \beta_i \cdot H_{qi} + \sum_k \beta_k \cdot X_{mk} + \varepsilon_{Campus_m} \quad (2)$$

where X_{mk} represent attribute k that describes mode m (e.g., travel time, fare/cost). The utility function of the *no work* alternative is expressed in equation (3):

$$U_{NoWork} = \beta_{NoWork} + \sum_j \beta_j \cdot Z_{qj} + \varepsilon_{NoWork} \quad (3)$$

It is important to note that respondents provided responses on the choice made each day of the 7-day week, and hence there are 7 choice sets per respondent. To recognise this, the error terms account for the panel structure of the data, i.e., varying across individuals but the same within individuals. The model results include those parameter estimates that were statistically significant with a 95% confidence level.

Table 3: Categories for activities that influence participants decision to attend campus

Description	Staff/Student	Category
I would like to participate in face-to-face classes	Student	Attend university because of classes/work
Attend a class	Student	Attend university because of classes/work
Elements of my program require in-person participation	Student	Attend university because of classes/work
It is a more effective way for me to learn	Student	Attend university because of classes/work
It is more enjoyable	Student	Attend university for social activities
I want to enjoy the facilities on-campus in full	Student	Attend university because of facilities or fits within my daily schedule
I want to meet or work with new people	Student	Attend university for social activities
I want to make friends and build my networks	Student	Attend university for social activities
It broadens my horizons	Student	Attend university for social activities
I want to participate in clubs and societies	Student	Attend university for social activities
It fits well with my daily schedule (e.g., childcare in or close to campus; my gym is close to campus)	Student	Attend university because of facilities or fits within my daily schedule
There are no delays/cancellations on public transport	Student	Attend university when there are no delays/cancellations on PT
I would like to participate in face-to-face discussions	Staff	Attend university because of classes/work
Teach a class	Staff	Attend university because of classes/work
My work requires me to be on campus	Staff	Attend university because of classes/work
It is a more effective way for me to work	Staff	Attend university because of classes/work
It is more enjoyable	Staff	Attend university for social activities
I want to enjoy the facilities on-campus in full	Staff	Attend university because of facilities or fits within my daily schedule
I want to meet or work with new people and build my networks	Staff	Attend university for social activities
I want to enjoy the social environment at work	Staff	Attend university for social activities
It is a nice change from working from home all the time	Staff	Attend university because it is a nice change from home all the time
It fits well with my daily schedule (e.g., childcare in or close to campus; my gym is close to campus)	Staff	Attend university because of facilities or fits within my daily schedule
There are no delays/cancellations on public transport	Staff	Attend university when there are no delays/cancellations on PT

4. RESULTS AND DISCUSSION

Preliminary model results are presented in Table 4, which combine Waves 1 and 2. All the parameter estimates are statistically significant at the 95% confidence level. Travel times and costs (including fuel costs) have a negative parameter estimate, as expected, and the value of travel time savings will be discussed below. Interestingly, the distance from home to campus has a very significant influence on the probability to use the train, suggesting that staff and students who live further away from campus are likely to use the train to go to campus.

Table 4: MML Model Results

Description	Alternative	Mean (t-value)
Alternative specific constant no work (base)	No work	-
Alternative specific constant WFH	WFH/SFH	0.27 (6.32)
Alternative specific constant commute by car driver	Car driver	0.61 (4.28)
Alternative specific constant commute by car pax	Car pax	-1.15 (5.08)
Alternative specific constant commute by taxi/rideshare	Taxi/Rideshare	-1.06 (1.70)
Alternative specific constant commute by train	Train	-0.50 (3.83)
Alternative specific constant commute by bus	Bus	-0.14 (1.15)
Alternative specific constant commute by light rail	Light rail	0.09 (0.38)
Alternative specific constant commute by ferry	Ferry	0.38 (1.24)
Alternative specific constant commute walking	Walking	-0.86 (3.70)
Alternative specific constant commute by bicycle	Bicycle	-1.52 (6.23)
Alternative specific constant commute by motorcycle	Motorcycle	0.66 (3.03)
Travel time (mins)	Car driver, pax and motorcycle	-0.02 (2.73)
Travel time (mins)	Public transport modes	-0.01 (3.47)
Travel time active modes (mins)	Active modes	-0.001 (1.95)
Fuel cost (AUD\$)	Car driver	-0.06 (2.05)
Fare (AUD\$)	Taxi/Rideshare	-0.05 (2.11)
Fare (AUD\$)	Public transport modes	-0.19 (5.88)
Distance from home to work (kms)	Train	0.05 (10.98)
Female (1,0)	WFH/SFH	-0.19 (3.45)
Personal income staff ('000\$AUD)	WFH/SFH	0.00 (3.43)
Personal income students ('000\$AUD)	WFH/SFH	0.00 (3.96)
Monday (1,0)	WFH/SFH	0.58 (10.61)
Wednesday (1,0)	WFH/SFH	0.13 (2.32)
Thursday (1,0)	WFH/SFH	0.23 (4.02)
Friday (1,0)	WFH/SFH	0.75 (13.73)
Attend university for social activities - students (1,0)	Car driver, pax and motorcycle	-0.38 (2.83)
Attend university because of facilities or fits within my daily schedule - students (1,0)	Car driver, pax and motorcycle	0.50 (3.84)
Attend university when there are no delays/cancellations on PT - students (1,0)	Car driver, pax and motorcycle	-0.77 (4.44)
Attend university because of work - staff (1,0)	Car driver, pax and motorcycle	0.63 (5.23)
Attend university for social activities - staff (1,0)	Car driver, pax and motorcycle	-0.33 (3.37)
Attend university when there are no delays/cancellations on PT - staff (1,0)	Car driver, pax and motorcycle	-1.11 (4.64)
Attend university because it is a nice change from home all the time - staff (1,0)	Car driver, pax and motorcycle	-0.38 (3.94)
Attend university because of classes - students (1,0)	Active modes	1.41 (5.95)
Attend university for social activities - students (1,0)	Active modes	0.27 (2.50)
Attend university because of facilities or fits within my daily schedule - students (1,0)	Active modes	0.32 (3.48)
Attend university because of work - staff (1,0)	Active modes	0.53 (2.28)
Attend university for social activities - staff (1,0)	Active modes	0.78 (4.91)
Attend university because of facilities or fits within my daily schedule - staff (1,0)	Active modes	0.37 (2.69)
Standard error	Public transport modes	1.37 (27.82)
Standard error	No work	0.62 (16.35)
Number of parameters		40
Log-likelihood equal shares L(0)		-20,097.11
Log-likelihood at convergence		-18,028.33
AIC/n		2.778

Results show that students that attend campus mainly for social activities are more likely to use active modes and less likely to use motorised private modes, relative to the other alternatives. Students that attend campus mainly because of classes, are more likely to use active modes; and those that go for the facilities/fits within their daily schedules are more likely to use private motorised vehicles, followed by active modes. Staff members that attend campus mainly due to work are more likely to use car, followed by active modes; and those that attend due to social activities are more likely to use active modes and less likely to use the car. When there are no cancellations/delays with public transport, both staff and students are less likely to use private motorised modes.

The value of travel time savings (VTTS in \$/person hour) estimates for car, public transport and taxi/rideshare are presented in Table 5. The value for car driver aligns with what is obtained from the broader population of commuters but is lower than expected for public transport (which is expected to be around \$5/person hour). This might be due to the fact that we are combining students and staff members and incorporating activities to attend campus which might have a significant influence on the VTTS, which is being analysed in more detail at the moment. The results presented in this short paper are preliminary, and we are currently working on understanding all the explanatory variables that might be explaining daily decisions to WFH/SFH or attend campus, providing separation of parameters between staff and students.

Table 5: Value of travel time savings results

Value of travel time savings (AUD\$/person hour)	Mean	Std Error
Car driver	16.74	22.13
Public transport	1.99	1.16
Taxi/Rideshare	20.18	15.66

With this model, we are also able to estimate elasticities, which represent the percentage change in the probability to choose an alternative given percentage changes in the explanatory variables, *ceteris paribus*. The elasticity estimates for this preliminary model are presented in Figure 5. The results show that the distance from home to campus has the greatest influence on the probability to choose train: if a student or staff member lives 10% closer to campus, they are 9.4% less likely to choose the train. A student that goes to campus because of classes is 80% more likely to use the bicycle and 72% more likely to walk than a student that is not motivated to go to campus because of class. As mentioned above, the model presented in this report is an initial model and more needs to be done to obtain a final model.

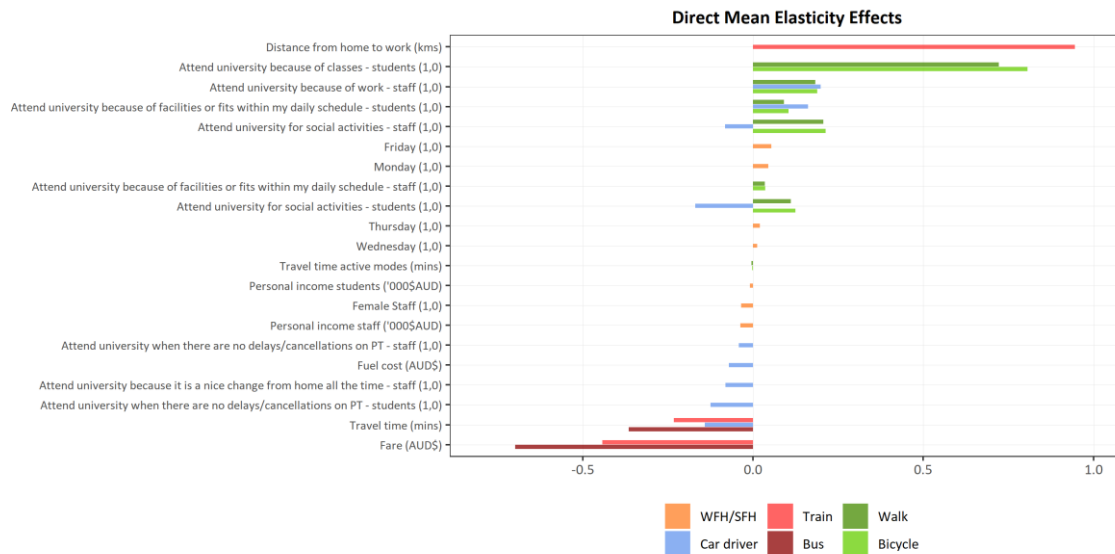


Figure 5: Elasticity results

5. CONCLUSIONS

This study aims to understand staff and students' preferences to work/study from home or to attend university campus using different modes of transport. The data is collected in two waves of data for staff and students that work/study at the University of Sydney, Australia. A mixed

logit is estimated to understand preferences for each day of the week. Results are preliminary, and initial findings suggest the importance of the distance from home to campus, fare, activities such as classes or work-related or social. Results suggest that those that live further away are more likely to use the train, while students that attend campus for social activities are more likely to use active modes and less likely to use car driver as their main mode of transport. These initial findings are encouraging as they are suggesting important drivers that should be considered when creating University travel demand management programmes to incentivise return to campus by sustainable modes of transport.

6. ACKNOWLEDGEMENTS

This research is part of iMOVE Cooperative Research Centre (CRC) Project 3-022 with Transport for New South Wales (TfNSW) on Promoting Sustainable University Travel Choices project (PSUTS). The findings reported are those of the authors and are not the position of TfNSW; but approval to present these findings is appreciated. The authors gratefully acknowledge financial support from ANID PIA/PUENTE AFB220003.

7. REFERENCES

- Balbontin, C., Hensher, D. A., Beck, M. J., Giesen, R., Basnak, P., Vallejo-Borda, J. A., & Venter, C. (2021). Impact of COVID-19 on the number of days working from home and commuting travel: A cross-cultural comparison between Australia, South America and South Africa. *Journal of Transport Geography*, 103188. <https://doi.org/10.1016/j.jtrangeo.2021.103188>
- Beck, M. J., & Hensher, D. A. (2021). Australia 6 months after COVID-19 restrictions part 2: The impact of working from home. *Transport Policy*, June. <https://doi.org/10.1016/j.tranpol.2021.06.005>
- Beck, M. J., Nelson, J. D., & Hensher, D. A. (2022). Attitudes toward public transport post Delta COVID-19 lockdowns: Identifying user segments and policies to restore confidence. *International Journal of Sustainable Transportation*, 1–18. <https://doi.org/10.1080/15568318.2022.2109083>
- Hensher, D. A., Beck, M. J., Nelson, J. D., & Balbontin, C. (2022). Reducing Congestion and Crowding with Working from Home. In C. Mulley & M. Attard (Eds.), *Transport and Pandemic Experiences* (pp. 235–255). Emerald Press. <https://doi.org/10.1108/S2044-994120220000017013>

Assessing expected ride-pooling performance with non-deterministic, heterogeneous travellers' behaviour.

Michal Bujak*^{1, 2} and Rafal Kucharski¹

¹Department of Mathematics and Computer Science, Jagiellonian University, ul. Prof. Lojasiewicza 6, Krakow, 30-348, Poland

²Doctoral School of Exact and Natural Sciences, Jagiellonian University, Golebia 24, Krakow, 31-007, Poland

SHORT SUMMARY

Ride-pooling remains a promising emerging mode with a potential to contribute towards urban sustainability and emission reductions. However, recent studies revealed complexity and diversity among travellers' ride-pooling aptitudes. So far, ride-pooling analyses assumed homogeneity and/or determinism of ride-pooling travellers. This, as we demonstrate, leads to a false assessment of ride-pooling system performance. We experiment with an actual NYC demand from 2016 and classify travellers into four groups of various ride-pooling behaviour (value of time and penalty for sharing), as reported in the recent SP study. We replicate their random behaviour to obtain meaningful distributions. Heterogeneity assumption proves to have a significant impact on the system. The performance indicators are shifted compared to the deterministic scenario. Albeit the high variability of travellers' preferences, system-wide results remain within reasonably narrow confidence intervals.

Keywords: ride-pooling, behavioural heterogeneity, shareability graph.

1 INTRODUCTION

Ride-pooling is a ride-hailing service which enables travellers to share a ride. Two or more trip requests are submitted to the platform and pooled into a single vehicle. Shared ride is typically longer (due to pick-up delay and detour) and yields a sharing discomfort, which has to be compensated by a lower ride fare.

Most ride-pooling algorithms (Alonso-Mora et al. (2017), Ke et al. (2021), Shah et al. (2020), Bilali et al. (2020), Wang et al. (2021)) operate on based on fixed time windows constraints. In reality, travellers are rational utility maximisers, who opt for the most attractive alternative. In this spirit, we proposed our previous ExMAS algorithm (Kucharski & Cats (2020)), which puts the traveller in the centre. Recent studies (Lavieri & Bhat (2019), Chavis & Gayah (2017), Alonso-González et al. (2020)) show that preferences towards pooling vary across population. Here, we introduce the behavioural heterogeneity (varying value of time, perceived discomfort of sharing) into the algorithm. While we are not aware of the individual traveller's preferences, we assume to know the distribution in the population. Applying results from Alonso-González et al. (2020) study on the actual taxi requests from NYC, we replicate the ride-pooling experiment 1000 times to obtain a meaningful estimation of the results.

We find that introduction of the heterogeneous preferences shifts performance indicators compared to the homogeneous scenario. While expected mileage reduction and profitability are worse than in the deterministic benchmark, traveller satisfaction with the service improves.

Our method allows the provider for more meaningful estimation of the system performance and its distribution. Results, apart from the expected values, reveal confidence intervals and tails. Estimates of performance (e.g. vehicle kilometres travelled, fares collected) and perceived attractiveness for travellers are now more meaningful and robust.

2 METHODOLOGY

To reach the objectives of this study, we extend the previously proposed off-line utility-based ride-pooling algorithm ExMAS. We introduce non-deterministic utility formulas to integrate them

with the previous deterministic approach and solve the ride-pooling problem for the heterogeneous population.

The original, traveller-oriented ExMAS is a utility-based algorithm. Traveller is assumed to share a ride only if it is more attractive than the alternatives (in our case, private ride). We measure attractiveness via utility, which has always a negative sign as it represents the perceived cost of a trip. The utility of a shared ride (for passenger i participating in ride r_k denoted as U_{i,r_k}^s) and of the non-shared ride (denoted U_i^{ns}) are as follows:

$$\begin{aligned} U_i^{ns} &= -\rho l_i - \beta_t t_i \\ U_{i,r_k}^s &= -(1-\lambda)\rho l_i - \beta_t \beta_s (\hat{t}_i + \beta_d (\hat{t}_i^p)) + \epsilon, \end{aligned} \quad (1)$$

where ρ stands for per-kilometre fare, while λ denotes discount for sharing a ride. Both are controlled by the operator. β^t , β^s , β^d are the behavioural parameters: value of time, penalty for sharing and delay sensitivity, respectively. t_i and \hat{t}_i stand for travel time of non-shared and shared ride, respectively, \hat{t}_i^p is updated for each evaluated shared-ride candidate and is typically greater than t_i due to both pooling detour (to pick-up and drop of other travellers) and delay (to wait for others). ϵ is a random term. By conducting the hierarchical search, while preserving exhaustiveness, the algorithm prevents the search space explosion and yields an optimal solution. Details are described in Kucharski & Cats (2020).

In this study, we extend the above approach to account for the heterogeneous behavioural characteristics of individuals. Namely, any of behavioural parameters (β 's in eq. 1) instead of being constants, can be now a random variable. We assume that preferences of an individual are unknown, while we know the distributions within the population. Moreover, to account for additional uncertainty, we introduce a panel noise ϵ_i (traveller-specific) and idiosyncratic noise $\epsilon_{i,r}$ (ride-specific for specific traveller), leading to the following formulas:

$$U_i^{ns} = \rho l_i + \hat{\beta}_t t_i \quad (2)$$

$$U_{i,r_k}^s = (1-\lambda)\rho l_i + \hat{\beta}_t \hat{\beta}_s (\hat{t}_i + \hat{\beta}_d (\hat{t}_i^p)) + \epsilon_i + \epsilon_{i,r}, \quad (3)$$

where all notions introduced in eq. 1 still apply, yet $\hat{\beta}$'s are now random variables. Consistently with a discrete choice theory, we assume that the probability that a traveller i finds a shared ride r_k attractive is expressed with:

$$P_{i,r_k}^s = \Pr(U_{i,r_k}^s > U_i^{ns}), \quad (4)$$

which, depending on the random variables distributions, may become e.g. a logit or a probit model. While in general any distribution of random variable can be applied in the method, in the experiment we use multiple classes \mathcal{C} and assume the random variables to follow a multimodal normal distribution (unimodal within the class):

$$\hat{\beta}_j = \sum_{\mathcal{C} \in \mathcal{C}} p_{\mathcal{C}} X(\bar{\beta}_{j,\mathcal{C}}, \sigma_{j,\mathcal{C}}) \text{ for } j \in \{t, s, d\}, \quad (5)$$

where $p_{\mathcal{C}}$ is the probability of belonging to the class \mathcal{C} , $X(\bar{\beta}_{j,\mathcal{C}}, \sigma_{j,\mathcal{C}})$ is a random variable following normal distribution with mean of $\bar{\beta}_{j,\mathcal{C}}$ and a standard deviation of $\sigma_{j,\mathcal{C}}$ (coefficient- and class-specific).

We measure the performance of ride-pooling solution with the four following indicators. For the environment and city perspective we look at the vehicle-kilometres saved due to pooling \mathcal{D} . For travellers, we observe on one hand the pooling costs (relative detours \mathcal{T}) and on the other benefits (relative improvement in utility \mathcal{U}). For the platform, we look at the profitability of ride-pooling service \mathcal{P} . All values are calculated as relative to the value of the private ride scenario. Furthermore, indicators are calculated not only system-wide, but on the levels of a single traveller (utility, travel time extension), ride (mileage reduction, profitability).

3 RESULTS AND DISCUSSION

Experiment settings

We illustrate how the proposed method enhances assessment of ride-pooling services with the case-study of New York City. We reproduce the likely case when the service provider can predict the demand and its behavioural structure, yet the actual traits of individual travellers remain



Figure 1: Demand dataset for experiments: ca. 150 trip requests from Jan 2016 in Manhattan. Green dots are origins, orange destinations.

latent. To obtain reliable estimates, we fix the demand (requests) and replicate the experiment (assuming various ride-pooling behaviour of individual travellers) 1000 times. Results are available to reproduce on the public repository¹

We experiment with the trips actually requested in the New York City on January 2016. We use the set 147 trips requested in the 30 minute batch (which is beyond the critical mass needed to induce pooling). Each trip request links origin with destination at a given time (as illustrated in Figure 1). We used the fare λ_p of 1.5 €/km (converted from \$, consistent with NYC Taxi & Limousine Commission (2022)), sharing discount λ is 30% (within the range suggested by Shaheen & Cohen (2019)). We assume $\beta_d = 1$ while $\hat{\beta}_t$ and $\hat{\beta}_s$ are derived from study by Alonso-González et al. (2020).

Value of time (VoT) found by Alonso-González et al. (2020) was explicitly applicable to our study. However, penalty for sharing (Pfs) needed to be adjusted. ExMAS assumes a penalty for sharing irrespective of the number of co-travellers. Moreover, the penalty is scaled with trip length, not fixed. Hence, we took value for 3 additional passengers, which was found for one class. We scaled the results to obtain values for the remaining three classes (based on proportions in the fixed penalty for sharing with one additional passenger). According to formulas by Seltman (2012), we retrieved the variance of the variables (which was missing in the original study).

Eventually, we successfully obtained: class membership probabilities (p_C), mean values of value-of-time and penalty for sharing for four different classes ($\beta_{s,C}$, $\beta_{t,C}$) and class-respective variances ($\sigma_{s,C}$, $\sigma_{t,C}$) as reported presented in Table 1.

Table 1: Ride-pooling behavioural parameterization from Alonso-González et al. (2020). Mean values and standard deviations of value of time (VoT) and penalty for sharing (Pfs) for four classes. *for class 4 we arbitrarily clipped the otherwise too wide st. dev. of VoT

parameter		class	C1	C2	C3	C4
		name	"It's my ride"	"Sharing is saving"	"Time is gold"	"Cheap and half empty"
VoT	$\beta_{t,C}$	mean	16.98	14.02	26.25	7.78
		st.dev.	0.318	0.201	5.777	1*
Pfs	$\beta_{s,C}$	mean	1.22	1.135	1.049	1.18
		st.dev.	0.082	0.071	0.06	0.076
Share	p_C		29%	28%	24%	19%

Expected ride-pooling performance with non-deterministic travelers

We dive into details how the behavioural heteroscedasticity impacts the ride-pooling performance. We assess the ride-pooling system's performance with the four selected indicators. Calculated first for consecutive replications (realisations of random variables) and then accumulated over all 1000 replications (one value corresponds to one realisation). Results are presented in Figure 2. Baseline for comparison is the deterministic ExMAS (with VoT and Pfs being weighted average of means introduced in Table 1) which yields mileage reduced by 30%, 9.8% detour, utility increased by 4.5% and profitability of 1.097.

¹Script to reproduce results is on the branch *probabilistic_topological* of original ExMAS (direct link: https://github.com/RafalKucharskiPK/ExMAS/blob/probabilistic_topological/Utils/Probabilistic_ExMAS_wrapper.py).

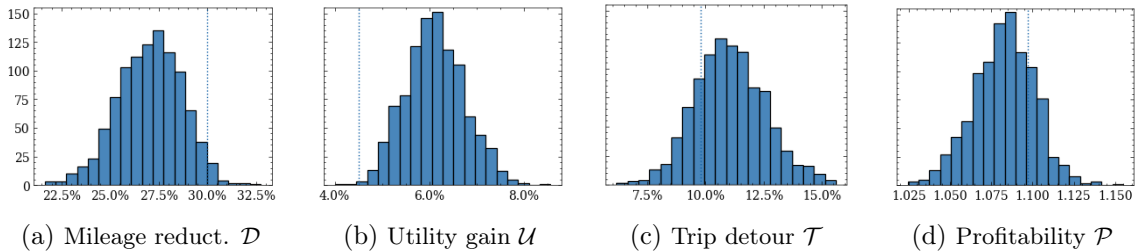


Figure 2: Distribution of ride-pooling performance resulting from 1000 replications. Dotted lines show the deterministic benchmark.

Regardless of behavioural variability, pooling always reduces the travel distance (\mathcal{D}). 90% of the observations range between 24.3% and 29.7% with the average of 27.1%, which is significantly worse than 30% reductions obtained in the deterministic setting (Figure 2a). While this is a significant reduction, which should be hugely appreciated from the city’s and environmental’s perspective, it also varies significantly, with observations ranging from 20% to 32.5%. The right tail seems to be fatter, with occasional outliers greater than 30% vehicle kilometres reductions.

The mean of average utility gain (which is user-subjective and depends on behavioural profile) is approximately 6.1%, which, in turn, is now significantly greater than in the deterministic benchmark. The 90% two-sided confidence interval spans between 5.1% and 7.2% utility gains.

A more physical measure of pooling performance, from the user’s perspective, is the trip detour, which does not depend on traveller behaviour. Notably, the acceptable detour (and delay), unlike the fixed time windows used in other studies, depends on VoT and PfS and can vary significantly with the behaviour. The mean travel time of pooled ride is approximately 11% longer than for the solo-rides only system (Figure 2c) and 90% of the observations fit between 8.6% and 13.6%, which, again, is worse than in the deterministic benchmark (9.8%).

Class dependent ride-pooling performance

Considering the importance of the behavioural preferences of individuals on pooling performance, we analyse further the data with respect to individuals and their classes. First, we show the cumulative distributions of individual indicators observed in 1000 replications. Along with overall distributions, we present the profiles associated with four distinct classes. While the CDF profiles serve for illustrative purposes mainly, we enhance them with tables showing differences between classes in the four consecutive indicators as they reveal intriguing patterns among the introduced behavioural classes.

Here, we aggregate all the rides, irrespective of in which replication they appear. In the Fig. 3, we present the cumulative distribution functions of the utility gains and detours. For classes C1 and C3, the utility gains are below average and detours are above average. For the class C4 the opposite is true, while class C2 is on average with the mean performance.

In the Table 2 we report the resulting trip detour for the passengers in the four respective behavioural classes. Apart from mean and standard deviation, we report values of 75th, 90th percentile and 95th percentile. First, we present the values obtained for all the rides and then, for illustrative purposes, only for shared rides (excluding the travellers who were not matched with anyone and travelled alone). While on average, pooled rides are 8% longer than private rides, it varies from 5% for class C1 to 15% for class C4. Notably, 5% of travellers in C4 decided to pool while having at least 61% longer travel times. Those trends are further pronounced when we restrict the analysis to the rides which were successfully pooled (right columns).

Similarly to travel time analysis in Table 2, we present data of relative utility gain in ride-hailing system with ride-pooling (relative to system without pooling). Surprisingly, the class which has the longest detours (C4) has also the greatest utility gains (11% on average). For the class unwilling to pool the benefits are only 4% (C1). For 10% of all the travellers ride-pooling led to at least 25% increase in perceived utility.

Ride-pooling performance vs ride-pooling behaviour

In this section, we provide an insight into the relation between sampled behavioural parameters (VoT, PfS) and performance of the individual travellers (\mathcal{U}_r , \mathcal{T}_r). In Figure 4, we scatter the values of VoT and PfS (first and second row respectively) on x -axis against the performance indicators,

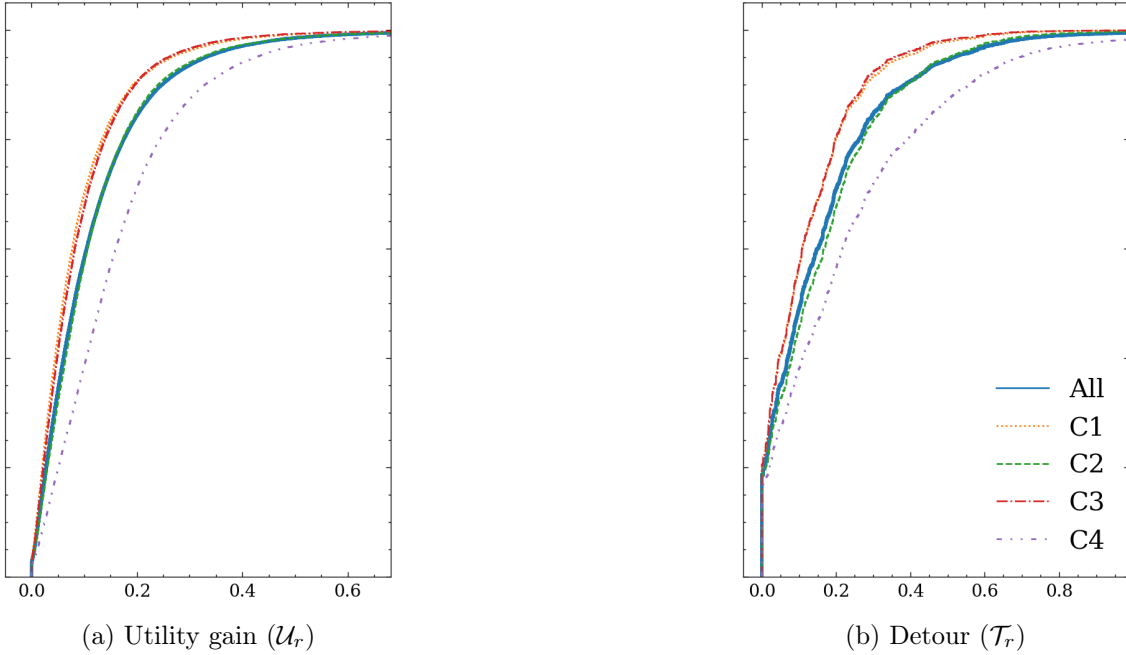


Figure 3: Cumulative distribution functions of utility gains and detours for the trips observed in 1000 replications, stratified into four classes of behaviour.

Table 2: Statistical properties of trip detour (relative) \mathcal{T}_r for travellers of four respective behavioural classes. Mean, variance and significant percentiles are shown first for all the travellers (left) and then (right) only for those who were pooled.

	All rides					Shared rides				
	Means	St.dev.	75	90	95	Means	St.dev.	75	90	95
All	0.08	0.16	0.11	0.27	0.39	0.16	0.19	0.22	0.39	0.52
C1 "It's my ride"	0.05	0.11	0.07	0.19	0.27	0.12	0.13	0.18	0.28	0.38
C2 "Sharing is saving"	0.09	0.15	0.14	0.28	0.41	0.16	0.17	0.23	0.39	0.50
C3 "Time is gold"	0.05	0.11	0.07	0.19	0.27	0.11	0.13	0.18	0.27	0.35
C4 "Cheap and half empty"	0.15	0.23	0.23	0.45	0.61	0.23	0.25	0.33	0.56	0.69

Table 3: Statistical properties of travellers gains (in relative increase of utility) \mathcal{U} for travellers of four respective behavioural classes. Mean, variance and significant percentiles are shown first for all the travellers (left) and then (right) only for those who were pooling.

	All rides					Shared rides				
	Means	St.dev.	75	90	95	Means	St.dev.	75	90	95
All	0.06	0.10	0.09	0.18	0.25	0.11	0.12	0.15	0.25	0.33
C1 "It's my ride"	0.04	0.08	0.05	0.12	0.18	0.09	0.10	0.12	0.19	0.26
C2 "Sharing is saving"	0.06	0.10	0.09	0.18	0.25	0.11	0.12	0.15	0.24	0.32
C3 "Time is gold"	0.04	0.08	0.06	0.14	0.19	0.09	0.10	0.12	0.20	0.26
C4 "Cheap and half empty"	0.11	0.14	0.16	0.28	0.37	0.16	0.15	0.22	0.33	0.42

i.e. \mathcal{T}_r and \mathcal{U}_r (first and second column) on y -axis. Each dot represents an individual traveller, coloured accordingly to her behavioural class.

We can observe a clear trend of ride-pooling detours decreasing with an increasing value of time (Fig. 4a), despite the longer detours for travellers with low value of time, their benefits of pooling (relative increase in utility) remains high (Fig. 4b), whereas for travellers with high value of time utility gains are significantly lower. The three classes (C1, C2 and C4) have relatively low variance and do not overlap, while for class C3 the variance of VoT is very big. Mind that in our experiment each traveller was assigned to different classes across replications, nonetheless the class membership strongly correlate with the resulting ride-pooling performance, despite the fixed spatiotemporal trip characteristics.

Those clear trends become blurred when we plot against penalty-for-sharing (PfS in Fig. 4c and 4d) where the Gaussian shape is observed and class memberships are indistinguishable. Both highest

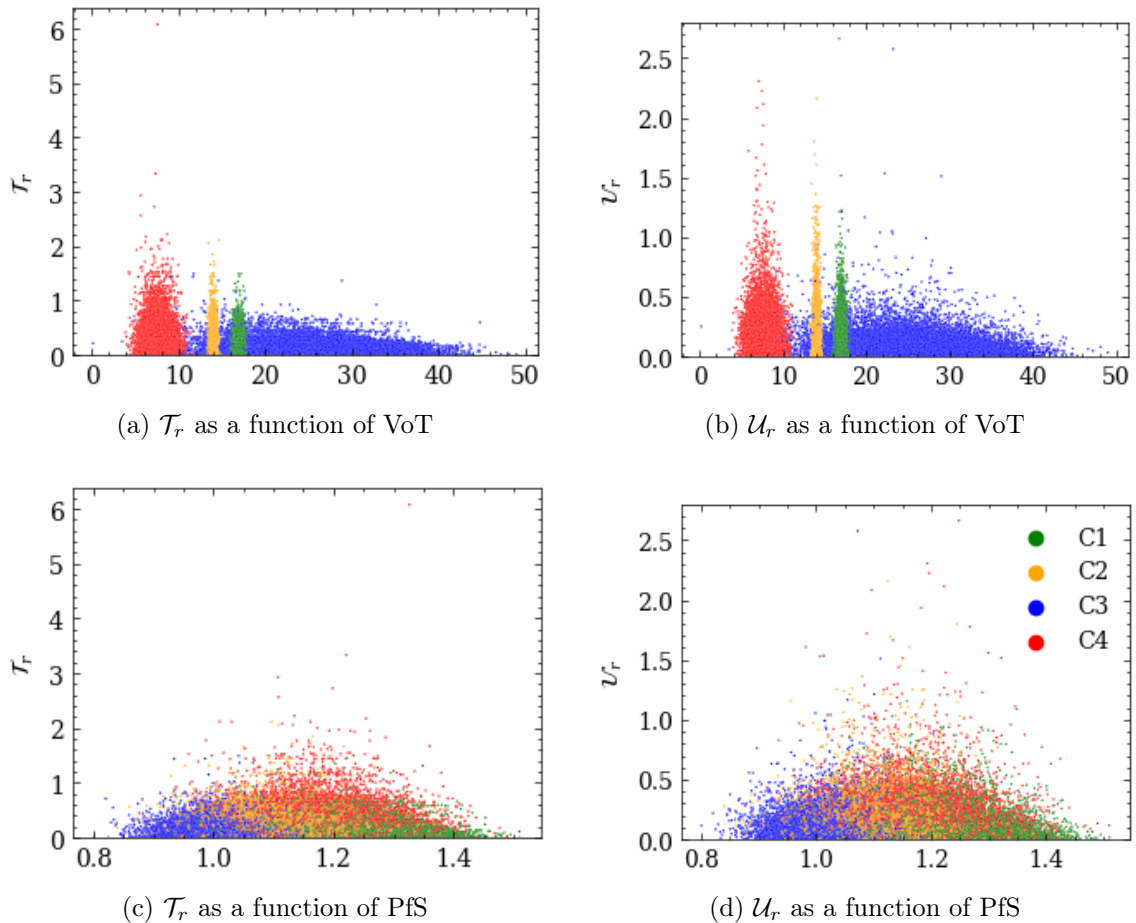


Figure 4: Scatter plot of travellers’ behavioural parameters (x -axis) against the resulting ride-pooling performance. Each dot denotes a single traveller, coloured accordingly to the class assigned to her in the respective replication. y -axis are scaled to represent relative change, e.g. 1 corresponds to increase of 100%.

detours and utility gains are obtained for intermediate values of PfS, while extreme cases have the lowest benefits. The travellers with a low penalty for sharing (even below 1 for some of class C3 travellers) seem to be often exploited by the system and their utilities are sacrificed to match preferences for travellers with lower preferences for sharing. Yet for those with high penalty the benefits are also limited, since they can hardly be pooled with others. The pooling seems to be best performing in the penalties ranging from 1.1 to 1.2 where it can reach the greatest expected benefits (Fig. 4d).

Scaling with the demand levels

The notion of critical mass is crucial to the ride-pooling and number of studies reported how the ride-pooling grows non-linearly with the demand size. Yet, up to now, those findings were reported in the deterministic setting only. Here, apart from reporting how system performance improves with growing demand, we provide insights on how the performance variability changes.

Demand set on which we run the experimental was selected to be slightly beyond the critical mass needed to induce the effective pooling, we extend it here from two sides: sub-critical (99 requests) and super-critical (198 requests). We report the previously introduced performance measures for three levels of demand in Figure 5. Unsurprisingly, we observe the critical mass effect as the performance significantly increases when the demand grows from 100 to 150 requests and then somehow stabilises when it reaches 200 requests (Fig. 3). Yet more importantly, we observe the trends with variabilities, which are alternative: either narrowing with growing demand (utility in Fig. 5b), widening (profitability in Fig. 5b), or remaining roughly constant (travel times and mileage savings in Fig. 5a and Fig. 5c).

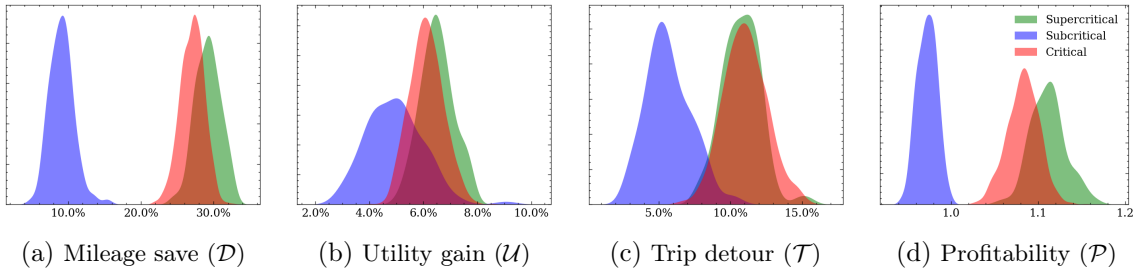


Figure 5: Distributions of four performance indicators with the three levels of demand: subcritical, critical and supercritical.

4 CONCLUSIONS

Most ride-pooling algorithms rely on the fixed constraints (usually time windows). To more objectively assess the attractiveness of a shared ride from the traveller’s perspective, we previously proposed the utility-based ExMAS algorithm. In this study, we take the matter further. We no longer consider travellers homogeneous and deterministic. Using recent empirical findings, we assume that while we still do not know preferences of the individuals, we know the general distributions.

To properly include this in our calculations, we introduced random variables representing the value of time and perceived sharing discomfort. Such reformulation allowed to reproduce the four distinct classes of travellers’ behaviour towards ride-pooling, as revealed in Alonso-González et al. (2020). Conducting sufficient number of replications, we obtain reliable estimations of the ride-pooling problem solution, which is no longer deterministic. Apart from mean values, we can now estimate the lower and upper bounds of the performance indicators, which substantially differ from the deterministic benchmark.

The additional probabilistic layer introduces a variability at the level of individuals and at the system performance. We find that despite the high variability of individuals, pooling performance remains stable and within reasonably narrow confidence intervals. Notably, probabilistic results are shifted compared to the deterministic benchmark. While the primary objective of minimising mileage is better met in the deterministic scenario, we observe a much higher satisfaction with the service in the heterogeneous setting.

Our study also provides an insight into how ride-pooling performs for travellers of certain preferences. We analyse the impact of the value of time and perceived sharing discomfort on trip detour and satisfaction with the service. We find that people with low value of time can be considered both the most flexible and the most beneficial travellers in the pooling system. However, those with intermediate penalty for sharing not only benefit more than those with high penalties (who does not want to share in general), yet also more than those with low penalties (willing to share with anyone and often exploited by the system).

Similarly to the deterministic case, we observe the critical mass effect and pooling becomes effective only when the demand levels reach the so-called critical mass. Now we enrich this notion with findings on variability, which may either decrease (in terms of utility gains) or increase (in terms of profitability for the provider) with growing demand levels.

The proposed method is general and can be easily applied to new cases, both for general demand patterns and different behavioural models. Also, the specific experimental setting used in this study may be reformulated, e.g. when we know individuals’ class membership, the demand is not predicted properly or when the behaviour is assumed fixed, but demand is varying (like in Kucharski et al. (2021)). In the future studies, those additional dimensions of variability may be included for even richer assessments. Finally, the proposed method may be valuable in the pandemic-analyses, when virus-averse behaviour drives the pooling behaviour of individuals.

ACKNOWLEDGEMENTS

This research was funded by National Science Centre in Poland program OPUS 19 (Grant Number 2020/37/B/HS4/01847)

REFERENCES

- Alonso-González, M., Cats, O., & van Oort, N. e. a. (2020). What are the determinants of the willingness to share rides in pooled on-demand services? *Transportation*, 1733–1765. doi: <https://doi.org/10.1007/s11116-020-10110-2>
- Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., & Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, 114(3), 462–467.
- Bilali, A., Engelhardt, R., Dandl, F., Fastenrath, U., & Bogenberger, K. (2020). Analytical and agent-based model to evaluate ride-pooling impact factors. *Transportation Research Record*, 2674(6), 1–12.
- Chavis, C., & Gayah, V. V. (2017). Development of a mode choice model for general purpose flexible-route transit systems. *Transportation Research Record*, 2650(1), 133–141.
- Ke, J., Zheng, Z., Yang, H., & Ye, J. (2021). Data-driven analysis on matching probability, routing distance and detour distance in ride-pooling services. *Transportation Research Part C: Emerging Technologies*, 124, 102922.
- Kucharski, R., & Cats, O. (2020). Exact matching of attractive shared rides (exmas) for system-wide strategic evaluations. *Transportation Research Part B: Methodological*, 139, 285–310.
- Kucharski, R., Cats, O., & Sienkiewicz, J. (2021). Modelling virus spreading in ride-pooling networks. *Scientific Reports*, 11, 1–11.
- Lavieri, P. S., & Bhat, C. R. (2019). Modeling individuals’ willingness to share trips with strangers in an autonomous vehicle future. *Transportation research part A: policy and practice*, 124, 242–261.
- NYC Taxi & Limousine Commission, u. (2022). *Taxi fare*. Retrieved from <https://www1.nyc.gov/site/tlc/passengers/taxi-fare.page>
- Seltman, H. (2012). *Approximations for mean and variance of a ratio*. unpublished note.
- Shah, S., Lowalekar, M., & Varakantham, P. (2020). Neural approximate dynamic programming for on-demand ride-pooling. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 34, pp. 507–515).
- Shaheen, S., & Cohen, A. (2019). Shared ride services in north america: definitions, impacts, and the future of pooling. *Transport reviews*, 39(4), 427–442.
- Wang, J., Wang, X., Yang, S., Yang, H., Zhang, X., & Gao, Z. (2021). Predicting the matching probability and the expected ride/shared distance for each dynamic ridepooling order: A mathematical modeling approach. *Transportation Research Part B: Methodological*, 154, 125–146.

Implementing an Agent-Based Formation of Social Networks for Joint Travel

Gabriel Hannon*¹, Joanna Ji², Qin Zhang³, Dr. Ana Tsui Moreno⁴, and Dr. Rolf Moeckel⁵

^{1,2,3,4,5}School of Engineering and Design, Technical University of Munich, Germany

SHORT SUMMARY

Introducing social networks to travel demand models could better capture socially induced travel behavior. This paper presents an agent-based approach to forming social networks that match important global characteristics and egocentric homophilies in distance, age, and gender for a population on the order of 10^6 . Based on data from an egocentric snowball sample, this methodology successfully reproduces homophilies in age and gender, as well as an expected power-law distribution of geographic distance between connections. An initial clique formation heuristic is implemented on top of the homophily calculations. The generated network exhibits preferential attachment between agents of higher degree, in line with more general literature on network formation.

Keywords: agent-based modeling, synthetic social networks, synthetic populations, transportation network modeling, travel demand modeling

1 INTRODUCTION

As Frei and Axhausen put it in their seminal 2007 paper on the geography of social networks, “travel is the price we pay to be with others” (Frei & Axhausen, 2007). Recently, researchers in transportation modeling have further investigated the marked interactions between social connections and travel activity participation (Carrasco et al., 2008; Kim et al., 2018). Social contacts affect multiple aspects of travel behavior, from activity generation to destination and mode choice (Kim et al., 2018). The literature has identified socio-demographic homophilies, geographic distance, and structural network properties, such as clique and degree distributions, as key characteristics of travel-focused social networks (Arentze et al., 2012; Illenberger et al., 2013; Dubernet, 2017). The synthesis of a static social network for a study area is a key next input for generating joint travel demand.

Social network structures are studied across fields. The development of a synthetic social network suitable for travel behavior modeling, however, has been challenging. Illenberger et al. (2013) approach this problem with an exponential random graph model, which accounts for homophily but has limited transitivity. Arentze et al. (2012) explicitly account for transitivity via a link probability model using binary logit estimations for forming friendships. Work by Dubernet (2017) generates a social network with a heuristic that accounts for socio-demographic and geographic homophily in addition to clique distribution, but does not account for agent-specific preferences. All three of the above utilize the snowball data sampled by Kowald & Axhausen (2012).

Building on this research, we attempt a scalable method to generate a synthetic social network that matches known socio-demographic homophily, connection distance, reciprocity, and transitivity properties. As the network is eventually intended to couple with a travel demand generation model, it is synthesized explicitly for the Munich metropolitan region.

2 METHODOLOGY

The fundamental methodology relies on a union of agent-based objects and network data structures. Agent objects are constructed from a synthetic population for the city of Munich generated by the open-source land-use model SILO (Moreno & Moeckel, 2018). Network and agent attributes - namely degree distribution, edge length preferences, and egocentric homophilies - are derived from ETH-Zurich’s snowball dataset (Kowald & Axhausen, 2012). This section describes these input

data and the social network formation algorithm.

Input data

Our study area is the Munich metropolitan region, including the City of Munich and the surrounding cities of Augsburg, Ingolstadt, Landshut, and Rosenheim, as seen in Figure 1. This area has a population of 4.5 million in roughly 2.1 million households. We use the output of the SILO land use model, which generates a synthetic population for the region based on census data and Iterative Proportional Updating (Moreno & Moeckel, 2018). Though a level of social connection can be inferred from shared households, workplaces and education places, there are no additional social or friendship connections. Therefore, we set out to generate a friendship network for a 5% sample of this synthetic population.

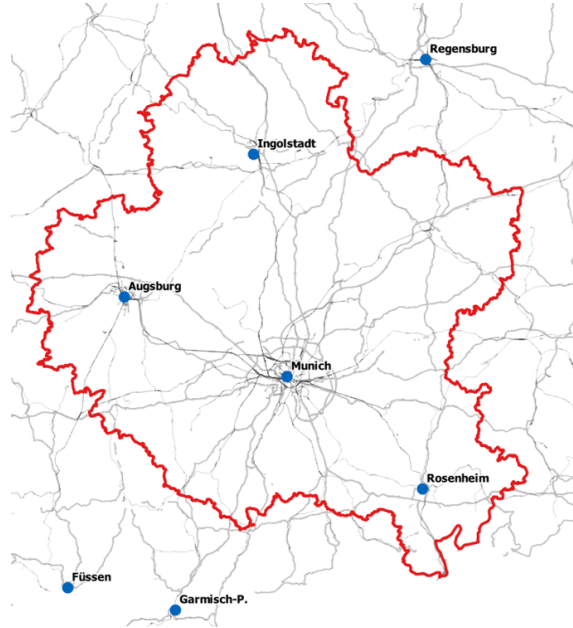


Figure 1: Munich metropolitan region.

Snowball Data

Despite the proliferation of data regarding large social networks, only a few datasets include the type of geographically-embedded demographic connection data that may impact joint leisure travel. Kowald & Axhausen’s (2012) snowball survey provides a key source for this topic. After data cleaning, this survey yields a total of 793 egos and 14,326 unique ‘names.’

While this data is extensive, it maps a sparse network, which tends to branch out to isolated alters. Therefore, any analysis of this data must focus on egocentric metrics, not global network characteristics. Previous work has established that age, gender, and distance homophilies are the most significant demographic attributes in leisure travel networks (Dubernet, 2017; Arentze et al., 2012; Illenberger et al., 2013). These characteristics, along with ego degree, form the main egocentric attributes for our formation model.

While previous work has used population-aggregated homophilies (Kowald & Axhausen, 2012; Dubernet, 2017), further examination suggests that demographic attributes affect an agent’s willingness to accept variation across these homophilies. After dividing the data into segments based on eight 10-year age brackets and two genders, Figure 2 illustrates the distribution in accepted average gender homophily, age difference, and degree by segment. Such segmenting shows, for example, that egos aged over 70 (70M, 70F, 80M or 80F) tend to accept ties across greater ranges of age differences and have smaller degrees. These insights motivate a formation model that accounts for the interaction between an agent’s demographic characteristics and homophily profile.

This segmenting leads to distributions for agent degrees by segment and a table of distributions for age and gender homophilies by segment pair. Figure 3 compiles the age and gender homophily distribution table. Rows represent network-level distributions of connections from egos in one

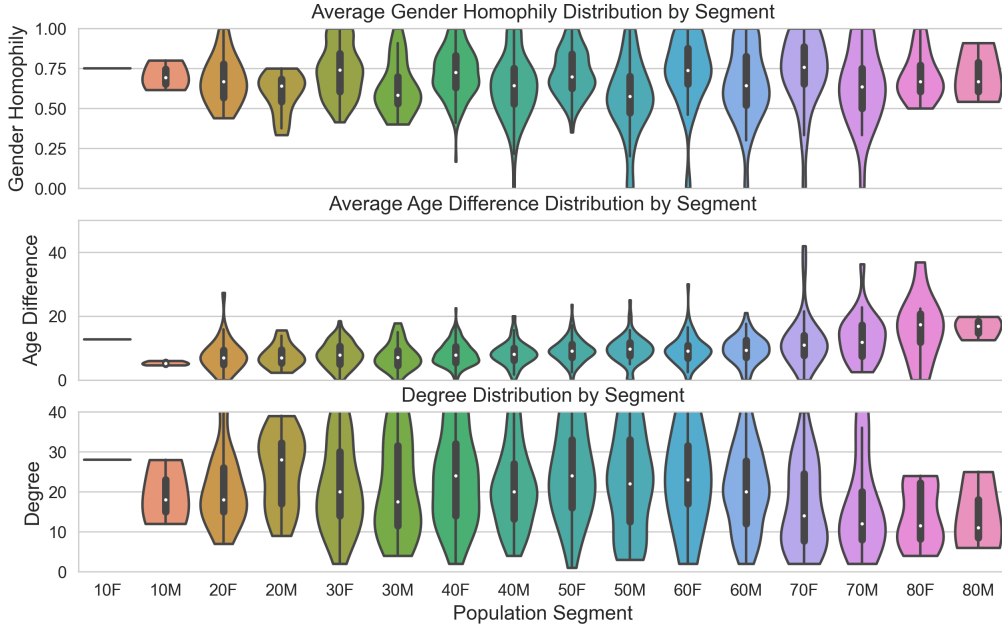


Figure 2: Segmented distribution in age difference, gender homophily, and degree for snowball egos by segment. A gender homophily of 1 indicates similar genders.

segment to alters in all other segments. This matrix is notably asymmetric; we can attribute this to sampling bias in the snowball data - e.g., 28% of egos are in their 50s while only 0.6% are younger than 20 - and fundamental asymmetries in the way that connections and popularity are distributed among agents in all social networks (Barabási & Albert, 1999).

10F	28	5.6	20	5	12	4.4	3.7	3.7	2.8	3.1	1.9	2.5	1.9	1.9	1.9	1.9	1.9
10M	13	31	4.2	13	3.5	7.7	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
20F	3.3	1.3	43	21	7.7	6.4	3.2	2.7	6.2	3.2	0.8	0.2	0.3	0.3	0.3	0.3	0.3
20M	3.5	2	24	39	7	12	0.6	1.7	2.6	3.8	1.2	0.6	0.6	0.3	0.6	0.9	0.9
30F	1	0.5	5.6	2.2	33	12	19	9.2	4.9	2.2	4.4	2.8	1.6	1.1	0.5	0.2	0.2
30M	0.2	1.4	5.6	7.3	15	27	8.4	17	2.3	5.2	2.8	4	0.9	1.6	0.9	0.5	0.5
40F	1.9	1.9	1.5	0.9	9.1	2.4	40	15	10	5.4	3.7	1.8	3.2	2.1	0.8	0.8	0.8
40M	0.5	1.4	1.6	2.3	7.8	7.2	18	35	4.2	9.5	2.5	4	2.5	2.9	0.5	0.7	0.7
50F	0.7	0.3	3.8	2.6	4.4	2.2	14	4	30	13	12	6.3	3.2	1.5	1.9	0.7	0.7
50M	0.4	1.5	3.8	4	3.4	3.8	11	12	14	25	4.3	11	1.4	2	1	1.1	1.1
60F	0.2	0.2	0.6	0.7	6	2.8	6.7	2.6	14	4.6	33	11	8.8	4.7	3	1	1
60M	0.4	0.3	1.1	0.9	4.9	4.4	5.3	5.9	8	15	11	28	3.4	7.6	1.2	2.5	2.5
70F	1	0.7	0.5	0.5	2.2	1	7.3	5.6	8.3	2.4	19	6.3	29	9.2	5.8	1.7	1.7
70M	0.3	1.3	0.3	1.8	1.3	2.3	7.9	9.2	2	14	7.2	14	9.8	21	3.6	4.6	4.6
80F	1	1	2.4	1	3.9	2.9	7.7	6.8	11	5.8	13	4.8	16	7.7	12	4.8	4.8
80M	5.9	2	5.9	2	5.9	3	5.9	4	5.9	4	5.9	28	4	7.9	4	5.9	5.9

Figure 3: Segment to segment homophily distributions. Rows egocentrically represent the desired distribution by percentage. Only rows sum to 100%.

Distance data, however, is addressed at the network-level through a power-law probability distribution function for all edges, as initially demonstrated in Illenberger et al. (2013). The probability of forming an edge of distance d between any two agents is described by the following distribution function:

$$P(\text{Edge Formation}|d) = \frac{(\alpha - 1)x_0^{(\alpha-1)}}{d^\alpha} \quad (1)$$

We specify the exponential and scale factor parameters, α and x_0 , by a least squares estimation on the distribution of snowball distance edges that would fit inside the Munich study area. An

additional normalizing factor is introduced in the code to ensure that the cumulative distribution function converges to one within the study area.

Social network formation

The social network generation algorithm uses the segmented snowball data distributions and the SILO synthetic population as inputs. The formation method broadly follows the iterative three-step process displayed in Figure 4. In short, a friend-goal dictionary is generated for each agent based on the agent’s segment-based age and gender homophilies. Then, the algorithm matches two mutually compatible agents based on these dictionaries. A connection is only formed if the agents pass a stochastic draw based on the geographic distance between them. If a connection forms, each agent updates their friend-goal dictionary before searching for further connections. After all agents have entered the matching stage or no new connections can be formed, unsatisfied agents redraw their friend-goal dictionaries, restarting the cycle.

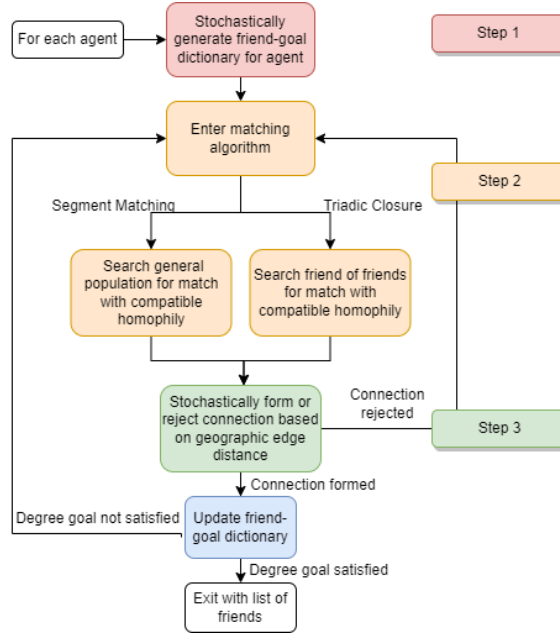


Figure 4: Social network generation steps for each iteration. Friend goal dictionary periodically re-drawn to avoid mismatches.

More specifically, in Step 1, each agent calculates how many degrees, or friends, k , they need to reach their total degree goal, which is assigned during population synthesis. Each agent then draws k times from their homophily distribution to form their ‘friend-goal’ dictionary. This provides the segments from which each agent is willing to search for connections during Step 2. Figure 5 illustrates two example agents after an initial draw.

Agent A	Gender	Age	Segment	*eligible segments highlighted												Current degree
	Female	41	40F													0
Segments eligible for connection	10F	10M	20F	20M	30F	30M	40F	40M	50F	50M	60F	60M	70F	70M	80F	Degree goal
Number of desired connection	1	0	1	0	0	2	8	1	2	0	1	0	1	0	0	17
Agent B	Gender	Age	Segment													Current degree
	Female	22	20F													0
Segments eligible for connection	10F	10M	20F	20M	30F	30M	40F	40M	50F	50M	60F	60M	70F	70M	80F	Degree goal
Number of desired connection	0	0	3	2	1	2	1	0	1	0	0	0	0	0	0	10

Figure 5: Sample agents after initial draw.

In Step 2, the algorithm matches mutually eligible agents. For example, in Figure 5, Agent A is in the 40F segment and is looking for one connection in the 20F segment; Agent B is in the 20F segment and is looking for a connection in the 40F segment. They are mutually compatible and could be matched.

During Step 2, agents search for matches from the general population *or* from their 2nd-degree connections - i.e., friends of their friends. The latter acts as a heuristic for triadic closure. Any con-

nection formed by triadic closure automatically creates a clique of at least three while maintaining the proper homophilies for all clique members; this provides an agent-based implementation of a concept explored in Asikainen et al. (2020).

The algorithm currently forms an initial proportion of agents’ degree goals by matching from the general population. It then switches to triadic closure, which reduces the search space but enables clique formation. When formation progress stagnates, the algorithm switches back to the general population search to complete the remaining connections.

After any potential match is found, this connection enters Step 3 where it is accepted or rejected based on the distance between the two agents. Equation (1) generates a connection probability for this distance, which is then compared to the result of a uniform draw from $(0, 1)$. If the edge is formed, each agent removes one from the appropriate segment of their friend-goal. If this completely satisfies either agent’s degree goal, that agent exits the algorithm. Otherwise, each agent returns to Step 2 to search for new connections.

After an iteration exhausts its possible matches, all remaining agents stochastically redraw their friend-goal dictionaries via Step 1 and the process repeats. Because of asymmetries in the homophily matrix and the demographic variation between the snowball and synthetic populations, these stochastic iterations are necessary for everyone to achieve their desired degree. By repeatedly drawing from the segment homophilies, the algorithm induces an equilibrium that balances these asymmetries.

3 RESULTS AND DISCUSSION

The current algorithm is built in Python and makes use of the Networkx package (Hagberg et al., 2008). The algorithm scales roughly as $\mathcal{O}(n^2)$ and runs in 30 minutes, generating 1,786,885 edges for 202,401 nodes.

Segment Based Homophilies

The aggregate segment-level connections are demonstrated in Figure 6. In general, the core adult population homophily distributions are well obeyed; this makes sense as they are well represented in the snowball data and have sufficient connection possibilities within their desired segments. The significantly older and significantly younger segments, which were underrepresented in the input data, perform worse. For example, segment $10F$ has too many internal connections and not enough connections to older segments; simultaneously, the other segments largely meet their goals regarding segment $10F$. This ties back to the asymmetries in the homophily matrix (Figure 3) where, for example, $M_{10F,30F} = 12\%$ but $M_{30F,10F} = 1\%$. In these cases, one segment meets its goal easily and has no need to surpass it, while the other tends to form more internal connections.

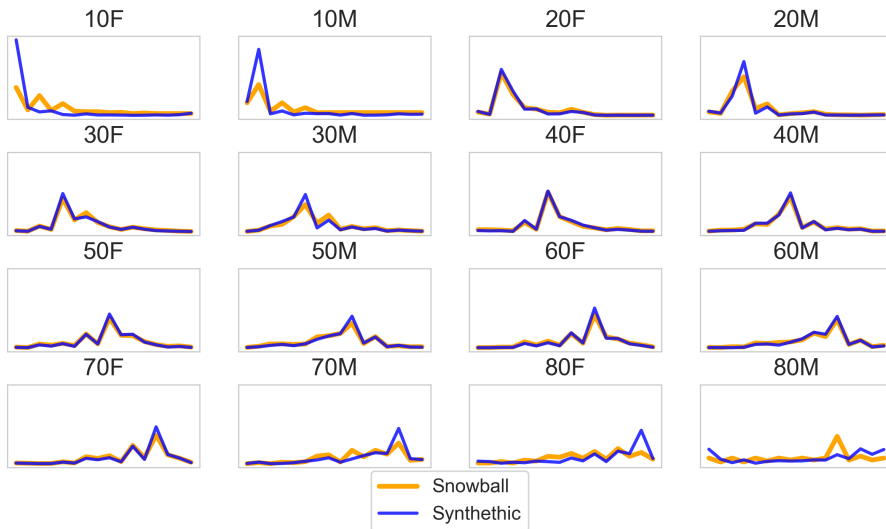


Figure 6: Input and realized homophily distributions by segment.

Figure 7 summarizes the egocentric distribution of gender similarity and age difference. While the synthetic network does exhibit slightly higher degrees of homophily - i.e., higher gender similarity and lower average age difference - it still captures activity at the tails of each distribution.

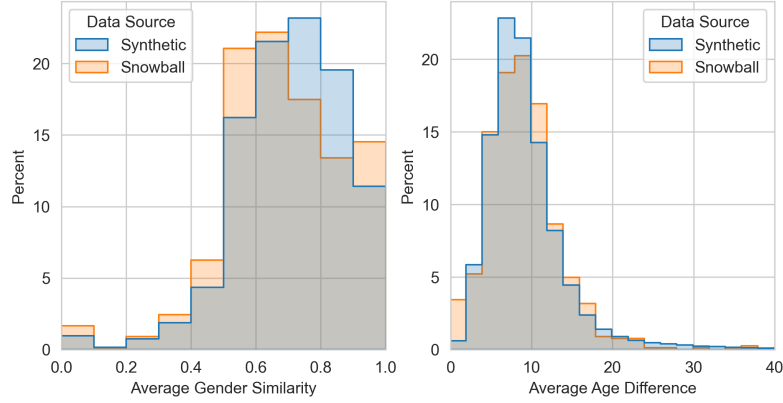


Figure 7: Average egocentric gender and age homophilies.

Distance Distribution Matching

Overall, the approach to distance formation reliably replicates the expected geographic edge length distribution. Figure 8 demonstrates the realized distribution of edge lengths for the entire network compared to the snowball data. There is a slight under-representation of edges less than 3 km as the basic power law formulation cannot be sensitive to distances smaller than its scale parameter, x_0 - which was set as ~ 1.5 km based on the snowball data - without causing its probability integral to diverge. Additionally, the synthetic distribution approaches 100% earlier than the snowball data, though this is likely an edge effect, given that only agents near the border of the study area can form the longest connections.

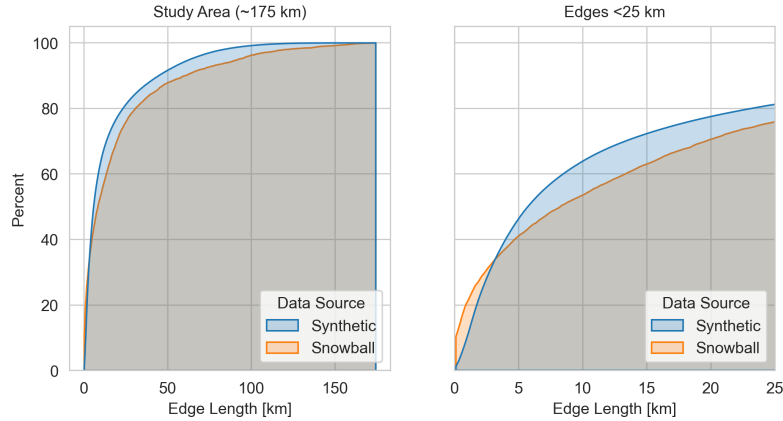


Figure 8: Edge length distribution for synthetic population.

Initial Clique Formation

The initial triadic closure methodology provides a basic mechanism for clique formation. Roughly 82% of agents have at least one clique in the full network, though the method significantly overforms small cliques compared to the input data. While 2nd-degree connections who have multiple mutual friends with the searching agent are duplicated in the triadic closure search space, providing a slight incentive to form larger cliques, these alters are given no prioritization in the search nor do they gain any additional homophily or distance flexibility. Future implementations will focus on strategies to align the clique distributions shown in Figure 9.

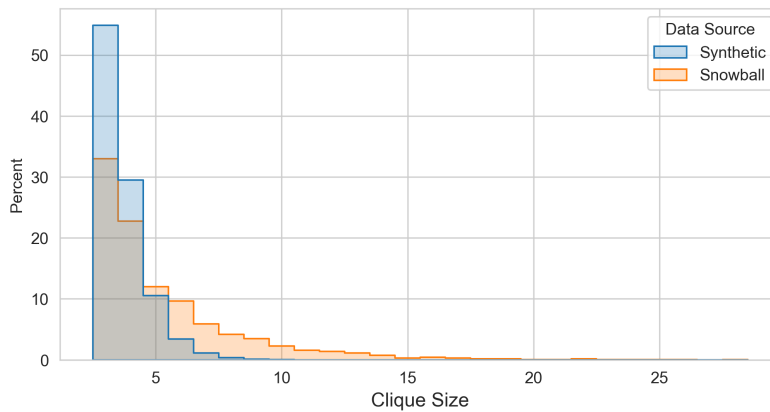


Figure 9: Clique size distribution.

Preferential Attachment Mechanism

Thanks to seminal work by Barabási & Albert (1999), many methodologies for social network formation rely on the concept of preferential attachment. Preferential attachment refers to the tendency of new nodes to connect to high-degree nodes during network formation or evolution. This gives many network degree distributions long right tails. While the snowball data, which limits respondents to 40 names, does not reflect this type of ‘scale-free’ distribution, the formation model still exhibits degree correlation. Figure 10 demonstrates a positive correlation between an agent’s degree and the average degree of its neighbors. This likely emerges from the algorithm’s iterative nature. Higher-degree agents generally remain in the algorithm for more iterations; the further into the simulation they get, the more often they encounter other agents with a similarly large degree goal.

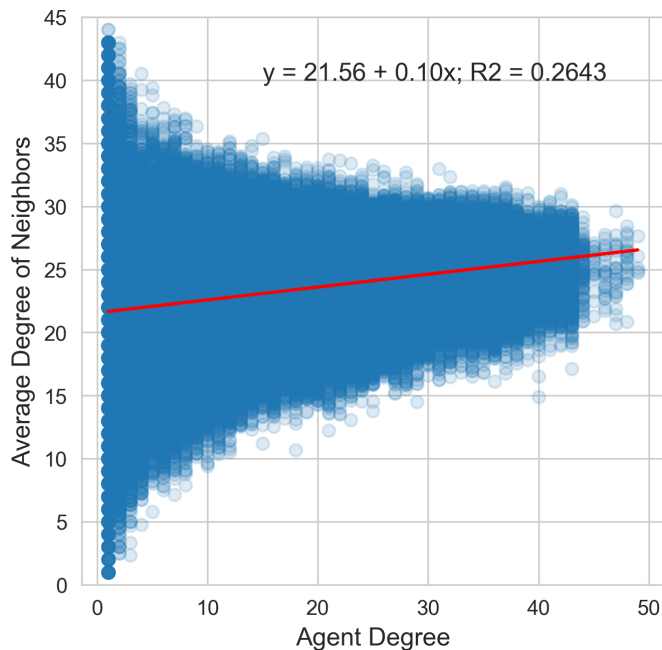


Figure 10: Agent degree versus average neighbor degree.

4 CONCLUSIONS

This paper presents a scalable method for generating a synthetic social network with relevant socio-demographic and geospatial characteristics based on a small, egocentric data sample. The generated network demonstrates homophily matching, has clique structure and shows preferential

attachment. A point of improvement would be better clique formation, as the current approach does not result in a clique distribution similar to the input data. Obeying egocentric homophilies while forming very large cliques is statistically unlikely in the current, agent-based framework; perhaps this approach can be blended with the clique-centric formation strategies of Dubernet (2017). Literature in network formation also suggests options related to clustering and community detection (Girvan & Newman, 2002). Additionally, the generated network is static, representing one point in time. Future research could focus on dynamic updating of the social network over time. Lastly, further work on travel behavior analysis should be conducted to assess how social networks influence joint travel decisions, as 45% of trips in Germany were performed with at least one companion (Infas et al., 2018). Further research in synthetic social network generation is a necessary step toward modeling the influence of social networks on travel behavior.

ACKNOWLEDGEMENTS

The research was funded by the German Research Foundation (DFG) under the research project TENGOS: “Transport and Epidemic Networks: Graphs, Optimization and Simulation” (Project number 458548755).

REFERENCES

- Arentze, T., Kowald, M., & Axhausen, K. (2012). A method to model population-wide social networks for large scale activity-travel micro-simulations. Retrieved from <https://doi.org/10.3929/ethz-b-000038374> doi: 10.3929/ethz-b-000038374
- Asikainen, A., Iñiguez, G., Ureña-Carrión, J., Kaski, K., & Kivelä, M. (2020). Cumulative effects of triadic closure and homophily in social networks. *Science Advances*, 6(19). Retrieved from <https://www.science.org/doi/abs/10.1126/sciadv.aax7310> doi: 10.1126/sciadv.aax7310
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509-512. Retrieved from <https://www.science.org/doi/abs/10.1126/science.286.5439.509> doi: 10.1126/science.286.5439.509
- Carrasco, J. A., Hogan, B., Wellman, B., & Miller, E. J. (2008). Collecting social network data to study social activity-travel behavior: An egocentric approach. *Environment and Planning B: Planning and Design*, 35, 961-980. doi: 10.1068/b3317t
- Dubernet, T. (2017). Explicitly correlating agent’s daily plans in a multiagent transport simulation: Towards the consideration of social relationships. Retrieved from <https://doi.org/10.3929/ethz-b-000165685> doi: 10.3929/ethz-b-000165685
- Frei, A., & Axhausen, K. W. (2007). Size and structure of social network geographies. Retrieved from <https://doi.org/10.3929/ethz-a-005562753> doi: 10.3929/ethz-a-005562753
- Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821-7826. Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.122653799> doi: 10.1073/pnas.122653799
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th python in science conference* (p. 11 - 15). Pasadena, CA USA.
- Illenberger, J., Nagel, K., & Flötteröd, G. (2013). The role of spatial interaction in social networks. *Networks and Spatial Economics*, 13(3), 255-282. Retrieved from <https://doi.org/10.1007/s11067-012-9180-4> doi: 10.1007/s11067-012-9180-4
- Infas, DLR, IBT, & infas 360. (2018). *Mobilität in deutschland*. Retrieved 14.06.2019, from <http://www.mobilitaet-in-deutschland.de/publikationen2017.html>
- Kim, J., Rasouli, S., & Timmermans, H. J. (2018, 7). Social networks, social influence and activity-travel behaviour: a review of models and empirical evidence. *Transport Reviews*, 38, 499-523. doi: 10.1080/01441647.2017.1351500

- Kowald, M., & Axhausen, K. (2012). Focusing on connected personal leisure networks: Selected results from a snowball sample. *Environment and Planning A: Economy and Space*, 44(5), 1085-1100. Retrieved from <https://doi.org/10.1068/a43458> doi: 10.1068/a43458
- Moreno, A., & Moeckel, R. (2018). Population Synthesis Handling Three Geographical Resolutions. *ISPRS International Journal of Geo-Information*, 7(5), 174. doi: 10.3390/ijgi7050174

Simulation of Mixtures of Legacy and Autonomous Mainline Rail Operations

Emily J. Morey*¹, R. Eddie Wilson², and Kevin Galvin³

¹PhD Student, Department of Engineering Mathematics, University of Bristol, Bristol, United Kingdom

²Prof. & Chair in Intelligent Transport Systems, Department of Engineering Mathematics, University of Bristol, Bristol, United Kingdom

³Systems Architect, Research, Technology & Innovation, Thales UK, Reading, United Kingdom

SHORT SUMMARY

This paper describes the analysis and development of a prototype (and first of its kind) simulator which models mixtures of legacy and autonomous mainline rail operations. This is achieved by using linked blocks to virtually superimpose occupancy information on two identical length tracks, one operating purely legacy trains and the other purely autonomous trains. These combine to form a single track running mixed operations. There are some surprising findings: the introduction of autonomous trains is not beneficial to system operations at all parameter settings, and moreover, there is a question of fairness to legacy train operations that may be adversely affected.

Keywords: Autonomous, Legacy, Modelling & Simulation, Railway.

1 INTRODUCTION & BACKGROUND

This paper explores how to simulate the introduction of Connected and/or Autonomous Trains (CATs) with existing Driver Operated and/or Guided trains (DOGs). Within the rail industry, there are a variety of software packages which simulate rail operations, such as OpenTrack (Huerlimann & Nash, 2017). However, although OpenTrack and others are able to model various levels of European Train Control Systems (ETCS) and moving blocks, they cannot model mixtures of different levels of ETCS on the same section of track. Hence, we have built a new train simulator from the ground up, to address this gap in extant capability.

In current rail operations, track is divided into sections known as *blocks*. To ensure safe separation, only one train is allowed in any one block at any given time. Therefore, block occupancy information needs to be conveyed to train drivers. This is achieved through trackside signalling or cab signalling, which provide movement authorisations (Pachl, 2020). Trackside signalling uses *aspects* to convey occupancy information at discrete points along the track, and depending on the number of aspects, governs the number of blocks the driver can (in effect) see ahead (Theeg & Vlasenko, 2020). Cab signalling delivers occupancy information directly to the driver, continuously in time. In effect, whether a train is a DOG or a CAT determines how much occupancy information is conveyed and the frequency at which it is delivered. This point is further explained in Section 2, which gives further detail on how we model CATs.

To ensure safe separation, any block-based simulator has to consider how a train reserves blocks on its braking path. In our simulator, this is achieved through a *watch-point* (see Fig. 1). When the watch-point progresses into a new block, it receives the new block's occupancy information, hence mimicking trackside signals (receiving information at discrete points). If the block is free, it is reserved and the train proceeds, whereas it must begin braking if the block is already occupied or reserved by another train. The key trick in this paper is to generalise this principle to model mixed DOG and CAT operations.

The paper is organised as follows. Firstly, Section 2 explains how virtual and linked blocks are employed to develop our mixed legacy-autonomous rail simulator. Section 3 then describes a highly simplified track simulation setup which we use for an initial inspection of mixed running performance. Section 4 presents and analyses the simulation results. Lastly, Section 5 provides

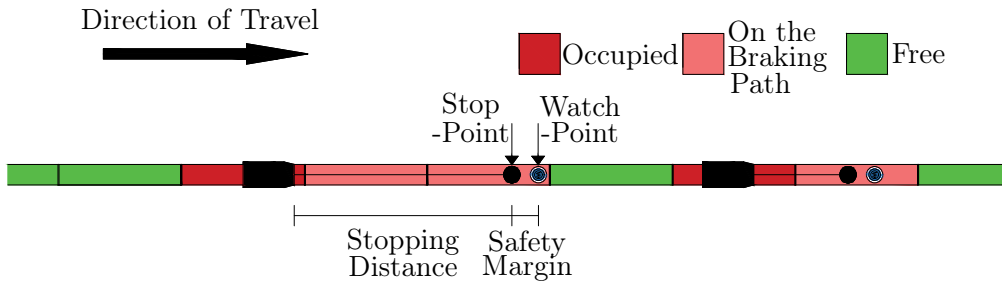


Figure 1: Simulation concepts. The train’s watch-point (eye symbol), stop-point (black circle), stopping distance (horizontal black line), and the block states differentiated by colour. Blocks on the braking path are reserved so that other trains may not enter them.

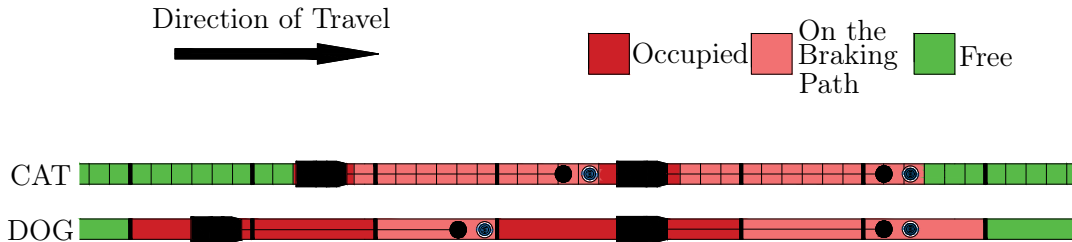


Figure 2: DOGs and CATs simulation setup. Both use fixed blocks, but CATs operate on shorter blocks than DOGs. Here six CAT blocks fit into a single DOG block. CATs can thus operate at reduced headway.

a wider discussion of linked-block principles and the simulation findings, and Section 6 describes conclusions and future work plans.

2 SIMULATION PRINCIPLES FOR MIXED OPERATIONS

This section explains the method we use to model mixed legacy-autonomous operations. As explained above, DOGs operate using fixed blocks, where occupancy information is conveyed at discrete points along the track via trackside signals.

In contrast, it is assumed that CATs can directly and continuously relay their speeds and positions to each other and this data is used to establish a buffer zone ahead of each CAT (the *moving block*) whose length equals the stopping distance plus a safety margin. The buffer zone moves along the track with the CAT and if it comes up against an obstacle, the CAT brakes to ensure safe separation (Pachl, 2020).

In our simulator, we model the CAT’s buffer zone as a union of *virtual* fixed blocks, of shorter lengths than used for the DOGs’ blocks. Due to the shorter block lengths, the CAT receives new occupancy information at a high temporal frequency and high spatial precision which approximates an update which is continuous in space and time. See Fig. 2. CATs can thus operate at reduced headway when following other CATs.

To model mixed running, we use two tracks of identical lengths, where one track operates purely DOGs and the other operates purely CATs with shorter virtual blocks. These tracks are *linked* through a *lookup* process, where one track’s occupancy information is virtually superimposed onto the other and vice versa. See Figs. 3& 4.

Fig. 3 shows a CAT following a DOG. Each virtual block on the CAT’s track performs a lookup on the corresponding block of the DOG’s track. The DOG occupies / reserves three blocks (Fig. 3(a)). Each of these DOG blocks is inspected by six virtual blocks on the CAT track (Fig. 3(b)), which are thus effectively considered reserved (Fig. 3(c)). The CAT will commence braking when its watch-point reaches the first of these reserved virtual blocks. See Fig. 3(d). Here the CAT may continue for another seven virtual blocks before it commences braking. Note that the minimum

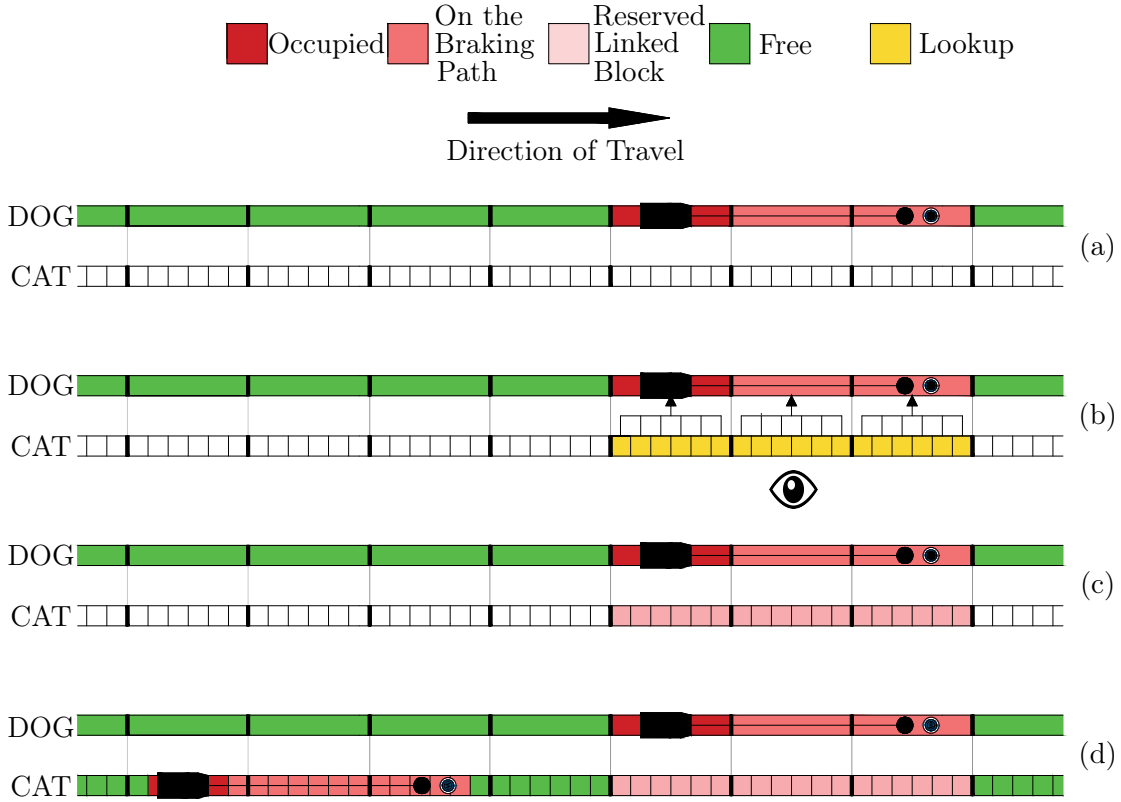


Figure 3: CAT follows DOG. Through the use of linked blocks and a lookup process (eye symbol, yellow blocks, and arrows) the occupancy of the corresponding blocks on the DOG’s track are virtually superimposed onto the CAT’s track.

headway is greater than if a CAT were following a CAT, because of the way that DOG blocks coarsely reserve corresponding virtual blocks irrespective of the continuous position of the DOG (which is not communicated and is thus unknown).

Fig. 4 shows a DOG following a CAT. Each DOG block performs a lookup on the six corresponding virtual blocks of the CAT track, and is marked occupied or reserved if any of those virtual blocks are occupied or reserved. As for CAT following DOG, the minimum headway is greater than for CAT following CAT, now because the following DOG is not able to receive the fine-scale position information communicated by the CAT.

3 SIMULATOR SETUP

Using the virtual and linked block principles introduced above and the architectural framework described by Morey et al. (2023), a simple time-stepping simulation has been developed to examine the dynamics of mixed legacy-autonomous operations at a converging track section.

The simulation follows many of the classical theoretical studies in car-following modelling (e.g., Bando et al. (1995)) where a set of identical trains repeatedly circulates around a *ring track*, which consists of identical blocks. Here the track setup consists of two ring track sections of equal length, which have an overlapping joint section, see Fig. 5. Each ring operates purely DOGs (section S1) or purely CATs (section S2), and their joint section (section S3) operates mixed running through the addition of linked virtual blocks (section S4). This simple layout induces the mixing, reordering, and self-organising of different train types, and therefore enables a first examination of some of the possible complex dynamics that might result from mixed operations.

In each simulation, the number of trains is prescribed and conserved over time — so (in effect) we prescribe the train density as an input parameter, and we examine the traffic patterns and

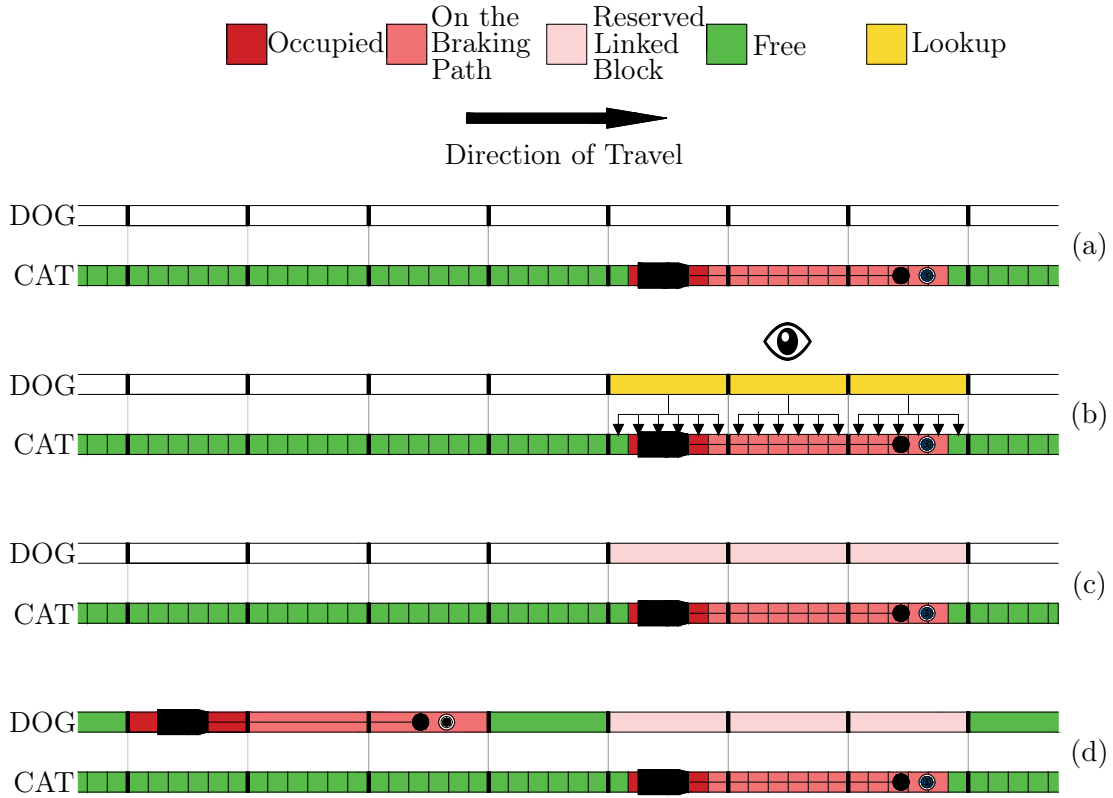


Figure 4: DOG follows CAT. Compare Fig. 3, where CAT follows DOG. The lookup process is now used to superimpose the occupancy of the blocks on the CAT's track onto the DOG's track.

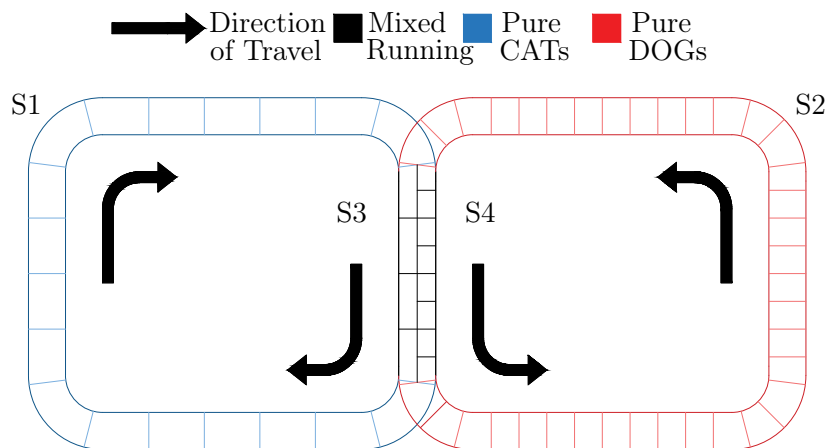


Figure 5: Simulator track layout. Two ring tracks overlap to form a joint section (S3 & S4, black) which operates mixed running. Sections S1 (blue) and S3 have pure DOG operations, whereas S2 (red) and S4 have pure CAT operations. Schematic only — lengths of sections not to proportion shown here, and we employ virtual blocks that are ten times shorter than DOG blocks.

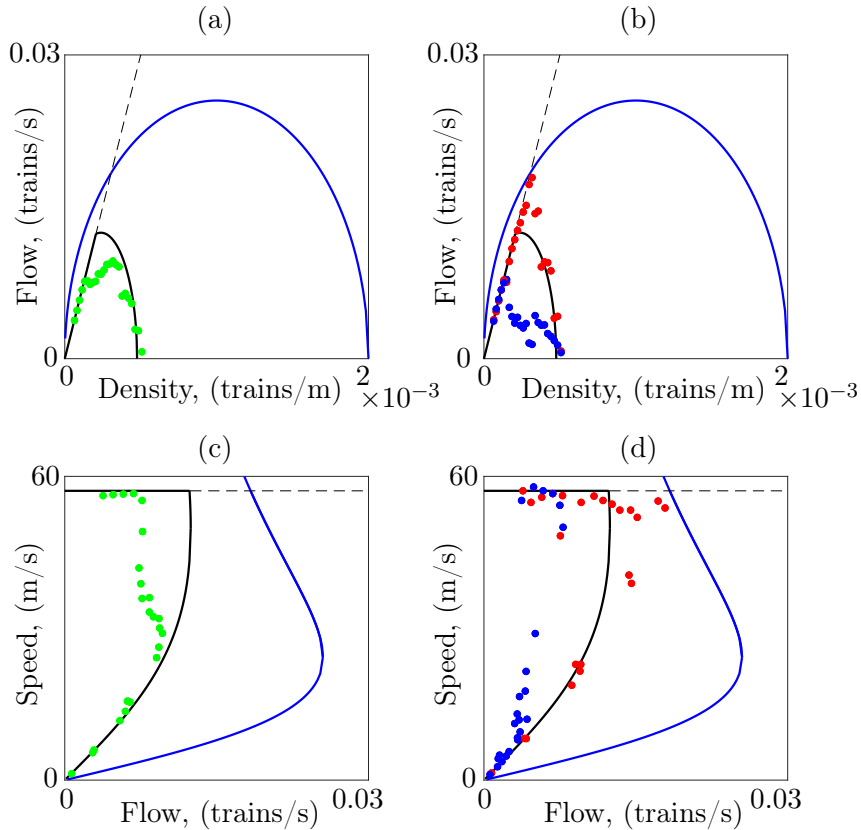


Figure 6: Fundamental Diagrams (FDs) for mixed mainline rail traffic. The black and blue lines show theoretical predictions for pure DOG and CAT operations respectively. Dots show mixed running simulation results: green (mixed population average), blue (DOG population average), and red (CAT population average). (a & b) flow-density graphs and (c & d) speed-flow graphs.

performance metrics (e.g., the time-average speed and the flow) that result. Our simulations begin with all trains at rest randomly spaced around their respective rings. For the simulations shown here, for simplicity, we prescribe an equal number of CATs and DOGs, however, of course, the penetration rate of CATs may also be varied as an input parameter.

The parameters used in the simulations are, see Morey et al. (2022a): trains’ acceleration 0.4 ms^{-2} ; braking rate 0.65 ms^{-2} ; maximum speed 60 ms^{-1} ; train length 400 m; number of blocks in sections, S1 & S3 20; S2 & S4 200; DOG block length 1600 m; virtual (CAT) block length 160 m. Thus each ring track is 64km long and the joint section is 32km long. We varied the number of trains from 4 to 27 of each type, yielding densities from 1.60×10^{-5} trains/m to 5.32×10^{-4} trains/m.

4 SIMULATION RESULTS

From simulator output, we derived fundamental diagrams (FDs) that relate density, flow, and speed averaged over the entire network. See Fig. 6. Results are compared with theoretical bounds derived by Morey et al. (2022a) and Morey et al. (2022b), which suggest that pure CAT operations might potentially have double the capacity (maximum flow) of pure DOG operations.

Note in our network that capacity is constrained by the mixed running section which is in essence a bottleneck and generally, flow falls short of even the pure-DOG bound, because there is interrupted flow induced by the point where sections S1 and S2 merge, which becomes quite acute at intermediate density ranges.

However, Figs. 6(b&d) show that the DOG and CAT populations have quite different experiences. Whereas the CAT population develops flow rates that exceed the pure-DOG bound, and even

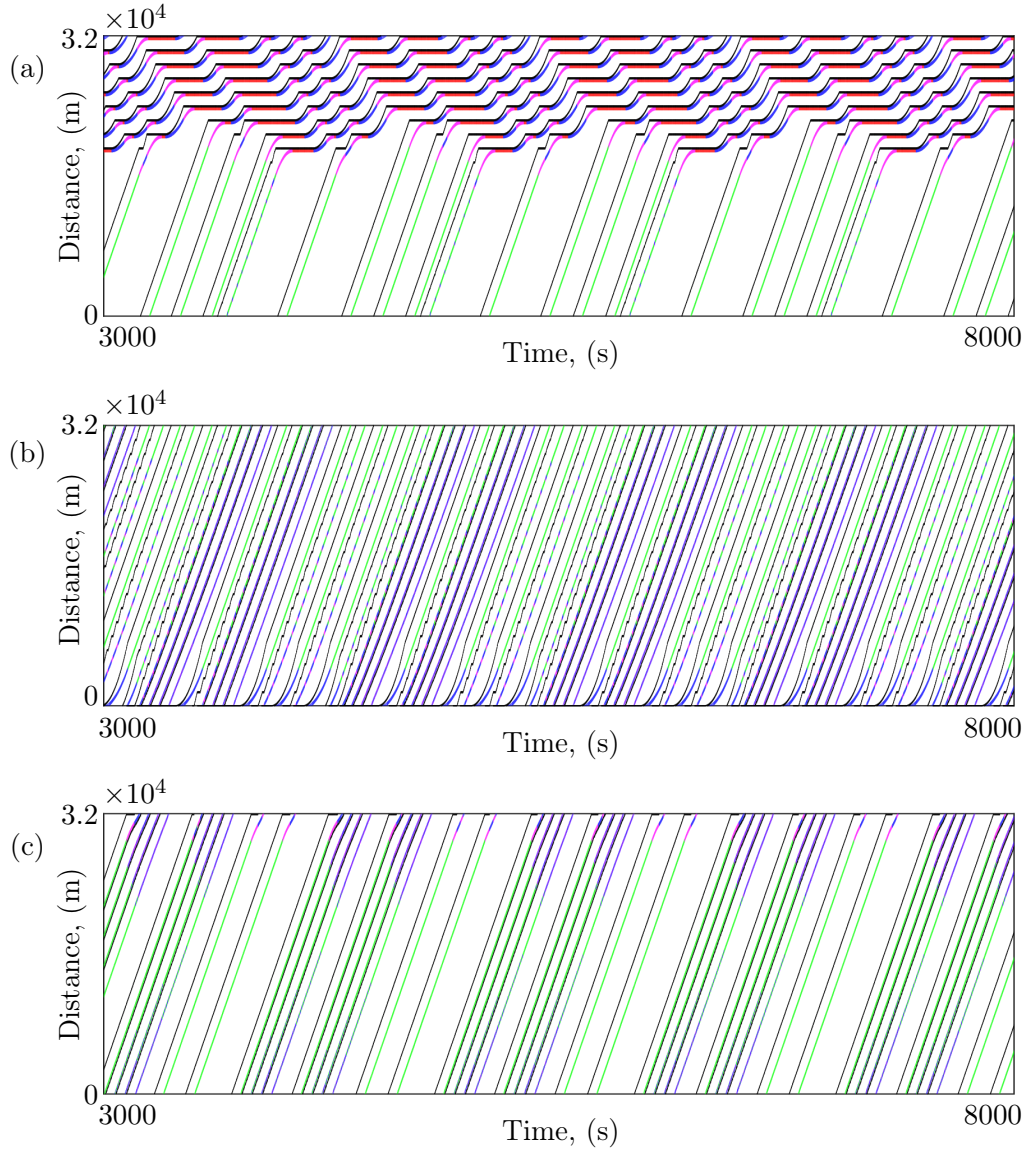


Figure 7: Trajectory plots showing mixed running of 10 DOGs and 10 CATs. Colour represents each train’s state: green (at goal speed), magenta (braking), red (halted), and blue (accelerating). Sections (a) S1, (b) S3 & S4, (c) S2.

approach the pure-CAT bound at lower density ranges, the DOG population suffers a dramatic reduction in speed and flow, which may be explained by the exemplar trajectory plots shown in Fig. 7).

Over time, the CATs tend to self-organise into platoons, see Fig. 7(c) which are never subsequently broken up by the DOGs. This is because the gaps between consecutive CATs are so short that a DOG approaching the end of section S1 must wait for the CAT platoon pass before it can join the mixed running section. In consequence, DOGs tend to form queues at the end of section S1, see Fig. 7(a), which can only be discharged in the gaps between the CAT platoons. Therefore DOGs also tend to form platoons over time, which fit into the gaps between the CAT platoons on the mixed running section.

The overall effect is unfair. Without additional control policies, the CATs achieve their enhanced throughput entirely at the expense of the DOGs.

5 DISCUSSION

The use of virtual and linked blocks that we have demonstrated here seems to be quite an elegant solution to extend existing fixed-block simulators to deal with mixed running scenarios. In fact, the

trick can also be used as a simple device to help simulate other network features, e.g.: 1. Two-way single-track sections (by a union of two parallel one-way sections where each block in one direction watches all of the blocks in the opposing direction — thus a train is prevented from entering the section if any block on it is occupied or reserved by a train travelling in the opposite direction); 2. More complicated mixtures of future train types, perhaps with different communication rates/latencies and/or braking capabilities, and thus different types of moving zones; 3. Complex routing patterns where an individual train might choose different paths at a diverge depending on previous history (without that history needing to be stored as a property of the train in question).

Clearly, the results shown here are only an initial analysis of the simplest possible network which allows mixed running with the potential for the ordering of CATs and DOGs to change over time. Of course, in the traffic flow theory community, there has been much interest in macroscopic/network fundamental diagrams that describe urban road network dynamics, and some recent papers have begun to explore these concepts in rail networks, e.g., Corman et al. (2019), Farhi et al. (2017), and Cuniasse et al. (2015). We have designed a sequence of test networks of increasing complexity and the immediate goal is to determine whether they display similar results to those we found here: i.e., that potential capacity gains might be quite unfairly distributed between CATs and DOGs.

A further question concerns the pathway to autonomy and the network performance as the penetration rate (i.e., proportion) p of CATs is increased slowly from zero. The anticipated benefits result from CAT-CAT leader-follower pairs operating at reduced headway. Without platooning or self organisation, the proportion of such pairs scales like p^2 — so the benefits might initially be very modest, to say the least. However, how this plays out in more complex networks with many merges and diverges remains to be seen.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we have described an elegant approach to simulate mixed legacy-autonomous rail operations. We have achieved this through the use of virtual blocks and a lookup process to link occupancy information. This technique provides the basic mechanism to simulate a wide variety of future rail scenarios. We have demonstrated exemplar results from a simple stylised network. The findings are surprising: it seems that the introduction of Connected and/or Autonomous Trains (CATs) does not necessarily increase capacity, and potentially, the system self organises so that high flow rates for CATs are achieved at the expense of lower flow rates for legacy trains.

Future work should of course involve experimenting with a wider range of track setups, incorporating more realism (e.g., station stops, complex routing) and more advanced control rules (e.g., prioritisation at track merges), and investigating varying penetration rates of CATs.

ACKNOWLEDGEMENTS

This work was funded and delivered in partnership between the Thales Group and the University of Bristol, and with the support of the UK Engineering and Physical Sciences Research Council Grant Award EP/R004757/1 entitled ‘Thales-Bristol Partnership in Hybrid Autonomous Systems Engineering (T-B PHASE)’.

REFERENCES

- Bando, M., Hasebe, K., Nakayama, A., Shibata, A., & Sugiyama, Y. (1995). Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, 51(2), 1035.
- Corman, F., Henken, J., & Keyvan-Ekbatani, M. (2019). Macroscopic fundamental diagrams for train operations — are we there yet? In *Proceedings of IEEE 2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)* (pp. 1–8).
- Cuniasse, P. A., Buisson, C., Rodriguez, J., Teboul, E., & De Almeida, D. (2015). Analyzing railroad congestion in a dense urban network through the use of a road traffic network fundamental diagram concept. *Public Transport*, 7(3), 355–367.

- Farhi, N., Van Phu, C. N., Haj-Salem, H., & Lebacque, J. P. (2017). Traffic modeling and real-time control for metro lines. Part I-A Max-plus algebra model explaining the traffic phases of the train dynamics. In *2017 American Control Conference (ACC)* (pp. 3834–3839).
- Huerlimann, D., & Nash, A. (2017). *Opentrack: Simulation of railway networks*. OpenTrack Railway Technology Ltd and ETH Zurich Institute for Transport Planning and Systems.
- Morey, E. J., Wilson, R. E., & Galvin, K. (2022a). Fundamental diagrams and emergent dynamics of mainline rail operations. Pure. Retrieved from <https://research-information.bris.ac.uk/en/persons/emily-j-morey/publications/>
- Morey, E. J., Wilson, R. E., & Galvin, K. (2022b). On a theory for potential capacity gains due to connected and autonomous trains. 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI). Retrieved from https://tristan2022.org/Papers/TRISTAN_2022_paper_0588.pdf
- Morey, E. J., Wilson, R. E., & Galvin, K. (2023). From problem space to executional architecture: how to develop a simulation to investigate the challenges associated with implementing autonomous rail to increase mainline railway capacity. Pure. Retrieved from <https://research-information.bris.ac.uk/en/persons/emily-j-morey/publications/>
- Pachl, J. (2020). *Railway signalling principles*. Technische Universität Braunschweig.
- Theeg, G., & Vlasenko, S. (2020). *Railway signalling & interlocking* (3rd ed.). Leverkusen: PMC Media International Publishing.

What do walking and e-hailing bring to scale economies? A general microeconomic model for on-demand mobility

Kenan Zhang^{1,*}, Andres Fielbaum^{2,3}, and Javier Alonso-Mora²

¹Automatic Control Laboratory, ETH Zürich, Switzerland

²Department of Cognitive Robotics, Delft University of Technology, The Netherlands

³School of Civil Engineering, University of Sydney, Australia

*Corresponding author and presenter: kenzhang@ethz.ch

SHORT SUMMARY

This paper investigates the impact of walking and e-hailing on the scale economies of on-demand mobility services. A microeconomic model is developed to explicitly characterize the physical interactions between passengers and vehicles in the matching, pickup, and walking processes under different market conditions and matching mechanisms. We show that passenger competition plays a critical role in scale economies. When unmatched passengers do not compete for idle vehicles, both street-hailing and e-hailing exhibit increasing returns to scale, although such property in e-hailing is less significant. In contrast, when there exists passenger competition, e-hailing service shows decreasing returns to scale. Street-hailing, however, is free of this detrimental effect thanks to its limited matching radius. While walking does not change the scale economies, it does benefit the system by reducing the total vehicle supply required to serve the same level of demand and improving the overall vehicle utilization rate.

Keywords: passenger-vehicle matching, on-demand mobility service, scale economies, walking.

1 INTRODUCTION

The emergence of ride-hailing companies, such as Uber and Didi, has revolutionized the industry of on-demand mobility services, which has long been dominated by taxis. With advanced mobile communication technologies, passengers and vehicles can now be connected online. Such an e-hailing matching mechanism is believed to be far more efficient than the traditional street-hailing that relies on visual contact on streets (Cramer & Krueger, 2016).

However, evidence from Shenzhen (China) suggests that, in a highly dense market, street-hailing can perform comparably well or even better than e-hailing (Nie, 2017). The unlimited connectivity in e-hailing has also shown negative impacts during demand peaks (Castillo et al., 2017). These observations motivate Zhang et al. (2019) to model the physical matching process in ride-hailing and analyze its scale effect. It concludes that street-hailing has better scale economies than e-hailing because i) its matching efficiency does not suffer from passenger competition as e-hailing, and ii) the nature of linear search is more prone to scale effect compared to spatial search in e-hailing. However, the analysis in Zhang et al. (2019) does not distinguish the matching and pickup processes and does not consider the network topology. Another missing factor is walking. In street-hailing, passengers often walk towards major streets to find taxis. On the other hand, e-hailing mostly serves door-to-door trips. Accordingly, sometimes drivers have to make long detours and enter local streets to pick up and drop off passengers. In fact, leaving the door-to-door scheme has shown great potential to increase the efficiency of shared on-demand systems by several previous studies (e.g. Fielbaum et al. (2021); Gurumurthy & Kockelman (2022)). These findings suggest walking might also play a role in the scale economies of ride-hailing.

This study thus sets out to analyze the scale economies in street-hailing and e-hailing at the system level and investigate the impact of walking. To this end, we model the ride-hailing market on a grid network with detailed specifications of each component in passenger and vehicle time. The scale economies are then evaluated according to how the system cost under optimal fleet size varies with the demand rate.

2 METHODOLOGY

Settings

Consider a grid network with two types of streets shown in Fig. 1. The major streets form the skeleton of the network. Between every two major streets, there are K local streets with equal spacing s . Accordingly, three types of intersections are identified: (i) Type-1: between major streets; (ii) Type-2: between local streets; and (iii) Type-3: between local and major streets.

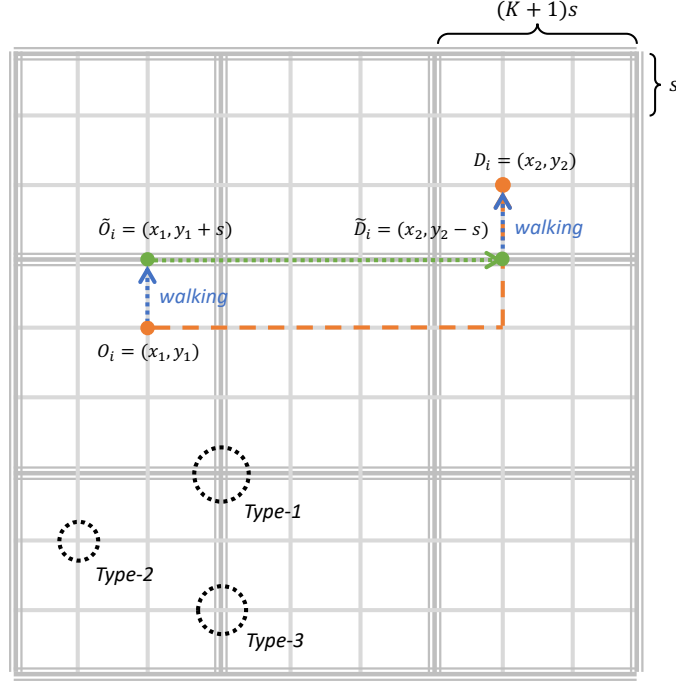


Figure 1: Illustration of a grid network.

Assume the travel speed on major streets is v and it takes $\delta = s/v$ to traverse a block. Two factors $\alpha_l, \alpha_a > 1$ are introduced to denote the vehicle speed on local streets and walking speed, respectively. In other words, it takes $\alpha_l \delta$ for a vehicle to pass a block on local streets and $\alpha_a \delta$ for a passenger to walk through a block.

Assume passenger demand and vehicle supply are both uniformly distributed in space and the market has reached a steady state. Then, the market conditions can be described by the idle vehicle density V , the unmatched passenger density W , and the demand rate q (i.e., the passenger arrival rate per unit area). As per the steady state condition, q also equals the matching and pickup rates.

Scenarios

In this study, we consider four scenarios: (i) street-hailing without walking (DS), (ii) street-hailing with walking (WS), (iii) e-hailing without walking (DE), and (iv) e-hailing with walking (DS). In all scenarios, trips are generated at Type-2 intersections (i.e., passengers travel between local blocks). When walking is considered, passengers are picked up and dropped off on major streets and thus pickups and dropoffs happen at Type-3 intersections. For simplicity, we assume passengers would randomly walk to one of the closest major streets, not necessarily the closest to their destinations.

On the supply side, we assume vehicles randomly cruise in the network. In DS , vehicles cruise on local streets to maximize the probability to find a passenger, while in other cases, they cruise on major streets.

Matching model

Here we present a general matching model and use $k \in \{DS, WS, DE, WE\}$ to denote the scenario. For both street-hailing and e-hailing, we first define the matching interval δ_k that denotes how

frequently matching is performed. We then define R_k and A_k as the matching radius and area, respectively, both in units of road links (arc). Specifically, street-hailing passengers can only see vehicles moving in four directions, and thus $A_k = 4R_k$, $k \in \{DS, WS\}$. For e-hail, the matching area can be derived as $A_k = 2(R_k^2 + R_k)$, $k \in \{DE, WE\}$.

To capture the possible competition among waiting passengers, we further introduce the notion of *dominant area* Yang et al. (2020), denoted by \tilde{A} , to represent the area within which any idle vehicle is for sure matched to the passenger. In this study, we specify the dominant area as follows:

$$\tilde{A}_k = \frac{A_k}{\gamma_k}, \quad \gamma_k = \max(A_k W, 1), \quad (1)$$

which implies the matching area is evenly distributed to unmatched passengers within A_k . With the assumptions introduced above, we can easily show the number of “matchable” vehicles at each matching instance follows a spatial Poisson process. Then, the expected matching time is derived as

$$w_m^k = \left(\frac{\gamma_k}{A_k \theta_k V} + \frac{1}{2} \right) \delta_k. \quad (2)$$

where the first term in the parentheses gives the expected number of matching intervals and the second accounts for the average elapsed time before the first matching instance. The density correction factor θ_k is introduced to reflect the accumulation of vehicles on a certain type of street. It only affects street-hailing whereas in e-hailing $\theta_k \equiv 1$.

The pickup time in street-hailing is simple thanks to the linear matching mechanism. In contrast, that in e-hailing is rather complicated because matching is performed in space and the potential passenger competition affects the effective matching radius R . With some algebra, we derived the expected pickup distance (in units of arcs) as follows:

$$D_p = \sum_{i=1}^{R-1} \left(1 - \frac{i^2 + i}{R^2 + R} \right)^{V(R^2 + R)}. \quad (3)$$

To make the model tractable, we introduce the following approximation and replace R with $R_k/\sqrt{\gamma_k}$ to capture the impact of passenger competition:

$$d_p^k = \frac{c_1}{V} - \frac{c_2 \sqrt{\gamma_k}}{V R_k}. \quad (4)$$

As shown in Fig. (2), where the approximation parameters are set to be $c_1 = c_2 = 1/6$, the approximation error diminishes rapidly with the idle vehicle density.

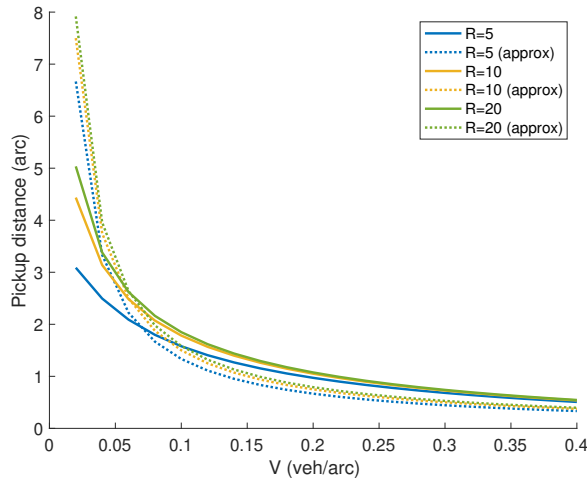


Figure 2: Comparison of exact and approximated pickup distance.

Finally, the expected pickup time is derived as follows

$$w_p^k = \alpha_p^k \left(d_p^k + d_a^k + \frac{1}{2} \right) \delta, \quad (5)$$

where

$$\alpha_p^k = \begin{cases} \alpha_l, & k = DS \\ 1, & k \in \{WS, DE, WE\} \end{cases}, \quad (6)$$

$$d_p^k = \begin{cases} R_k - 1, & k \in \{DS, WS\} \\ \frac{c_1}{V} - \frac{c_2\sqrt{\gamma_k}}{VR_k}, & k \in \{DE, WE\} \end{cases}, \quad (7)$$

$$d_a^k = \begin{cases} 0, & k \in \{DS, WS, WE\} \\ (\alpha_l - 1)D_a, & k = DE \end{cases}, \quad (8)$$

where D_a is the average walking distance derived as $D_a = \frac{(K+1)(K+2)}{6K}$. Finally, the total passenger waiting time is written as

$$w_k = w_a^k + w_m^k + w_p^k, \quad (9)$$

where the walking time is $w_a = \alpha_a D_a \delta$ in scenarios WS and WE , otherwise zero.

Another adjustment to make in the scenarios of walking is the in-vehicle time. Let $\bar{\tau}$ be the average door-to-door trip duration. Then, the in-vehicle time is given by

$$\tau_k = \begin{cases} \bar{\tau}, & k \in \{DS, DE\} \\ \bar{\tau} - 2\alpha_l D_a \delta, & k \in \{WS, WE\} \end{cases}. \quad (10)$$

Analysis of scale economies

Consider demand rate q as the only system input. Then, the scale economies can be expressed by how the system cost varies with q under the optimal fleet size N^* , which is solved from the following optimization problem:

$$\min_N c(N, q), \quad (11a)$$

$$s.t. \quad N = V + q(w_p + \tau), \quad (11b)$$

$$W = qw_m, \quad (11c)$$

$$w_p = f_{w_p}(V, W), \quad (11d)$$

$$w_m = f_{w_m}(V, W), \quad (11e)$$

The objective of (11) is the sum of vehicle operating cost and passenger travel cost:

$$c(N, q) = c_0 N + q(\beta_m w_m + \beta_p w_p + \beta_a w_a + \beta_\tau \tau), \quad (12)$$

where c_0 is the operation cost of each vehicle and β_j , $j \in \{m, p, a, \tau\}$ is the value of time specified for different legs of a trip.

To get N^* , we need to solve the following two implicit functions

$$W = qf_{w_m}(V, W) \Rightarrow W = f_W(V, q), \quad (13)$$

$$N = V + q(f_{w_p}(V, f_W(V, q)) + \tau) \Rightarrow V = f_V(N, q). \quad (14)$$

Accordingly, N^* can be derived from the first-order condition, which reads

$$0 = \frac{\partial c}{\partial N} = c_0 + q [\beta_m (\partial_V f_{w_m} + \partial_W f_{w_m} \partial_V f_W) + \beta_p (\partial_V f_{w_p} + \partial_W f_{w_p} \partial_V f_W)] \partial_N f_V. \quad (15)$$

Let $c^*(q)$ denote the system cost with the optimal fleet, then a market exhibits increasing (constant/decreasing) returns to scale if its marginal cost decreases (does not change/increases) with input.

3 RESULTS AND DISCUSSION

Scale economies in different scenarios

Note that the walking distance d_a is exogenously determined by the network property (K) rather than the demand-supply relationship. Neither does it affect other endogenous variables in the

market (see Eqs. (11b)-(11e)). Hence, walking does not fundamentally change the property of returns to scale. On the other hand, the passenger competition does make a difference, particularly for e-hailing. Hence, in what follows, we investigate the scale economies for street-hailing without passenger competition, e-hailing without passenger competition, and e-hailing with passenger competition. The case of street-hailing with passenger competition is neglected because it rarely happens in practice, which will be further discussed in the next section.

Street-hailing without passenger competition

In this case, the matching time is independent of W while the pickup time is independent of V . Hence, we can easily derived the optimal fleet size is solved as

$$N^* = \sqrt{\frac{\beta_m \delta_k q}{c_0 A_k \theta_k}} + \left[\alpha_p^k \left(R_k - \frac{1}{2} \right) \delta + \tau_k \right] q, \quad (16)$$

which yields the system cost

$$c^* = 2c_0 \sqrt{\left(\frac{\beta_m \delta_k}{c_0 A_k \theta_k} \right)} q + \left[\frac{\beta_m \delta_k}{2} + (c_0 + \beta_p) \alpha_p^k \left(R_k - \frac{1}{2} \right) \delta + \beta_a w_a^k + (c_0 + \beta_\tau) \tau_k \right] q \quad (17)$$

E-hailing without passenger competition

In this case, the matching time is still independent of W whereas the pickup time becomes a function of V . Thus, we first solve the implicit function Eq. (14). With some algebra, we derive the following equations:

$$V = \frac{N - \hat{\tau}_k q}{2} + \sqrt{\left(\frac{N - \hat{\tau}_k q}{2} \right)^2 - \left(c_1 - \frac{c_2}{R_k} \right) \delta} q, \quad (18)$$

$$\partial_N f_V = \frac{V^2}{V^2 - \left(c_1 - \frac{c_2}{R_k} \right) q \delta}, \quad (19)$$

where $\hat{\tau}_k = (d_a^k - \frac{1}{2}) \delta + \tau_k$ can be interpreted as the exogenous trip duration. Plugging Eqs. (18) and (19) into Eq. (15) yields

$$V^* = \sqrt{\left[\frac{\beta_m \delta_k}{c_0 A_k} + \left(1 + \frac{\beta_p}{c_0} \right) \left(c_1 - \frac{c_2}{R_k} \right) \delta \right]} q. \quad (20)$$

Accordingly, the optimal system cost is given by

$$\begin{aligned} c^* &= 2c_0 V^* + \left[\frac{\beta_m \delta_k}{2} + (c_0 + \beta_p) \left(d_a^k + \frac{1}{2} \right) \delta + \beta_a w_a^k + (c_0 + \beta_\tau) \tau_k \right] q \\ &= 2c_0 \sqrt{\left[\frac{\beta_m \delta_k}{c_0 A_k} + \left(1 + \frac{\beta_p}{c_0} \right) \left(c_1 - \frac{c_2}{R_k} \right) \delta \right]} q \\ &\quad + \left[\frac{\beta_m \delta_k}{2} + (c_0 + \beta_p) \left(d_a^k + \frac{1}{2} \right) \delta + \beta_a w_a^k + (c_0 + \beta_\tau) \tau_k \right] q \end{aligned} \quad (21)$$

E-hailing with passenger competition

In this case, we need to solve both implicit functions:

$$W = \frac{\delta_k q V}{2(V - \delta_k q)} \quad (22)$$

$$V = \frac{N - \hat{\tau}_k q}{2} + \sqrt{\left(\frac{N - \hat{\tau}_k q}{2} \right)^2 - \left(c_1 - \frac{c_2 \sqrt{A_k W}}{R_k} \right) q}. \quad (23)$$

Note that Eq. (23) still involves W and thus $\partial_N f_V$ cannot be directly solved as in the previous cases. Instead, we consider the right-hand-side of Eq. (23) as a function $h_V(N, W)$ and conduct implicit differentiation. Accordingly,

$$\begin{aligned} \partial_N f_V &= \frac{\partial_N h_V}{1 - (\partial_W h_V)(\partial_V f_W)} \\ &= \frac{\partial_N h_V}{V^2 - \left[c_1 - \frac{c_2}{R_k} \left(1 + \frac{W}{V} \right) \sqrt{A_k W} \right] q \delta} \end{aligned} \quad (24)$$

Plugging into Eq. (15) yields

$$V^* = \sqrt{\left\{ \frac{\beta_m \delta_k W^*}{c_0} \left(1 + \frac{2W^*}{V^*} \right) + \left(1 + \frac{\beta_p}{c_0} \right) \left[c_1 - \frac{c_2}{R_k} \left(1 + \frac{W^*}{V^*} \right) \sqrt{A_k W^*} \right] \delta \right\} q}, \quad (25)$$

where W^* is the unmatched passenger density under V^* as per Eq. (22). Thus, solving Eq. (25) gives us V^* as a function of q . However, the exact form of such an equation is challenging to derive. Instead, we numerically solve V^* with various q and explore their relationship. The default parameters used in these numerical experiments are reported in Tab. 1.

Fig. 3 illustrates both sides of Eq. (25) as a function of V under different demand rates q . V^* is then given by the intersection of two curves, which increases with q as expected. We then numerically solve V^* at each q by bisection search and plot the results in Fig. 4, along with those in e-hailing without competition. It can be observed that V^* with passenger competition first increases sublinearly below the case without passenger competition (Eq. (20)). This is expected as fewer vehicles are required when holding some passengers waiting. As q continues to increase, however, the relationship becomes linear and V^* with competition exceeds the other case. As will be shown later, this seemingly counter-intuitive result is due to the violation of assumption $A_k W < 1$ in the case of no competition.

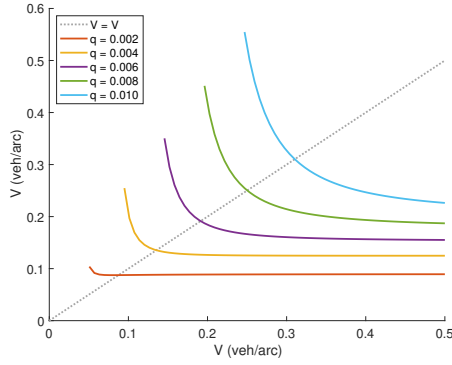


Figure 3: V^* as a fixed point.

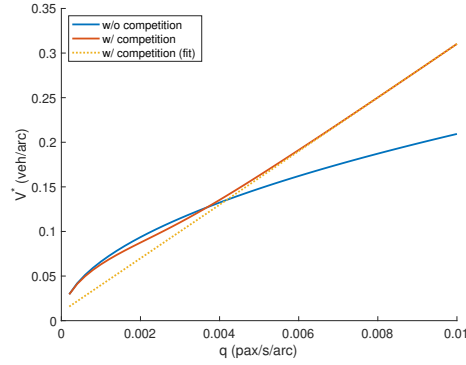


Figure 4: V^* in e-hailing.

A simple linear approximation is also plotted in Fig. (4) and fits well when q is relatively large (> 0.004 pax/s/arc) with an intercept close to zero (0.0098). Also, in theory, V^* should reduce to zero when $q = 0$ (when there is no demand, the optimal fleet size is also zero). Therefore, we propose to approximate V^* by a simple linear function

$$V^*(q) = \eta q, \quad (26)$$

and thus the unmatched passenger density is also a linear function of q , which reads

$$W^*(q) = \frac{\eta \delta_k}{2(\eta - \delta_k)} q. \quad (27)$$

Finally, the optimal system cost is derived as

$$c^* = c_0 \left\{ \left[\eta + \frac{\beta_m \delta_k}{c_0} \frac{\delta_k}{2(\eta - \delta_k)} \right] q + \frac{\delta}{\eta} \left(1 + \frac{\beta_p}{c_0} \right) \left[c_1 - \frac{c_2}{R_k} \sqrt{\frac{\eta \delta_k A_k q}{2(\eta - \delta_k)}} \right] \right\} + \left[\frac{\beta_m \delta_k}{2} + (c_0 + \beta_p) \left(d_a^k + \frac{1}{2} \right) \delta + \beta_a w_a^k + (c_0 + \beta_\tau) \tau_k \right] q \quad (28)$$

Comparing Eqs. (17), (21) and (28), one can easily observe the optimal system cost consists of two parts: (i) cost related to matching that can be represented as the extra supply cost to sustain a certain demand rate, and (ii) cost independent of matching that is proportional to the demand rate.

The impact of walking reflects in the second part. On the one hand, it helps reduce pickup and in-vehicle times and thus saves these costs on both sides of the market. On the other hand, passengers endure an extra walking cost. As will be shown in the numerical results, when the walking distance is reasonable and the vehicle unit cost is high, the former effect is dominant. Besides, the benefit is more significant in street-hailing because walking also helps increase vehicle cruising efficiency.

The more intriguing findings regard the impact of passenger competition. The marginal costs in the three scenarios discussed above are given by

$$\text{Street-hailing w/o competition: } mc_{(s,w/o)}^* = \frac{B_{(s,w/o)}}{\sqrt{q}} + C_{(s,w/o)}, \quad (29)$$

$$\text{E-hailing w/o competition: } mc_{(e,w/o)}^* = \frac{B_{(e,w/o)}}{\sqrt{q}} + C_{(e,w/o)}, \quad (30)$$

$$\text{E-hailing w competition: } mc_{(e,w/)}^* = -\frac{B_{(e,w/)}}{\sqrt{q}} + C_{(e,w/)}, \quad (31)$$

where $B_k, C_k, k \in \{(s, w/o), (e, w/o), (e, w/)\}$ are constants determined by the exogenous variables. It thus concludes that both street-hailing and e-hailing exhibit increasing returns to scale when there is no passenger competition, whereas e-hailing leads to decreasing returns to scale when there exists passenger competition. Further, we can compare $B_{(s,w/o)}$ and $B_{(e,w/o)}$ to which service has a more considerable scale economies:

$$B_{(s,w/o)} = c_0 \sqrt{\frac{\beta_m \delta_k}{c_0 A_k}}, \quad (32)$$

$$B_{(e,w/o)} = c_0 \sqrt{\frac{\beta_m \delta_k}{c_0 A_k} + \left(1 + \frac{\beta_p}{c_0}\right) \left(c_1 - \frac{c_2}{R_k}\right) \delta}. \quad (33)$$

As per Eqs. (32) and (33), the two scalars are only different in the term $\left(1 + \frac{\beta_p}{c_0}\right) \left(c_1 - \frac{c_2}{R_k}\right) \delta$. Our numerical results suggest the parameters $c_1 = c_2 = 1/6$ and the matching radius in e-hailing is often large (e.g., $R_k = 14$ with a threshold pickup time of 4 min). Therefore, it is safe to conclude $B_{(s,w/o)} < B_{(e,w/o)}$ and thus the street-hailing enjoys higher economies of scale. This result also aligns with empirical evidence (e.g., Zhang et al., 2019; Frechette et al., 2019).

Numerical experiments

In this section, we compare the system performances in different scenarios under the optimal fleet size. The values of exogenous variables and approximation parameters are reported in Tab. 1.

Existence of passenger competition

Recall that e-hailing presents the opposite scale economies with and without passenger competition. Hence, we first examine whether passenger competition often exists in an e-hailing market. Fig. 5 plots the number of unmatched passengers within a matching area (i.e., $A_k W$) solved for DS and DE under the assumption of no passenger competition. Clearly, for most tested demand levels, the condition $A_k W < 1$ holds for DS . In contrast, the assumption is easily violated for DE due to its much larger matching area. In other words, the increased matching radius of e-hailing not only reduces the matching friction between passengers and vehicles but also induces considerable competition among passengers. This phenomenon has also been recognized in some previous work (e.g., Zhang et al., 2019), but unfortunately not yet been widely adopted in recent studies on e-hailing services. Due to this observation, in what follows, we only present results of the model for e-hailing with passenger competition.

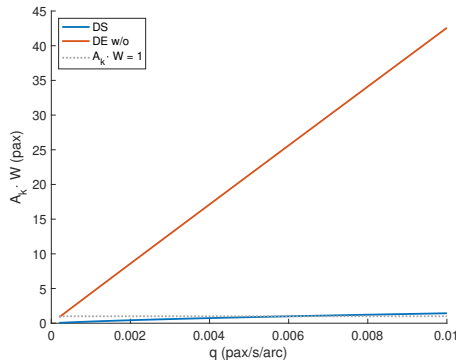


Figure 5: Violation of assumption on passenger competition.

System performance

Figs. 6 and 7 plot the composition of passenger waiting time and vehicle operation time at different demand levels. For e-hailing, the pickup time is deducted from the original walking time. This is because, in practice, passengers usually start walking after they are matched. Accordingly, the total waiting time is the matching time plus the maximum between the walking time and the pickup time. Because of the small matching radius in street-hailing, passengers spend most of their waiting time in matching while vehicles spend most of their vacant time (idle or pickup) in cruising. On the other hand, the pickup time takes a majority of passenger waiting time and vehicle vacant time in e-hailing, whereas its fraction reduces rapidly with the demand rate.

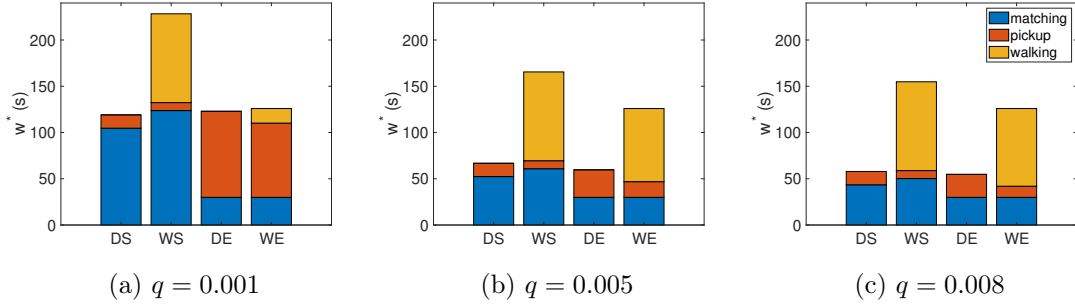


Figure 6: Passenger waiting time at different demand rates (pax/s/arc).

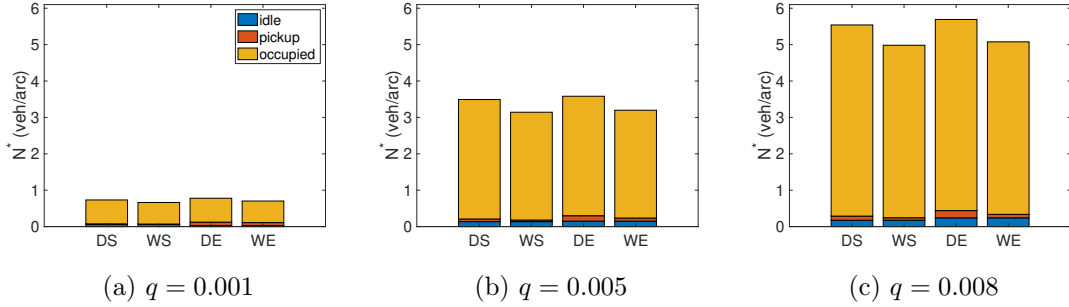


Figure 7: Vehicle time at different demand rates (pax/s/arc).

In the tested scenarios, walking does not really help passengers reduce their total waiting time because its benefit in improving the matching and pickup times is rather minor. Specifically, it does not help street-hailing passengers reduce their matching time as expected. A closer investigation reveals that this is mainly due to the decreased idle vehicle density. Although walking induces a higher concentration of vehicles on major streets, which yields a larger θ_k in Eq. (2), the optimal idle vehicle density becomes further lower, and thus the matching time increases.

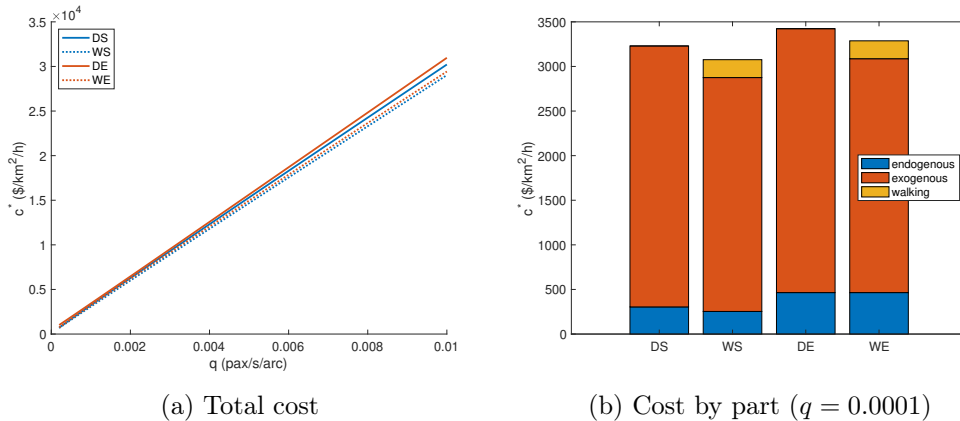


Figure 8: System cost under optimal fleet size.

Nevertheless, as shown in Fig. (7), walking does help reduce the required vehicle supply to sustain the same level of demand, both for street-hailing and e-hailing. Besides, it can be found that most of the vehicle time is occupied. This finding deviates from the empirical observations that the vehicle utilization rate is often lower than 60% (Schaller, 2017). We note that this discrepancy is largely due to the fleet sizing objective. In 11, we aim to minimize the system cost, whereas, in practice, the fleet size is often determined to maximize the operator’s profit or a consequence of drivers’ competition. In these cases, the fleet size is normally smaller than that at system optimum (Douglas, 1972).

The large contribution of occupied vehicle time also leads to a quite linear system cost illustrated in Fig. 8a. In brief, the four studied scenarios share very similar costs when the demand is relatively low. As the demand rate increases, street-hailing presents a better efficiency. Walking benefits both service types, while, as expected, it brings a larger cost saving to street-hailing, which is better illustrated in Fig. 8b. Nevertheless, all these differences are rather small compared to the total system cost.

Adaption of autonomous vehicles (AVs)

In face of the increasing labor cost, many e-hailing platforms are proposing to replace human-driven vehicles with autonomous vehicles (AVs), which are believed to have a lower operation cost meanwhile fully controllable. In what follows, we compare the system performance of *DE* and *WE* by only changing the vehicle unit cost c_0 to reflect the adaption of AVs. Here, the demand rate is set to be $q = 0.001$ pax/s/arc.

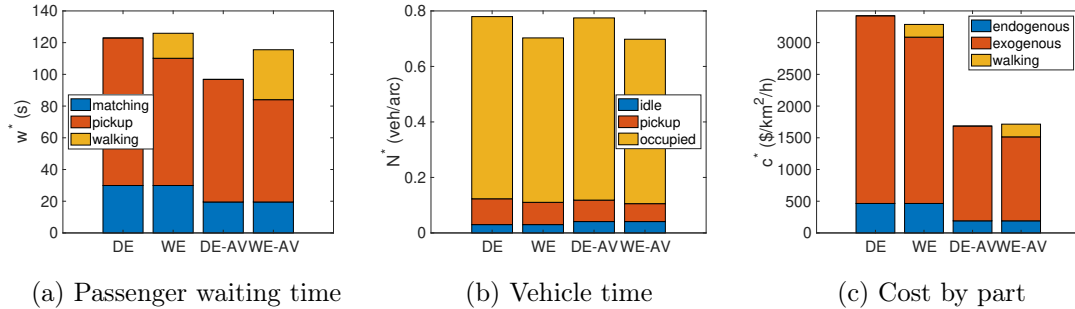


Figure 9: System performance with AVs.

As shown in Fig. 9a, the adaption of AVs reduces both the matching time and the pickup time because more idle vehicle time is devoted to cruising. Consequently, passengers enjoy a shorter waiting time in *DE* with an AV fleet. However, the benefit of walking diminishes, both for passengers (Fig. 9a) and for the system as a whole (Fig. 9c). On the supply side, the impact of AVs is rather minor. The lower cost of AVs does not induce a much larger supply size and the benefit of walking in terms of reducing the fleet size remains similar. Moreover, although the unit cost of AVs is one-fourth of human-driven vehicles, the system cost only cuts in half.

4 CONCLUSIONS

In this study, we model the street-hailing and e-hailing services on a grid network and analyze their scale economies with system optimum fleet size. We show the existence of passenger competition plays a critical role in the returns to scale. Without passenger competition, both street-hailing and e-hailing exhibit increasing returns to scale, while the scale effect in street-hailing is more significant. However, when subject to passenger competition, e-hailing shows decreasing returns to scale. Through numerical experiments, we show this is very likely to happen due to the large matching radius of e-hailing.

Although walking does not fundamentally change the scale economies, it produces two opposite impacts on the system cost. On the one hand, it reduces the pickup and in-vehicle times and specifically increases the matching efficiency in street-hailing. On the other hand, it imposes an extra cost on passengers. Our numerical results show that the cost-saving effect is in general more profound. However, when AVs are adapted with a much lower unit cost, the benefit of walking diminishes.

As a future direction, we will continue validating our findings with simulations on general road networks and demand profiles. It is also interesting to further analyze the scale economies with

different objectives in the fleet sizing problem (e.g., profit-maximization as in a monopoly market).

Table 1: Notations and default values

Variable	Description	Unit	Value
K	number of local streets between every two major streets		3
s	length of each street segment	m	120
v	vehicle travel speed on major street	km/h	25
α_l	speed scaling factor for local street		1.67
α_a	speed scaling factor for walking		5
R_s	matching radius in street-hailing		1
R_e	matching radius in e-hailing		14
A_s	matching area in street-hailing		4
A_e	matching area in e-hailing		420
δ_s	matching interval in street-hailing	s	20
δ_e	matching interval in e-hailing	s	20
δ	time to drive through one segment of major street	s	17.28
$\bar{\tau}$	average door-to-door trip duration	s	656.5
c_0	operation cost per human-driven (autonomous) vehicle	\$/h	20 (5)
β_m	value of time for matching	\$/h	15.00
β_p	value of time for pickup	\$/h	12.51
β_a	value of time for walking	\$/h	14.51
β_τ	value of time for in-vehicle time	\$/h	10.00
$\theta_{DS}(\theta_{WS})$	vehicle density correction factor in street-hailing		1.345 (1.94)
c_1	first approximation parameter for pickup distance		1/6
c_2	second approximation parameter for pickup distance		1/6
η	approximation parameter for the optimal idle human-driven (autonomous) vehicle density		30.05 (41.00)

REFERENCES

- Castillo, J. C., Knoepfle, D., & Weyl, G. (2017). Surge pricing solves the wild goose chase. In *Proceedings of the 2017 acm conference on economics and computation* (pp. 241–242).
- Cramer, J., & Krueger, A. B. (2016). Disruptive change in the taxi business: The case of Uber. *American Economic Review*, 106(5), 177–182.
- Douglas, G. W. (1972). Price regulation and optimal service standards: The taxicab industry. *Journal of Transport Economics and Policy*, 116–127.
- Fielbaum, A., Bai, X., & Alonso-Mora, J. (2021). On-demand ridesharing with optimized pick-up and drop-off walking locations. *Transportation Research Part C: Emerging Technologies*, 126, 103061.
- Frechette, G. R., Lizzeri, A., & Salz, T. (2019). Frictions in a competitive, regulated market: Evidence from taxis. *American Economic Review*, 109(8), 2954–92.
- Gurumurthy, K. M., & Kockelman, K. M. (2022). Dynamic ride-sharing impacts of greater trip demand and aggregation at stops in shared autonomous vehicle systems. *Transportation Research Part A: Policy and Practice*, 160, 114–125.
- Nie, Y. M. (2017). How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China. *Transportation Research Part C: Emerging Technologies*, 79, 242–256.
- Schaller, B. (2017). Empty seats, full streets: Fixing manhattan’s traffic problem. *Schaller Consulting*, 1(3), 1–27.

- Yang, H., Qin, X., Ke, J., & Ye, J. (2020). Optimizing matching time interval and matching radius in on-demand ride-sourcing markets. *Transportation Research Part B: Methodological*, 131, 84–105.
- Zhang, K., Chen, H., Yao, S., Xu, L., Ge, J., Liu, X., & Nie, M. (2019). An efficiency paradox of uberization. *Available at SSRN 3462912*.

Text-aided Group Decision-making Process Observation Method (x-GDP): A novel methodology for observing the joint decision-making process of travel choices

Giancarlos Parady*¹, Yuki Oyama², Makoto Chikaraishi³

¹ Lecturer, Department of Urban Engineering, the University of Tokyo, Japan

² Associate Professor, Department of Civil Engineering, Shibaura Institute of Technology, Japan

² Associate Professor, Graduate School of Advanced Science and Engineering, Hiroshima University, Japan

SHORT SUMMARY

Joint travel decisions remain poorly explained in behavioral models due to lack of empirical data. To address this problem, we propose a novel survey methodology to collect data on joint activities, from all members of a given clique. Through this method we are able to observe not only the outcome, but also the decision-making process itself, including the alternatives that compose the choice set, individual and clique characteristics that might affect the choice process, and the discussion behind the choice via texts. This will allow researchers to gain a deeper understanding of the joint decision-making process, including how alternatives are weighted, how members interact with each other, and how joint choices are made. Here we introduce the results of an implementation focusing on joint eating-out activities in Tokyo, focusing on survey components, execution, and insights on the data.

Keywords: Activity-based modeling, group behavior, joint decision-making, leisure travel behavior, social networks, survey methods

1. INTRODUCTION

Many of our behavioral decisions are made in coordination with members of the social networks we are embedded in. However, joint decision-making processes, particularly related to social activities, remain poorly explained in traditional behavioral models. A key reason for this is the lack of empirical data. While some studies have indeed focused on modeling joint activities, these studies rely on agent-based simulations (Arentze & Timmermans, 2008) and still require empirical data for parameter estimation and validation.

In recent years, egocentric network data-collection efforts have been conducted to get a better understanding of ego-centric social networks characteristics and social interactions such as surveys in Canada (Carrasco & Miller, 2006), Switzerland (Kowald & Axhausen, 2012), The Netherlands (van den Berg et al., 2012), Chile (Carrasco & Cid-Aguayo, 2012) and Japan (Parady et al., 2020) and the U.K. (Calastri et al., 2020). A key limitation of these efforts is that since data is collected using an ego-centric approach, the data that can be collected on other group members is limited to what ego can recall. This limitation is particularly critical for modeling travel behavior as spatio-temporal constraints are key constraints defining travel behavior (Hagerstrand, 1970). Han *et al.* (2023) has shown in the context of group eating-out destination choices that

considering the average travel times of all participating members of a clique increases the predictive ability of the model by up to 49% against a model considering only ego's travel times, a considerable increase in performance.

Against this background, this study proposes x-GDP (Text-aided Group Decision-making Process Observation Method), a novel survey method to collect data on joint activities and their underlying joint decision-making process of any dimensions of travel choice. We implemented the method for joint leisure activities with a focus on destination choice. Through this method we are able to observe not only the outcome but also the decision-making process itself, including the alternatives that compose the choice set, individual and clique characteristics that might affect the choice process, as well as the discussion behind the choice via texts. Observing such a process will allow us to first understand the decision-making process qualitatively, including how alternatives are weighted, how members interact with each other, and finally how the choice is made.

2. METHODOLOGY

The main objective of x-GDP method is to collect data on the joint decision-making process of travel-related activities of a given clique, a group where all members know each other. The general idea of x-GDP is to ask participant cliques to plan (and later actually execute) an actual activity or set of activities in the virtual presence of the researchers, using a chat-group interface. In this study we use as a case study eating-out activities due to its high frequency in joint activities (Stauffacher et al., 2005). Since participants have to actually conduct the activity decided in the group discussion, there are real incentives to guarantee a real discussion that considers the preferences and constraints of clique members. Fig 1 illustrates the flow of an x-GDP survey.

Step 1: Recruitment and pre-registration

x-GDP requires participation of existing cliques and registration of all members for schedule coordination. This study targeted cliques composed of at least one University of Tokyo student to simplify the sampling process. This was also done to limit to some extent the spatial distribution of participants to cliques with similar daily life activity spaces. Provided this condition was met, no constraints were imposed on the eligibility of other members.

Recruitment was done via social media (the Urban Transportation Research Unit Twitter account). In spite of the nonprobability sampling method, it is important to point out that the student population of the University of Tokyo is not that large (27,233 students) and is rather homogeneous in terms of sociodemographics. In total, data on 816 individuals in 217 cliques was collected. Out of the 816 participants 76% were University of Tokyo students, 20% students from other universities and 4% non-students.

Step 2: Virtual meeting schedule coordination

Scheduling coordination was conducted via online forms. As shown in Fig 2, the Schedule Coordinator matched Experiment Moderators (the person in charge of guiding the experiment over Zoom) with cliques. Once matched, all members were informed of the date and time and other details of the experiment such as conditions for payment of participation reward, etc.

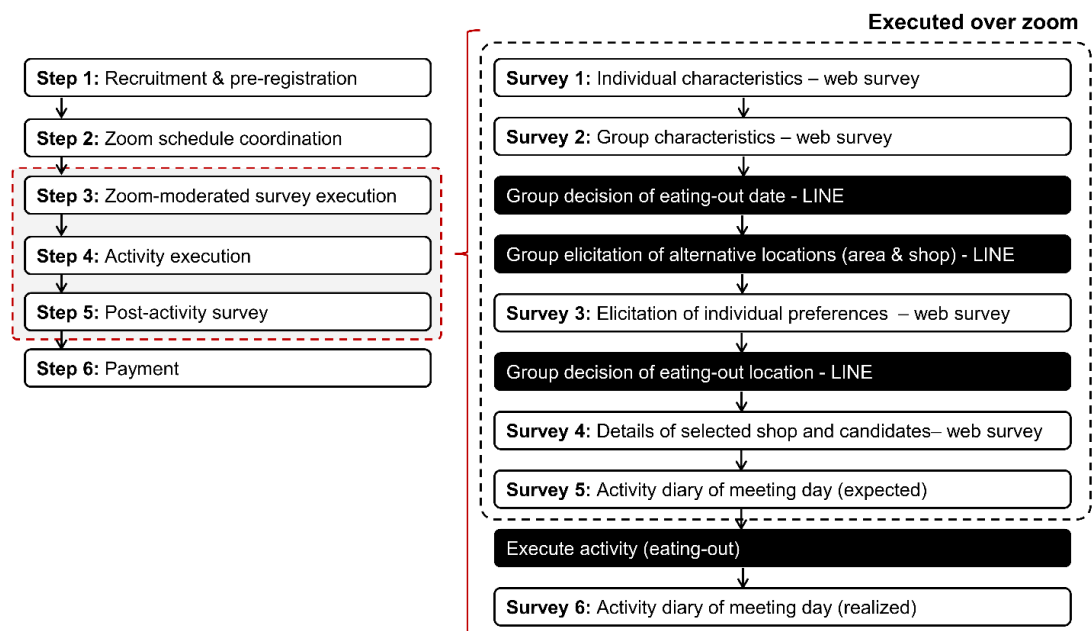


Fig 1. Flow of an x-GDP survey

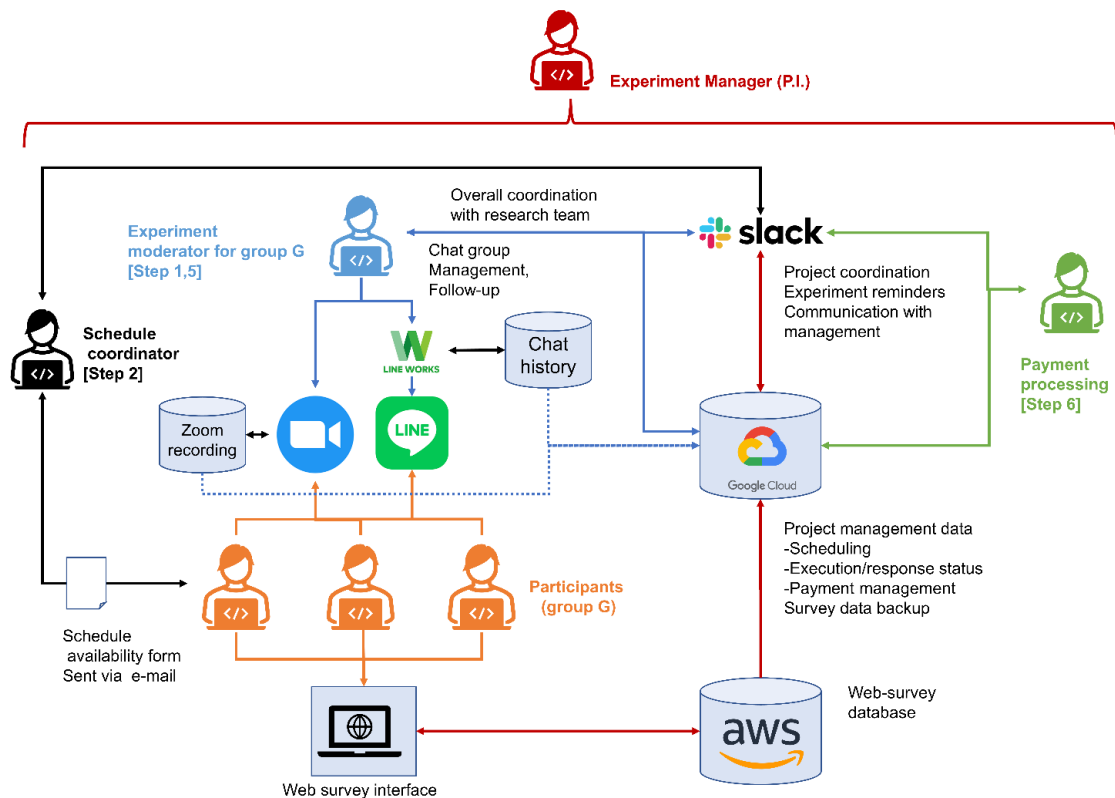


Fig 2. Simplified diagram of the logistics of the x-GDP experiment after recruitment.

Step 3: Zoom-moderated survey execution

This step was the crux of the experiment. Guided by the Experiment Moderator, participants were first asked to respond to Survey 1 and Survey 2 via an online survey platform. Survey 1 collected

data on individual socio-demographic characteristics. Survey 2 collected data on clique characteristics.

After Survey 2 was completed, the Experiment Moderator invited all members to a LINE group chat (LINE is the most popular instant communication app in Japan). The Experiment Moderator joined via a Line Works account (a cloud-based business chat tool that can link to LINE) for privacy, ethical and data management reasons. In the LINE group chat, the Experiment Moderator asked the clique to first decide the date and time of the activity. Two constraints were imposed. First, for management reasons, the date of the activity must be within a maximum of two weeks from the day of the experiment. Second, the activity must be done from 17:00 on to reduce the temporal variability of activities and simplify the modeling process later on. Note that these constraints can be generalized depending on the interests of the researcher.

Date and time defined, participants were asked to elicit potential areas and shops to execute the activity. There was no upper bound on how many candidates could be elicited but participants were asked to propose at least one location per person. Before moving on to the discussion phase to choose the activity location, respondents were asked to respond to Survey 3, which asked them to rank the elicited candidate locations in order of their personal preference. This was done anonymously so that responses were not affected by the opinions of others.

After completing Survey 3, participants were asked to discuss and decide the location of the eating-out activity. No guidance was given regarding how to make this decision, so each clique was free to choose their own method. No time constraint was imposed. The average duration for the LINE discussion section including time decision, preference elicitation and location decision was 35 minutes (S.D. 16.42 mins). The moderator then asked participants to respond to Survey 4 via a web-survey (at the clique level), to collect data on the chosen location as well other candidate locations. To avoid the issue of untraceable locations, participants were asked to use store links from either Tabelog (a restaurant review site in Japan) or Google maps. Out of the 1,188 unique shops elicited in the experiment, we were able to identify 99.5% of the shops via their public links and collect additional data on these shops.

Finally, once Survey 4 was completed, participants were asked to report their expected schedule for the day of the activity in the form of an activity diary (Survey 5) via a visual and interactive interface that greatly reduced the response burden.

Step 4: Activity execution

On the morning of the day of the planned activity, participants were sent a reminder via LINE and were given explanations about proof-of-execution submission such as location pin, a picture in front of the shop with a mobile phone showing date and time, A group picture inside the restaurant and the receipt.

Step 5: Post-activity survey

Using the same interface as Survey 5, data was collected on the actual schedule executed on the day of the activity.

Step 6: Payment

A monetary incentive of JPY 4000 (approx. USD 29.80) was provided for participants who responded to all surveys and provided proof-of-execution. For participants who did not provide proof-of-execution or did not complete Survey 6 after participation, the incentive was JPY 1080 (approx. US\$8).

3. RESULTS AND DISCUSSION

For brevity, we will not introduce the details of the individual and clique characteristics and focus explicitly on the scheduled joint activity and its decision-making process. Fig 3 illustrates both the chosen restaurant location as well as other considered candidates. The first thing to point out is the agglomeration of locations around Tokyo sub-centers such as Shibuya, Shinjuku and Tokyo connected via the Yamanote loop line, in addition to areas around the University of Tokyo's Komaba and Hongo Campuses. Historically, the Tokyo sub-centers have exhibited high degrees of agglomeration of commercial and other facilities due to their high levels of access both from the railway-connected suburbs as well as the city center. In addition, smaller agglomerations can be seen around the intersection of railway lines even though they are not central.

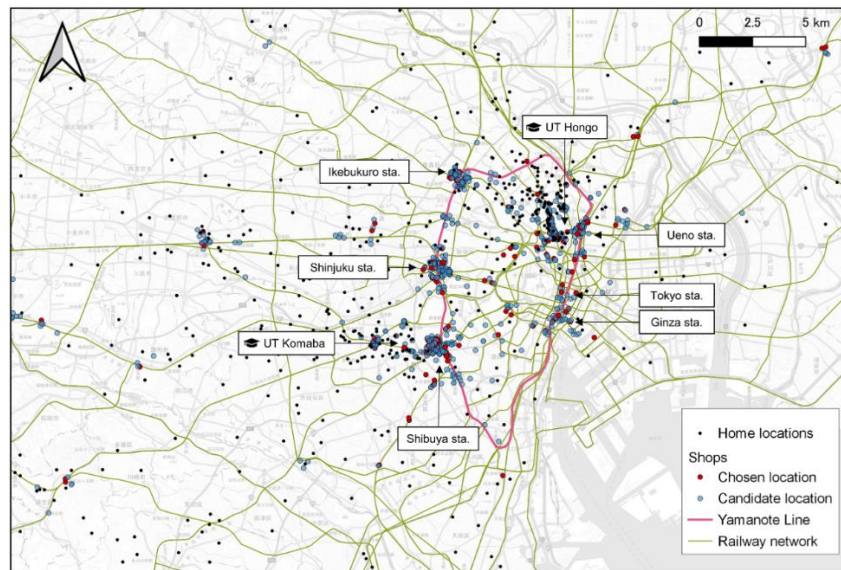


Fig 3. Location of chosen restaurants and alternatives considered during the experiment.

When asked the main reasons for choosing the locations each clique chose, restaurant quality and accessibility were the most frequently mentioned factors (78.8% and 57.1%, respectively). This is also consistent with the attitudinal responses collected in the individual survey (Survey 1) where respondents were asked to rate on a 7-point Likert scale (1 being not important at all, 7 being extremely important), the importance they place on different factors when eating out with a group (Fig 4). Group evaluation of shop and group transit access were rated six or seven by 71.2% and 76.7% of the individuals, respectively. What these answers do not capture is whose accessibility is being prioritized, or whose preferences. As shown in Table 1, in less than 12% of cases, all members' individually top-ranked locations were actually chosen, with this percentage reducing as clique size increases. Furthermore, irrespective of clique size in around 17% to 20% of cases, no one's top-ranked location was chosen by the clique, suggesting a considerable degree of compromise among members. This underscores the importance of observing the actual decision-making process to gain a better understanding of within-group dynamics.

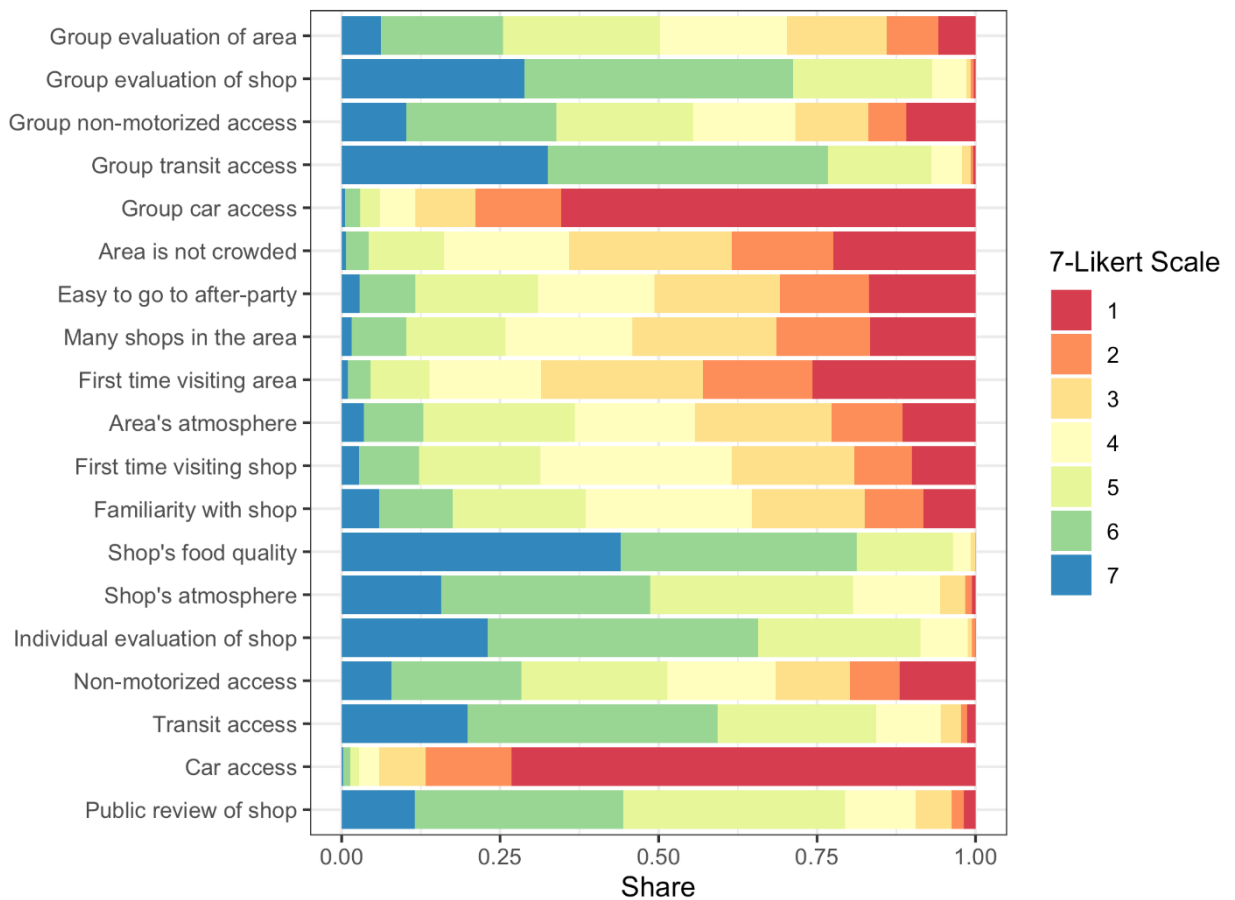


Fig 4. Factors considered important for group-level restaurant choice by individuals (n=816)

Table 1. Degree of matching between individually top-ranked locations and clique choice

		Number of individuals whose top-ranked locations are chosen by the clique					
		0	1	2	3	4	5
Clique size	3	18.6%	37.1%	33.0%	11.3%		
	4	14.7%	29.3%	28.0%	18.7%	9.3%	
	5	20.0%	37.8%	24.4%	11.1%	6.7%	0.0%

Two case studies

To further elucidate the properties of the data collected we will briefly introduce the decision-making process in two particular cases, as summarized in Fig 6 and Fig 7 using information from Surveys 1 to 5 as well as the LINE group discussion text record (a). The plots of members' schedules (b) and activity places (c) were created using data from the individual preference elicitation survey (Survey 3) and the expected activity diary of the meeting day (Survey 5).

The first clique (Fig 6) is composed of five same-year students. Two of the members had previous commitments on the suburbs of Tokyo on the day of the activity (1b and 1c). In this particular case several features of the decision-making process can be highlighted (1a). For instance, Mr. A pushed from early in the discussion for his preference, eating French food at Ginza, an upscale district in central Tokyo. Other members, like Mr. C, had a personal preference but

showed high degree of agreeableness and willingness to compromise for the group stating: “My preference is for meat, but if everyone is in for French at Ginza, I don’t mind.” While other alternatives were raised during the discussion such as Japanese BBQ or and oyster bar in Shibuya, Mr. A kept insisting on his preference by posting a link to the shop’s online site and menu: “Let me give you an idea of what French at Ginza will be like.” It should be noted that most members’ individually top-ranked locations were close to their expected origin locations on the day of the activity. Another constraint in the process was that some students were under 20 years old, hence could not drink alcohol, which tilted the choices towards restaurants rather than bars or Japanese izakaya. In the end the group agreed on Mr. A’s preference. In this particular case, Mr. A’s strong opinion clearly influenced the final decision, given the other member’s agreeableness and willingness to compromise. In other words, the weight of Mr. A’s opinion was larger than other members. At the same time, we can speculate that had other members had similarly strong opinions, the resulting outcome might have been different. Such information cannot be observed from the outcome alone, but we were able to capture it with the proposed x-GDP method.

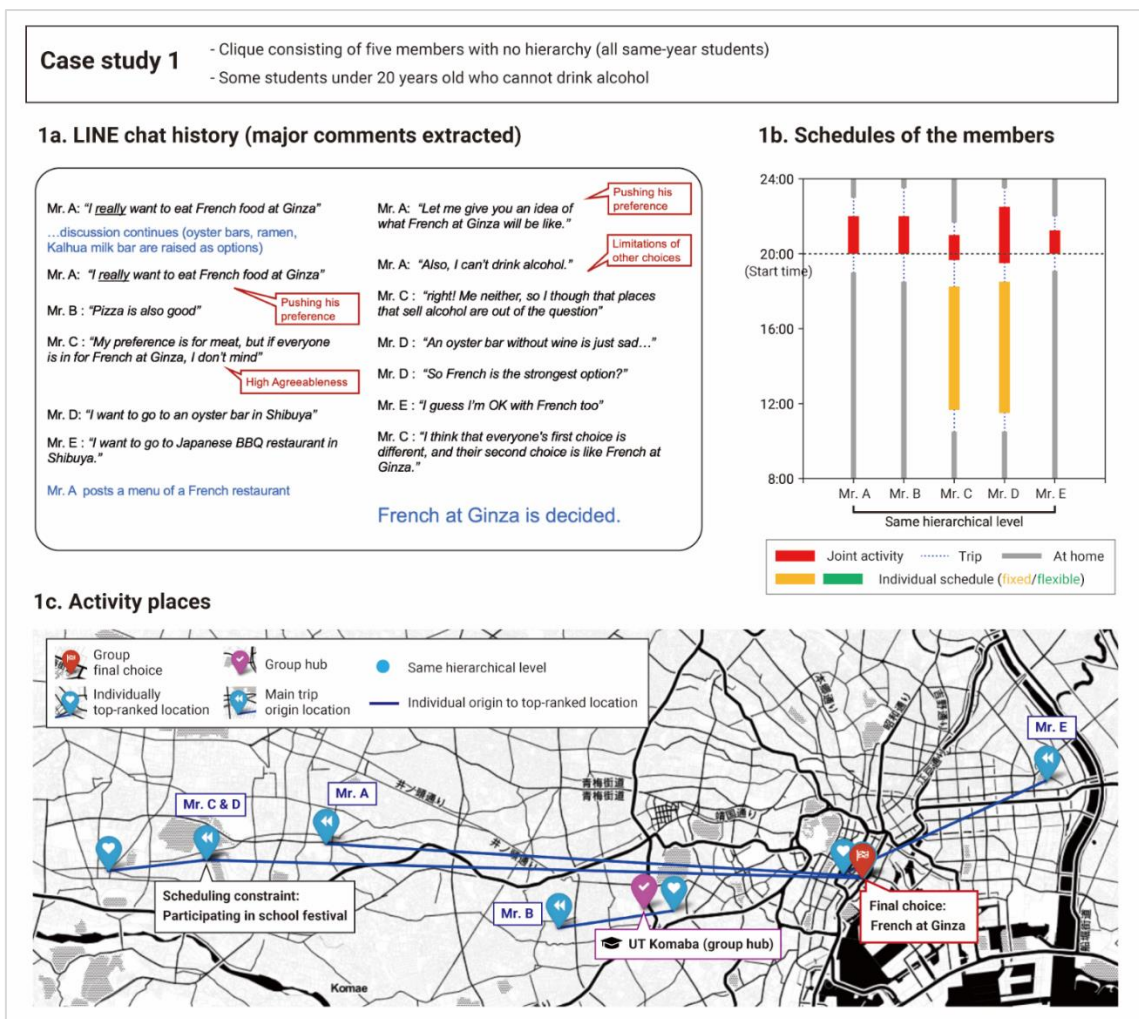


Fig 6. Extract of collected data for a clique 1. 1a. LINE chat excerpt. 1b. Schedule of members on activity day. 1c. OD lines to individually top-ranked location.

The second clique is composed of three futsal club friends, one of them being one year more senior than the other two. First, the joint activity time was set based on two time constraints. First, Mr. A had a part-time job at Shinjuku until 19:00 and second, all members wanted to watch the

FIFA World Cup (Qatar 2022) after dinner. Once the time slot was defined, several candidate locations were proposed, but they were all in Shinjuku. A possible reason for this is that Mr. A had a non-flexible activity schedule during that day. It is also worth noting that Mr. A was the more senior member of the group. The rest of the discussion focused on the restaurant type, such as hotpot, Brazilian BBQ and gibier. In this case, economic constraints were taken into consideration and Brazilian BBQ was selected.

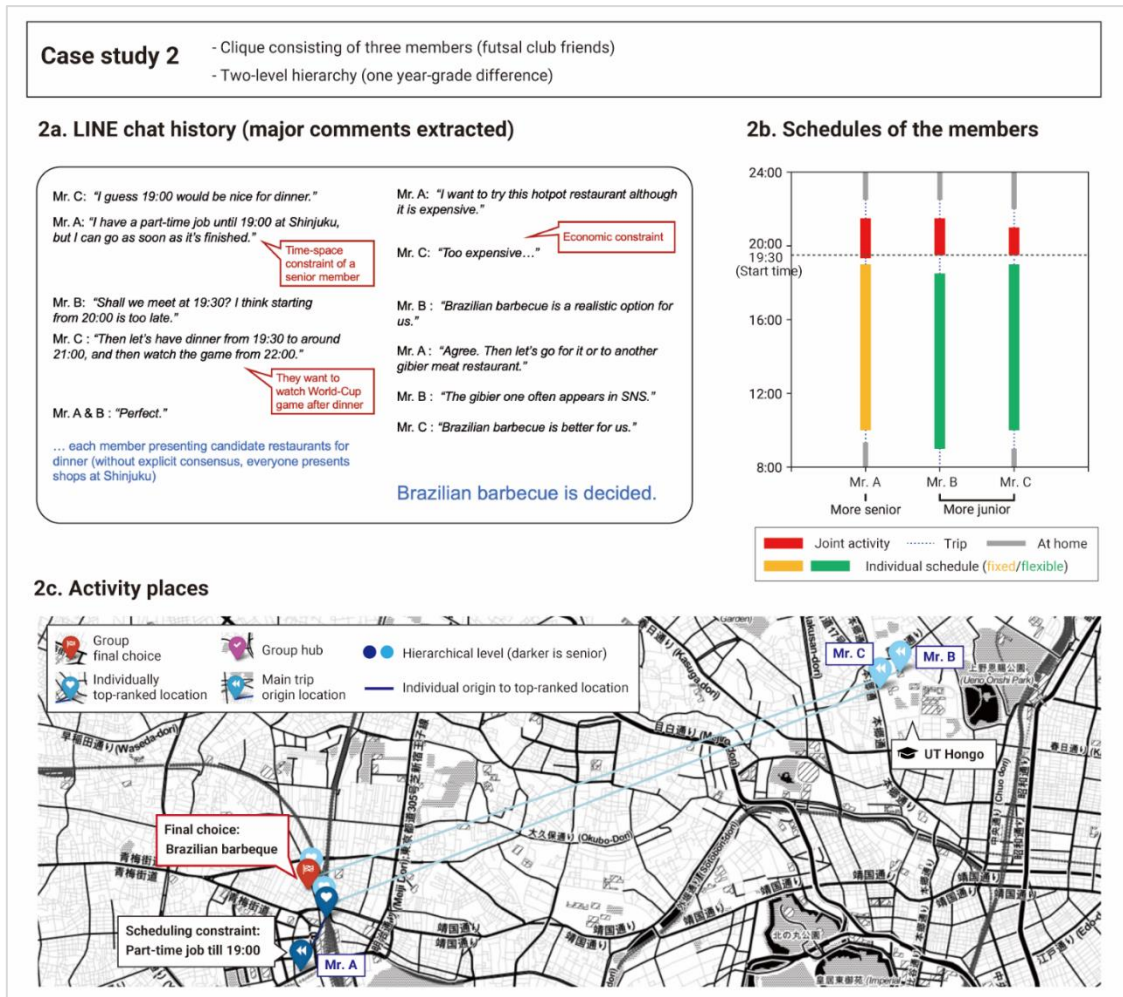


Fig 7. Extract of collected data for clique 2. 2a. LINE chat excerpt. 2b. Schedule of members on activity day. 2c. OD lines to individually top-ranked location.

4. CONCLUSIONS

To conclude we want to point out potential avenues of research that can be pursued with this kind of data. First, we have illustrated with only a few examples, that clique-level decision-making is rather heterogeneous. As such, a first necessary step is a qualitative analysis of the group discussion text records collected to formulate hypothesis regarding decision-making patterns. Such qualitative analysis can be complemented with quantitative methods such as natural language processing and cluster analysis.

Another potential avenue of research is the empirical estimation of joint accessibility and respective parameters. Theoretical joint accessibility methods have been proposed by Neutens *et al.*

(2008) however, empirical data is required to estimate model parameters. Joint accessibility estimates can be used to further investigate agglomeration effects in cities, as well as estimate joint activity destination choice models.

Finally, based on the above, we expect to build a theoretical framework to quantitatively model the joint decision-making process considering clique-level dynamics.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grants 20H02266.

REFERENCES

- Arentze, T., & Timmermans, H. (2008). Social networks, social interactions, and activity-travel behavior: A framework for microsimulation. *Environment and Planning B: Planning and Design*, 35(6), 1012–1027. <https://doi.org/10.1068/b3319t>
- Calastri, C., Crastes dit Sourd, R., & Hess, S. (2020). We want it all: experiences from a survey seeking to capture social network structures, lifetime events and short-term travel and activity planning. *Transportation*, 47(1), 175–201. <https://doi.org/10.1007/s11116-018-9858-7>
- Carrasco, J. A., & Cid-Aguayo, B. (2012). Network capital, social networks, and travel: An empirical illustration from Concepción, Chile. *Environment and Planning A*, 44(5), 1066–1084. <https://doi.org/10.1068/a43222>
- Carrasco, J. A., & Miller, E. J. (2006). Exploring the propensity to perform social activities: A social network approach. *Transportation*, 33(5), 463–480. <https://doi.org/10.1007/s11116-006-8074-z>
- Hagerstrand, T. (1970). What about people in the regional sciences? *Papers of the Regional Science Association*, 24.
- Han, C., Luo, L., Parady, G., Takami, K., & Chikaraishi, M. (2023). *Modeling joint eating-out destination choices incorporating group-level impedance : A case study of the Greater Tokyo Area* . <https://doi.org/10.48550/arXiv.2302.11400>
- Kowald, M., & Axhausen, K. W. (2012). Focusing on connected personal leisure networks: Selected results from a snowball sample. *Environment and Planning A*, 44(5), 1085–1100. <https://doi.org/10.1068/a43458>
- Neutens, T., Schwanen, T., Witlox, F., & Maeyer, P. De. (2008). My space or your space? Towards a measure of joint accessibility. *Computers, Environment and Urban Systems*, 32(5), 331–342. <https://doi.org/10.1016/j.compenvurbsys.2008.06.001>
- Parady, G., Takami, K., & Harata, N. (2020). Egocentric social networks and social interactions in the Greater Tokyo Area. *Transportation*. <https://doi.org/10.1007/s11116-020-10079-y>
- Stauffacher, M., Schlich, R., Axhausen, K. W., & Scholz, R. W. (2005). The diversity of travel behaviour: Motives and social interactions in leisure time activities. *Arbeitsberichte. Verkehrs.- Und Raumplanung*.
- van den Berg, P., Arentze, T., & Timmermans, H. (2012). A multilevel path analysis of contact frequency between social network members. *Journal of Geographical Systems*, 14(2), 125–141. <https://doi.org/10.1007/s10109-010-0138-0>

A fast algorithm to optimize meeting-point-based electric first-mile feeder services with capacitated charging stations

Tai-Yu Ma*¹, Yumeng Fang², Richard Connors³, Francesco Viti⁴, Haruko Nakao⁵

¹Research Scientist, Department of Urban development and Mobility, Luxembourg Institute of Socio-Economic Research, Luxembourg

²PhD candidate, Luxembourg Institute of Socio-Economic Research, Luxembourg

³Research Scientist, University of Luxembourg, Luxembourg

⁴Associate Professor, University of Luxembourg, Luxembourg

⁵Research associate, University of Luxembourg, Luxembourg

SHORT SUMMARY

This paper addresses the meeting-point-based electric demand-responsive-transport routing and charging scheduling problem under charging synchronization constraints. The problem considered exhibits the structure of the location-routing problem, which is more difficult to solve than conventional electric vehicle routing problems. We propose to model the problem using a mixed integer linear programming approach based on a layered graph structure. A two-stage simulated annealing-based algorithm is proposed to solve the problem efficiently. A mixture of randomness and greedy partial recharge scheduling strategy is proposed to find feasible charging schedules under the synchronization constraints. The algorithm is tested on 20 instances with up to 100 customers and 49 bus stops. The results show that the proposed algorithm outperforms the best solutions found by a commercial mixed-integer linear programming solver (with a 2-hour computational time limit imposed) for 12/20 test instances and with less than 1-minute computational time on average.

Keywords: charging synchronization, demand responsive transport, electric vehicle, feeder service, meeting point, metaheuristics (topics: electrification and decarbonization of transport, operations research application)

1. INTRODUCTION

Electric vehicle routing problems consist of deciding vehicle routes and charging schedules to serve a set of customers while satisfying constraints regarding vehicle capacity, time windows and vehicle energy (Kucukoglu et al., 2021). For passenger transportation, the problem is related to the electric dial-a-ride problem (Bongiovanni et al., 2019). Most of the literature assumes that vehicles can be recharged anytime with unlimited capacity of charging stations (Schneider et al., 2014). This assumption is often violated in practice as the number of rapid chargers is very limited due to their high installation costs. The electric vehicle routing problem with capacitated charging stations (EVRP-CS) is yet more difficult, as it needs to synchronize the charging operations of vehicles to save waiting time at charging stations. Recent research efforts have mainly focused on developing exact methods based on the mixed linear integer programming (MILP), by assuming that the vehicles recharge to full when arriving at charging stations and considering a linear charging speed (Bruglieri et al., 2019). To allow multiple visits of vehicles to the chargers, each charger node has several dummy copies. Charging capacity is ensured by deferring the current

visit of a charger to at least the full charging time of the previous visit of another vehicle in ascending order of visits. The authors propose a path-based MILP formulation and use the cutting plane method to solve exactly the test instance with less than 20 customers within 1-hour computational time. Froger et al. (2021) propose a path-based MILP formulation by considering piecewise linear charging functions and partial recharge, and propose a matheuristic method to solve the EVRP-CS exactly. Their solution method first generates a pool of initial routes without considering the capacity constraints of the charging stations, and then in the second step, they try to recombine these routes to find a solution satisfying the capacity constraints. Their problem assumes that the vehicles are homogeneous (battery size) and fully recharged before starting the service. They were able to solve exactly most of the test instances with 10 customers. Lam et al. (2022) propose a branch-and-cut-and-price algorithm to solve the EVRP problem with time windows and capacitated charging station constraints (EVRPTW-CS) by considering both vehicle partial recharge and piecewise linear charging functions. The charging scheduling synchronization subproblem is handled by applying the constraint programming technique. Their exact method can solve the problem with the larger test instances of up to 100 customers. To the best of our knowledge, the existing literature mainly focuses on exact methods that can be adopted to small-scale instances. There are still no efficient algorithms to address the related EVRP or electric dial-a-ride problem (e-DARP) at large scale.

In this paper, we aim to address the above issue and focus on a variant of EVRPTW-CS, an electric DRT (feeder) system which provides passenger transport service to connect to transit stations. This type of service is mainly applied in rural areas where public transport service is poor (Ma et al., 2021). To enhance efficiency of the electric DRT system, the meeting point concept (customers may board/alight at pre-defined stops near to their origins/destinations) is considered (Czioska et al., 2019). The problem needs to decide jointly where to pick up customers and how to route vehicles under various constraints. The problem is more complicated due to interactions between customer-to-bus-stop assignment and the subsequent vehicle routing and charging synchronization.

2. METHODOLOGY

2.1. Problem description

We consider a DRT feeder service in a rural area provided by an operator using a heterogeneous (in terms of capacity, battery size, and energy consumption rate) fleet of electric buses (also called vehicles hereafter) to complement the public transport system. To enhance system efficiency and reduce operational costs, the DRT system adopts the concept of **meeting points** i.e. customers are offered a limited number of pick-up/drop-off meeting points, rather than a door-to-door service (Czioska et al., 2019; Ma et al., 2021) and the service is **punctuated** (e.g. the vehicle arrives at a transit station every 10-20 minutes to drop off the transit passengers). The system is operated as follows. For a given planning period, customers submit their ride requests in advance indicating their origin, the transit station to be dropped off, and their desired arrival time (corresponding to the pre-defined arrival timetable of the DRT buses). The operator collects these ride requests and communicates whether customers' ride requests are accepted, their pickup time, and suggested bus stop (meeting point). The operator's objective is to optimize vehicle routes so as to arrive at transit stations within a fixed buffer time (e.g. ≤ 10 minutes before the announced arrival timetable at transit stations). We assume that customers are willing to walk from their origins to the suggested meeting points, up to some maximum acceptable walking distance. The state of charge of the vehicles cannot fall below the reserve battery level throughout the route. Vehicles can be recharged only at operator-owned charging stations; each station has a limited number of

chargers. Charging operations cannot overlap at any charger, i.e., a vehicle is not allowed to wait at a charger/charging station.

The meeting-point-based electric feeder service problem with charging synchronization constraints (**MP-EFCS**) problem is formulated as a mixed-integer-linear programming (MILP) problem as an extension of the electric dial-a-ride problem (e-DARP)(Bongiovanni et al., 2019), but adopting the concept of meeting points, allowing customers to be rejected (with a high penalty costs) under vehicle charging synchronization constraints. Given a set of customer requests, the objective is to optimize vehicle routes to meet these requests while considering the trade-off between system costs and customer inconvenience. The objective function minimizes the weighted sum of total vehicle travel time and total vehicle charging time, customer’s total walking time, total vehicle waiting time at transit stations before the acceptable fixed buffer time, and the total penalty of unserved requests. The computational time for solving the MP-EFCS exactly needs to enumerate all possible customer-bus-stop assignments and then solve each corresponding e-DARP with charging synchronization constraints problem (**e-DARP-CS**) to find the global minimum. This is possible only for very small problem size. To solve it efficiently, we propose a layered (directed) graph model (Fang and Ma, 2022) according to the sorted arrival timetable at transit stations and prune infeasible arcs or layers to reduce the problem size (see an illustrative example in Figure 1).

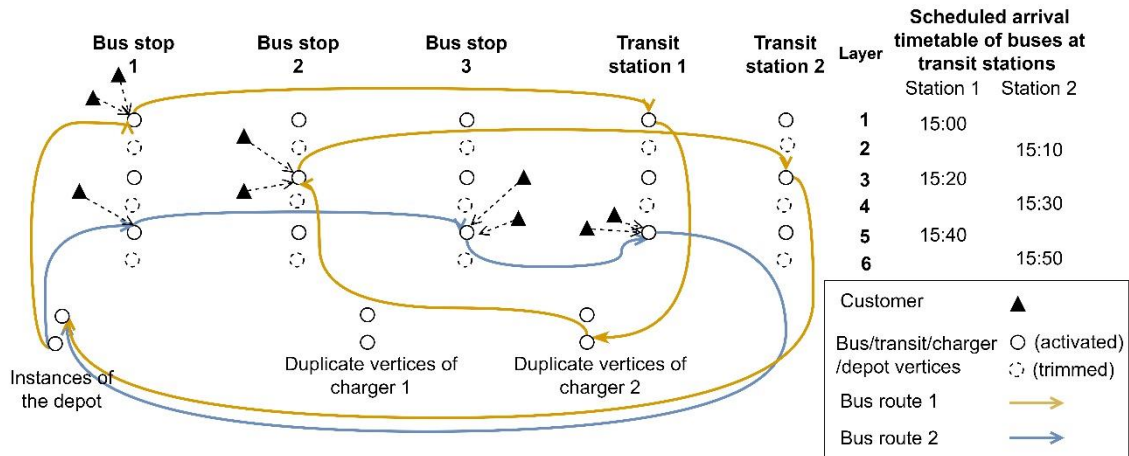


Figure 1. An illustrative example of the layered directed graph (arcs are omitted) for modeling the meeting-points-based electric feeder service with the charging synchronization constraints.

2.2. Solution algorithm

We propose an efficient **two-stage solution scheme** by finding a good customer-bus-stop assignment in the first stage. In the second stage, a **simulated annealing (SA) based metaheuristic** (Braekers et al., 2014) **with a post-optimization** procedure is proposed to solve the routing problem with charging synchronization constraints. The new challenge is how to optimize vehicles’ charging schedules with synchronization constraints. The **customer-bus-stop assignment** problem on the first stage is formulated as an MILP formulated, as a variant of the capacitated facility location problem, to minimize the weighted sum of the total customer walking time and bus travel time between the activated (with positive assigned customers) bus stops. Given the solution obtained from the customer-bus-stop assignment problem, we construct an e-DARP-CS instance by trimming off unused bus stop nodes and arcs connected to them, based on the layered graph model. An initial feasible solution is generated as the best feasible solution found for n random solutions using a greedy insertion approach. The SA-based algorithm applies a randomly selected local search operator on the current solution and obtains a temporary solution. If the cost of the

temporary solution is smaller than that of the current solution plus the threshold value T (temperature), and there are no charging operation conflicts, a vehicle exchange operator is applied on the temporary solution to further reduce the charging time of the vehicles of the temporary solution. If the resulting vehicle exchange and rescheduled charging operations (if any) has an improved cost without charging conflicts for all vehicles, then update the current solution. We track the number of times that the current best solution has stagnated. If this exceeds a pre-defined limit, the algorithm returns the current best solution. Otherwise, randomly selected local search operators are applied until a maximum number of iterations is achieved. This early stop criterion helps to reduce the computation time. As we allow customer requests to be rejected, unserved customers are managed in a pool, which is regarded as a virtual route, allowing customers to be removed from the vehicles. Note that charging schedule updating is applied at the end of each local search operation. Given the layered graph structure, we can efficiently screen-out infeasible insertion positions by checking whether the layer of a customer to be inserted is (in)compatible with the layer of the current inserted position of the route. This conflict check can be done in $O(1)$ and reduces the computational time significantly. We propose seven local search operators, including relocate ensemble, two-opt*, two-opt, exchange-segment, exchange-customer, four-opt, and create-route.

The e-DARP-CS instance is optimized based on the first stage customer-bus-stop assignment. It might be possible to accommodate unserved customers by changing their assigned bus stops, then re-inserting them into the current bus routes. In doing so, bus routes and charging schedules need to be updated accordingly. In the case that there are unserved customers, we propose an efficient post-optimization procedure to re-optimize the best solution obtained from the SA algorithm. Our numerical results show that this post-optimization procedure can improve the final solution and reduce the number of unserved customers with little additional computational effort.

3. RESULTS AND DISCUSSION

To test the algorithm, we consider two 4-hour scenarios, corresponding to peak (P) and off-peak (OP) demand profiles. Scenario P simulates a peak-hour situation where customers' desired arrival times at transit stations are concentrated around a peak hour, while OP reflects the opposite situation when customers' desired arrival times are uniformly spread over a longer operating period. In each scenario, we generate 10 instances spanning the range of 10-100 customers. These test instances have a single vehicle depot, two train stations, and four chargers. Meeting points (potential bus stops) are generated as a grid with a separation distance of 1 km and customers maximum walking distance is 1.5 km. Punctuated services are provided for the two train stations with three services per hour throughout the analysis period. In total, there are 26 layers with 25 to 49 activated bus stops per layer (bus stops within the maximum walking distance of the customers). A customer may have up to 7 potential bus stops within walking distance. Consequently, the possible customer-bus-stop assignment combinations are very large, providing non-trivial tests of algorithm performance. We consider two types of vehicles with different passenger capacity, battery capacity, and energy consumption rate.

The performance of the algorithm is compared with the solution obtained by a state-of-the-art MILP solver (Gurobi, version 9.1.2) with a 2-hour computational time limit. Our algorithm and the MILP model are both implemented using the Julia programming language. We run the experiments on a laptop with Intel(R) Core(TM) i7-11800H processor and 64 GB memory using a single thread.

MILP solutions obtained by Gurobi are reported in Table 1. The instance name c_{xx} means that there are xx customers in that instance. To ensure scenarios where vehicles need to recharge, initial battery levels of vehicles are set as low as 20%, 30%, ..., and 80% of the battery capacity.

For each instance, the MILP results give the best feasible solution found within 2 hours, along with the lower bound. The third column shows the number of unserved customers. The solver can obtain (near) optimal solutions for small instance of 10 customers. The number of unserved customers increases dramatically for instances with more than 50 customers. To account for the random elements within the two-stage SA-based algorithm, results are based on the average over 5 runs with random seeds. For each instance, we report the average objective function value and its gap to the best-known solution (BKS) found by the solver. The last two columns report the number of unserved customers and the average computational time (per run). The results show that the proposed algorithm outperforms the BKS for 12/20 test instances with less than 1-minute computational time on average.

Table 1: Computational results obtained using Gurobi solver and the two-stage SA-based algorithm on the test instances.

In- stances	MILP			Two-stage SA based algorithm			
	Best known solution	Gap to the lower bound	Num. of un- served custom- ers	Avg. obj. value	Gap to BKS	Num. of un- served cus- tom- ers	cpu time (sec.)
Scenario Off-Peak (OP)							
c10	107.91	1.10%	0	107.91	0.00%	0	4
c20	233.02	14.40%	0	234.96	0.83%	0	11
c30	327.46	16.30%	0	340.08	3.85%	0	36
c40	450.17	31.30%	1	424.03	-5.81%	0	39
c50	696.13	44.70%	2	611.32	-12.18%	0.4	36
c60	816.88	43.04%	4	659.31	-19.29%	0	120
c70	755.46	26.35%	0	782.41	3.57%	0	84
c80	1107.52	43.35%	7	906.83	-18.12%	0	123
c90	1525.81	58.33%	15	969.28	-36.47%	0	155
c100	1689.13	55.57%	20	1074.65	-36.38%	0	154
Scenario Peak (P)							
c10	112.31	18.00%	0	112.56	0.22%	0	5
c20	284.79	41.49%	1	311.36	9.33%	0	9
c30	340.85	33.16%	1	365.03	7.09%	1	22
c40	455.81	44.10%	2	472.6	3.68%	1	27
c50	705.36	55.10%	5	698.18	-1.02%	2.2	36
c60	996.41	60.11%	11	692.2	-30.53%	0	37
c70	1023.21	55.22%	15	769.55	-24.79%	0	60
c80	1514.25	66.05%	24	892.82	-41.04%	0	42
c90	1600.89	68.04%	24	1009.5	-36.94%	0	3
c100	1369.96	53.34%	13	1101.09	-19.63%	0	15

4. CONCLUSIONS

Electric vehicle routing with charging synchronization (under capacitated charging stations) are more difficult to solve and efficient solution algorithms are still underdeveloped for solving medium/large problem instances. In this study, we consider the problem of an electric dial-a-ride feeder system with charging synchronization based on the concept of meeting points and propose a layered graph model and a mixture of randomization and greedy strategy within a two-stage SA-based algorithm framework to solve this problem efficiently. We test the algorithm on 20 test instances with up to 100 customers and 49 bus stops. Results show that the proposed algorithm can find solutions efficiently with good solution quality. Several research directions are ongoing, including algorithmic parameter calibration, sensitivity analysis, charging infrastructure and fleet size planning, and integrated DRT system operational policy optimization.

ACKNOWLEDGEMENTS

The work was supported by the Luxembourg National Research Fund (C20/SC/14703944).

REFERENCES

1. Bongiovanni, C., Kaspi, M., & Geroliminis, N. (2019). The electric autonomous dial-a-ride problem. *Transportation Research Part B: Methodological*, 122, 436–456. <https://doi.org/10.1016/j.trb.2019.03.004>
2. Braekers, K., Caris, A., & Janssens, G. K. (2014). Exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. *Transportation Research Part B: Methodological*, 67, 166–186. <https://doi.org/10.1016/j.trb.2014.05.007>
3. Bruglieri, M., Mancini, S., & Pisacane, O. (2019). The green vehicle routing problem with capacitated alternative fuel stations. *Computers and Operations Research*, 112. <https://doi.org/10.1016/j.cor.2019.07.017>
4. Czioska, P., Kutadinata, R., Trifunović, A., Winter, S., Sester, M., & Friedrich, B. (2019). Real-world meeting points for shared demand-responsive transportation systems. In *Public Transport (Vol. 11, Issue 2)*. Springer Berlin Heidelberg. <https://doi.org/10.1007/s12469-019-00207-y>
5. Fang, Y., Ma, T-Y. (2022). Demand responsive feeder bus service using electric vehicles with timetabled transit coordination. In: *Proceedings of the 6th Conference on Sustainable Urban Mobility*, Skiathos, Greece (to appear).
6. Froger, A., Jabali, O., Mendoza, J. E., & Laporte, G. (2021). The electric vehicle routing problem with capacitated charging stations. *Transportation Science*. <https://hal.archives-ouvertes.fr/hal-02386167>
7. Kucukoglu, I., Dewil, R., & Cattrysse, D. (2021). The electric vehicle routing problem and its variations: A literature review. *Computers and Industrial Engineering*, 161(July), 107650. <https://doi.org/10.1016/j.cie.2021.107650>
8. Lam, E., Desaulniers, G., & Stuckey, P. J. (2022). Branch-and-cut-and-price for the Electric Vehicle Routing Problem with Time Windows, Piecewise-Linear Recharging and Capacitated Recharging Stations. *Computers & Operations Research*, 145, 105870. <https://doi.org/10.1016/J.COR.2022.105870>
9. Ma, T. Y., Chow, J. Y. J., Klein, S., & Ma, Z. (2021). A user-operator assignment game with heterogeneous user groups for empirical evaluation of a microtransit service in Luxembourg.

Transportmetrica A: Transport Science, 17(4), 946–973.
<https://doi.org/10.1080/23249935.2020.1820625>

10. Schneider, M., Stenger, A., & Goeke, D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science*, 48(4), 500–520.
<https://doi.org/10.1287/trsc.2013.0490>

Bayesian Optimization of Road Pricing using Agent-based Mobility Simulation

Dimitrios Argyros^{*1}, Renming Liu¹, Ravi Seshadri¹, Filipe Rodrigues¹, and Carlos Lima Azevedo¹

¹Department of Technology, Management and Economics, Technical University of Denmark, Denmark

SHORT SUMMARY

Road pricing policies are frequently debated but not widely adopted. Tools for designing near practice-ready policies are still missing, especially considering the complex dynamics between the different levels of traveller decision-making and the networks' performance. We couple an agent- and activity-driven mobility simulator with a Bayesian Optimization (BO) framework for designing optimal road pricing policy in a daily mobility and transportation network system. We extend the literature with a BO-framework application to distance-based road pricing under a departure-time and route-choice sensitive demand model combined with a detailed mesoscopic network. We then tested a general BO and a recently proposed contextual BO algorithm for SimMobility and computational performance. Both identified a similar optimum distance-based pricing, with the second being more computationally efficient. Nonetheless, iterations number, increasing search space and dimensionality could limit their performance. Lastly, the effects of the identified policy were analyzed by leveraging the outcome capabilities of SimMobility.

Keywords: agent-based modelling, Bayesian optimization, congestion pricing, machine-learning, SimMobility

1 INTRODUCTION

Numerous strategies have been explored to address traffic congestion and its associated repercussions, which include increased air pollution, accidents, and negative impacts on the overall city's quality of life. Congestion pricing control constitutes one of the main traffic management measures that have been researched but not widely applied during the past decades (Wang et al., 2022; de Palma & Lindsey, 2011). However, certain implementations have produced positive results, such as the Stockholm (Börjesson et al., 2012), and the Milan case (Gibson & Carnovale, 2015). As initially stated by Pigou (1920), congestion pricing serves as a technique for internalising the cost of externalities caused by road users. Since then, various tolling schemes, have been researched and implemented over the years. Regarding the classification of tolls, this occurs according to numerous factors. These can be the kind of scheme (e.g., facility-based, area-based, or distance-based), the level to which charges fluctuate across periods, additional characteristics of toll variation, and equipment (de Palma & Lindsey, 2011).

The most typical price optimisation approaches are either with nonlinear programming or a bi-level optimisation setup, both of which entail a high degree of complexity (NP-hard). Moreover, because of the practical implications they face, the first-best solution is generally used as a benchmark in many studies (Verhoef, 2005; de Palma et al., 2005). As a result, second-best prices, have been most researched, with commonly used methods for calculating prices being heuristics, meta-heuristics, approximations, and trial and error (Verhoef, 2005; Ekstrtopm et al., 2009; Luo et al., 2019).

Bayesian optimisation (BO) is a comparatively recent pricing optimisation approach. Since the pricing problem is NP-hard and does not have simple analytical solutions, black-box optimisation approaches are often used, and the BO has been demonstrated to be a promising one. First, Zhong et al. (2021) presented a BO technique for solving a bi-level optimisation issue where the effects of road pricing on both flow times and land usage were investigated. Particularly, an integrated transport assignment and land usage model was employed in the study and was combined with an active learning algorithm that featured the multivariate-multi-objective BO method. In the same vein, Liu et al. (2021) investigated a BO approach combining a trip-based Macroscopic

Fundamental Diagram (MFD) model with distance-based area-based pricing schemes. The toll profile was created based on time-of-day pricing, with the dimensionality impact on the BO under question. As a result, it was numerically discovered that as the search space expands, performance decreases progressively.

Considering the complexity of the problem, and the lack of tools for designing near practise-ready policies, this study aims to develop approaches for optimal design and evaluation of congestion pricing policies. Following the available research and leveraging the current knowledge, the study brings emerging dynamic congestion pricing control formulations closer to practice, by developing and extending state-of-the-art simulation frameworks. Two frameworks were inspired by the working paper of Liu et al. (n.d.), where BO algorithms were tested on a MFD model for the morning commute problem. However, this study uses a different simulation model which provokes higher complexity and differentiation in concepts but provides the possibility for thorough evaluation and analysis of policies. The study also extends the optimal toll query to a full-day mobility problem with multiple trip purposes.

2 METHODOLOGY

We assume the decision maker needs to define a distance-based road pricing scheme for an entire network, which can be time-varying. Since we aim at finding the best toll for an average day, the network equilibrium process and the optimization process may interact with each other. In a simpler and theoretically sound mobility model, it was recently shown that adding context data to BO related to the equilibrium process can help in reaching optimal solutions faster (Liu et al. (n.d.)). We test again such a hypothesis in a more complex and closer to reality simulator. In the following paragraphs, we will describe (1) the overall simulation platform, SimMobility; (2) its learning (equilibrium) process; (3) the general BO approach; (4) the contextual BO; and (5) our specific combined framework for both the general and the contextual BO cases.

SimMobility ¹, a detailed agent-based mobility simulator, was used in the context of this study. Specifically, SimMobility constitutes a large-scale simulator that includes many mobility-sensitive behavioural models in a multi-scale dimension. It was built from three distinct sub-models: long-term (LT), mid-term (MT), and short-term (ST). The MT model is made up of three interconnected simulators: 1) the pre-day simulator, which calculates the individual daily activity schedule (DAS), 2) the within-day simulator, which simulates departure times and route choice behaviour including en-route behaviour, and 3) the supply mesoscopic simulator, which handles network characteristics and the supply for different modes of transport. The reader is referred to (Adnan et al., 2016; Lu et al., 2015) for more details on SimMobility’s models and implementation.

The MT provides two learning approaches, with the one used known as within-day learning. In more detail, an activity schedule is initially given and the default travel time is derived for each link. Then, and for each time a within-day simulation is conducted, the link travel times experienced by each vehicle update the database accordingly. The second, namely day-to-day learning, occurs when the agents’ transport mode and their relevant travel information are updated. By iterating the pre-day, a new DAS is produced, and used as input to the within-day one, which consequently allows a learning process (Lu et al., 2015).

As for the BO, its main components are a model of the objective function and an acquisition function. Specifically, it develops a surrogate for the objective and quantifies the uncertainty in that by utilising a Gaussian process (GP) regression, fitting the data, and then selecting the next sampling point based on the acquisition function (Frazier, 2018). GP regression is described by a prior mean function $\mu_0(x)$ and a covariance function $\kappa(x, x')$, where x denotes the input variables. For the simplification of the GP training, $\mu_0(\cdot) = 0$ was used (Rasmussen & Williams, 2005). As for the covariance function, the Marten Kernel, defined below, can be used, where $\Gamma(\cdot)$ is the Gamma function and H_ν is the modified Bessel function. Here $\nu = 5/2$ is used.

$$W_{(x)} \sim GP(\mu_0(x), \kappa(x, x')) \quad (1)$$

$$W_{(x_{n+1})} | W_{(x_n)} \sim \mathcal{N}(\mu(x_{n+1}), \sigma^2(x_{n+1})) \quad (2)$$

¹<https://github.com/smart-fm/simmobility-prod>

$$k(x, x') = \frac{2^{\nu-1}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}}{l} \|x - x'\| \right)^\nu H_\nu \left(\frac{\sqrt{2\nu}}{l} \|x - x'\| \right) \quad (3)$$

The acquisition function is used to determine the next point to assess, given the posterior mean function and the variance of the GP. The upper confidence bound (UCB) is a widely used one (Srinivas et al., 2012) where ρ is a hyperparameter that determines the ratio of exploration to exploitation, with a bigger value indicating higher exploration. Here $\rho = 2$ was chosen.

$$n_{(UCB)}(x; \rho) = -\mu(x) + \rho\sigma(x) \quad (4)$$

$$x_{i+1} = \underset{x}{\operatorname{argmax}} n_{(UCB)}(x) \quad (5)$$

Regarding the contextual BO, Krause & Ong (2011) presented that the BO can take into consideration different environments through the usage of contextual variables. Thus, a contextual Gaussian process (CGP) was proposed. Here, the contextual variable which is used for the CGP is the day d . In that setting, and as the product of two Merten kernels is still a Merten kernel, the composite kernel occurs accordingly.

$$\bar{k}(x, d), (x', d') = k(x, x') * k(d, d') \quad (6)$$

As for the evaluation of the system, the main key performance indicators used were the travel time index (TTI), the Social welfare (SW) and the consumer surplus (CS). For the computation, only the trips by car are considered. TTI is the ratio of average travel time to free flow travel time. It is a metric of congestion that can be used to assess overall network performance and, here, is calculated at the 5-minute level, where $TT_{observed_{n_i}}$ denotes the travel time observed, $TT_{freeflow_{n_i}}$ the travel time in free flow settings and $TtripLength_{n_i}$ the total trip length, for the traveller n in the time interval i . SW is used for assessing the overall performance of the system and is computed based on the total of observed individual travel disutilities, where D_n denotes the disutility, $TravelCost_n$ the travel cost and β_{Cost_n} the cost coefficient for the traveller n . The CS of the system is calculated below and represents the total travellers' expenses.

$$TTI_i = \frac{\sum_{n_i=1}^{N_i} \frac{(TT_{observed_{n_i}} * TtripLength_{n_i})}{TT_{freeflow_{n_i}}}}{\sum_{n_i=1}^{N_i} TtripLength_{n_i}} \quad (7)$$

$$SW = \sum_{n=1}^N \frac{(D_n - TravelCost_n * \beta_{Cost_n})}{\beta_{Cost_n}} \quad (8)$$

$$CS = \sum_{n=1}^N \frac{D_n}{\beta_{Cost_n}} \quad (9)$$

Regarding the frameworks, both extend the SimMobility simulator and interact with it by updating its database; thus, their integration happens. Their objective function is the SW, whereas the acquisition function predicts the (distance-based) toll's control variables. Noteworthy is that the tolling rates are imputed to the database at 5-minute intervals based on a multi-modal Gaussian distribution resulting from the BO algorithms, where the number of modes corresponds to the number of network traffic peaks (p). In this setting, the amplitude (A_p), represents the height/highest price, the mean (μ_p) indicates the time of the highest price, and the standard deviation (σ_p) defines the width of the profile.

As for framework 1 (fig. 1), the simulation starts with a sampled toll and when the system has reached equilibrium, the BO forecasts the next test points/toll. Then the database is updated, by initialising the link travel times at no-toll conditions and incorporating the new toll prices, and the within-day simulator is iterated until the equilibrium is reached. The process is repeated until the GP has converged by reaching the optimum toll profile.

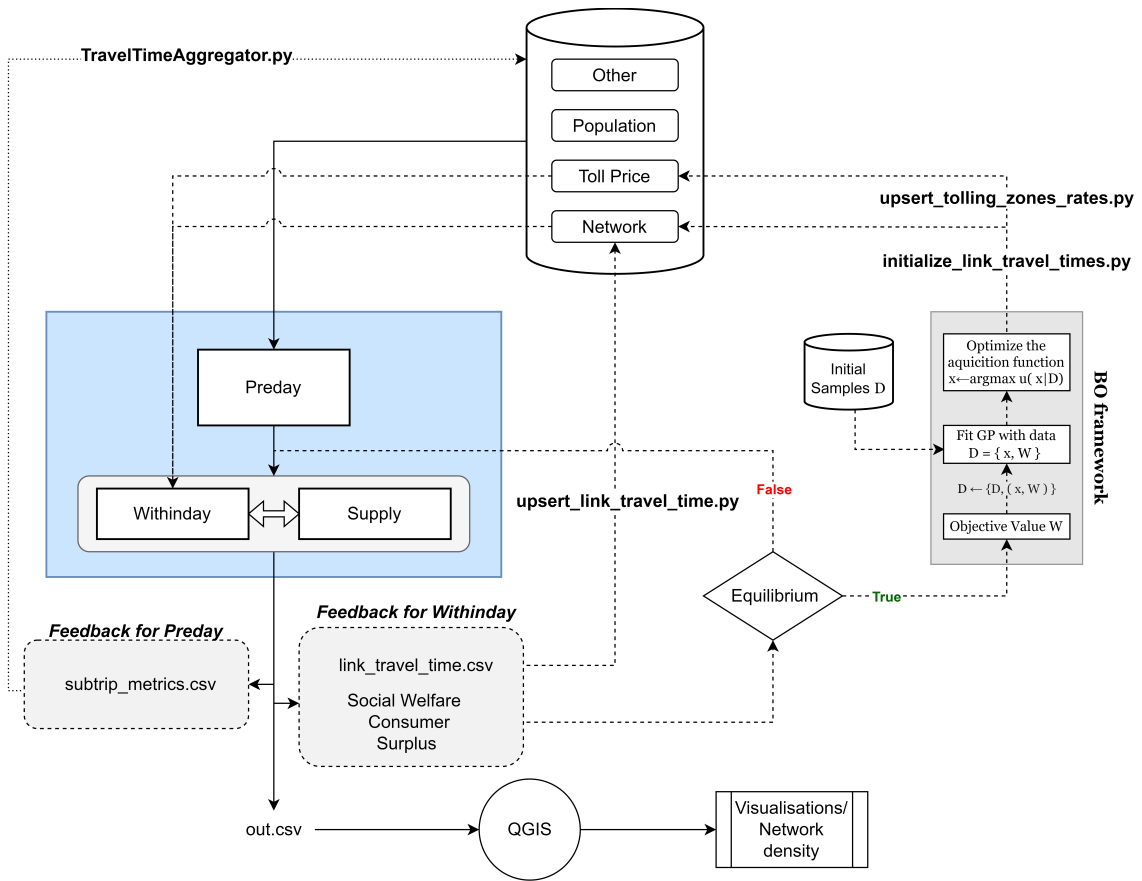


Figure 1: Framework 1. Bayesian optimisation and SimMobility

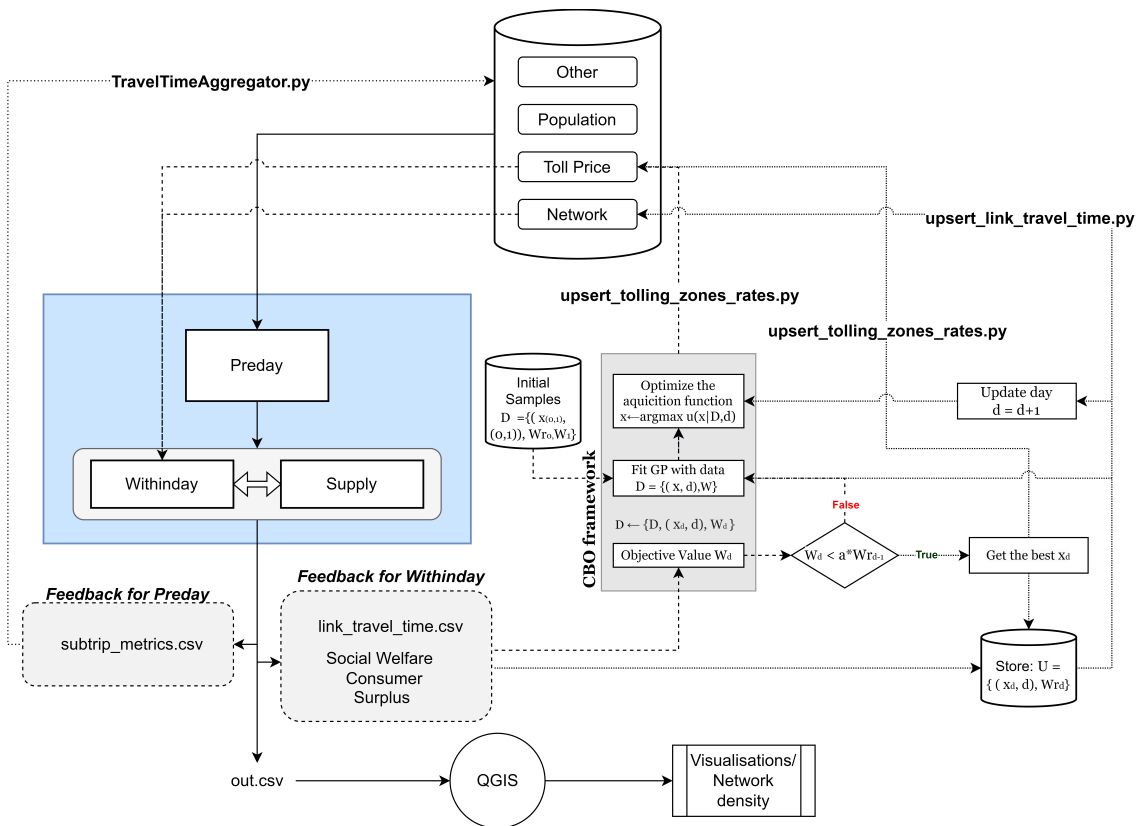


Figure 2: Framework 2. Contextual Bayesian optimisation and SimMobility

Framework 2 (fig. 2) follows a similar design but aims to offer greater flexibility as it can alter the day, based on the condition that the day’s SW is less than the preceding one multiplied by a scaling factor (a). In more detail, the framework considers as initial inputs the SW of day zero (SW of the no-toll case), and a few input points for day 1. Then the contextual BO proposes the next toll for the same day and SimMobility calculates its SW. The same procedure continues until a toll which gives a better SW than the previous day’s has been identified. Then the realised SW is calculated for the day and stored. Simultaneously, the framework updates the perceived link travel times of the network. Then, it updates the day and the algorithm proposes the first toll for the new day. This approach would be repeated until an overall equilibrium is attained for both the system and the GP.

Lastly, an initial set of points is essential for the initialization of the BO algorithms. Thus, a space-filling experiment design which uses Latin hypercube sampling (LHS) was included in both frameworks (Mckay et al., 2000).

3 RESULTS AND DISCUSSION

SimMobility’s built-in prototypical city was used for the experiments. Its network has 254 links, which correspond to 286 segments. The total length of the links is 279 km with an average maximum speed of 65.27 km/h and a standard deviation of 12.77 respectively. The number of nodes of the network is 95 with 1918 turning intersections. Lastly, the city can be separated into 24 zones. Also, a fixed demand of 19000 travellers, corresponding to 51071 trips, were employed. Moreover, to attain a congested condition of the system, the network’s initial capacity was reduced to 10%, despite a few bottlenecks, which remain at 30% of the original capacity. Based on the new capacity, the average critical density of the system per link is at 6,62 passenger-car units(PCU)/km with a standard deviation calculated at 2,03.

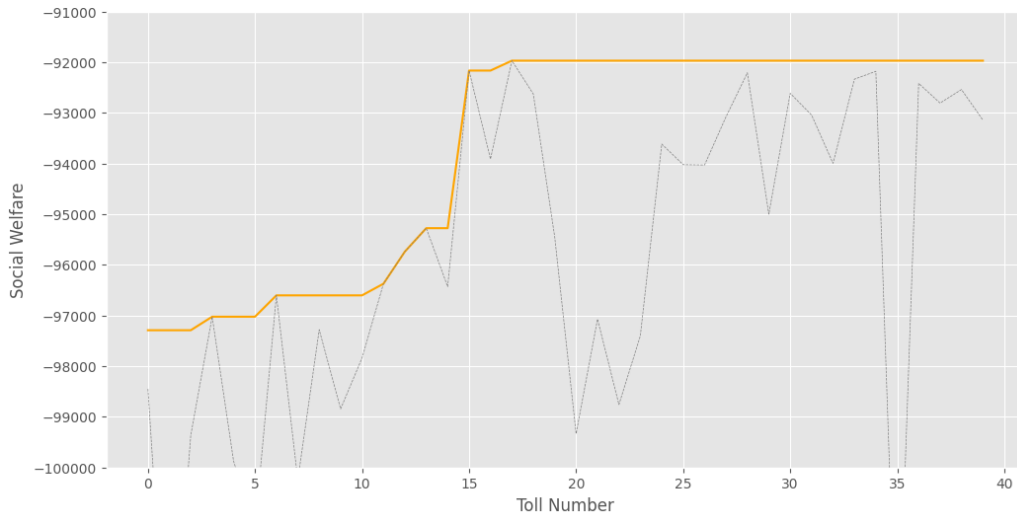
First, an experimental design took place to verify and evaluate the frameworks. The frameworks set to design a distance-based toll policy with an upper payment bound (UB) at \$10. The purpose of this experiment was to prove that the developed frameworks are adequate to design and assess a toll which optimises the system’s SW. On top of that, the comparison of the frameworks’ performance was set under question. Moreover, it should be noted that for the experiments, the pre-day simulation is inactive and considered that it has already provided the modal split and the respective DAS file.

Table 1: Optimisation bounds

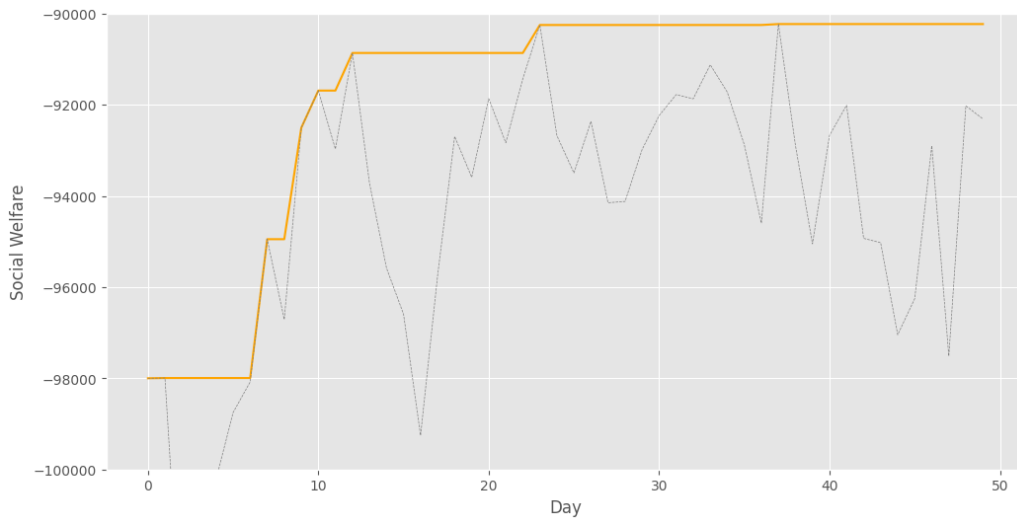
Bounds	μ_1 (5-min)	σ_1 (5-min)	A_1 (\$/m)	μ_2 (5-min)	σ_2 (5-min)	A_2 (\$/m)
Upper	108	24	0.003	234	24	0.003
Lower	84	3	0.00001	210	3	0.00001

It has been found that both frameworks can indeed design a toll that raises the SW of the systems. Also, the fact that both frameworks reached a similar toll profile is a clear sign that the global optimum, within the respective constraints, has been discovered. Regarding their performance (see fig. 3 and fig. 4), the number of iterations required for framework 1 was for the 12 sampled points 240 (~70 hours), and for the algorithm, until the optimum been found after 15 repetitions, was 300 (~88 hours). However, considering the number of needed iterations for the system to reach stationary, 126 and 150 iterations, respectively, are determined as a "fair" number. As for CBO, the setting immediately offers many advantages as just 5 iterations were required for initialization. The framework then ran with varying iterations per day, reaching the ideal toll region after ~110 iterations (~32 hours).

As for the designed toll, the one with the biggest SW provided by framework 2 (μ_1 : 108.0, σ_1 : 14.39, A_1 : 0.003, μ_2 : 234.00, σ_2 : 24.00, A_2 : 0.00001) was selected for further analysis. The toll has a very high peak, similar to the morning traffic peak, which has pushed some trips earlier in the morning. In this way, the vehicle accumulation and the TTI have decreased compared to the no-toll case, from ~1200 to ~900 and from 1,65 to 1,35 respectively (see fig. 5). As for the afternoon peak, CBO has not defined a toll profile that affects the peak’s characteristics, while the amplitude was set at the lower bound. Regarding SW, this was calculated to \$92611 implying a 5,5% improvement



(a) Framework 1



(b) Framework 2

Figure 3: Best social welfare (\$) (yellow line) and social welfare (\$) evolution for tested tolls (grey line)

Table 2: Base case and optimum toll case metrics

	Base case		Toll case		MD
	Mean	SD	Mean	SD	%
Social welfare (\$)	-98006.82	2707.91	-92611.51	927.19	5.51
Consumer surplus (\$)	-98006.82	2707.91	-231406.13	963.59	-136.11
Trip avg. travel time (sec)	458.38	6.23	410.07	2.43	10.54
Trip avg. scheduled delay (sec)	332.74	5.44	515.95	1.62	-55.06
Trip avg. paid toll (\$)	0.00	0.00	3.31	0.00	-

compared to the base case (see table 2). Similarly, the mean travel time improvement is around 10,5%, where the average scheduled delay has become 55% longer. Moreover, the most gains in travel time have been identified in the southeast part of the city, while the bigger CS losses were presented in the northern region. Furthermore, the link densities during the morning peak hour (08:00–09:00 AM) have been decreased. Specifically, the number of links during the peak hour that presented a density bigger than the average critical density of the system, decreased from 64 to 39 due to the implementation of the toll. Similarly, the average system density in this period declined from 2.3 to 1.9 PCU/km/lane. Additionally, for travellers, it has been calculated that in the equilibrium state, $\sim 20\%$ pays the maximum limit of \$10 and $\sim 55\%$ less than \$1 (see fig. 6).

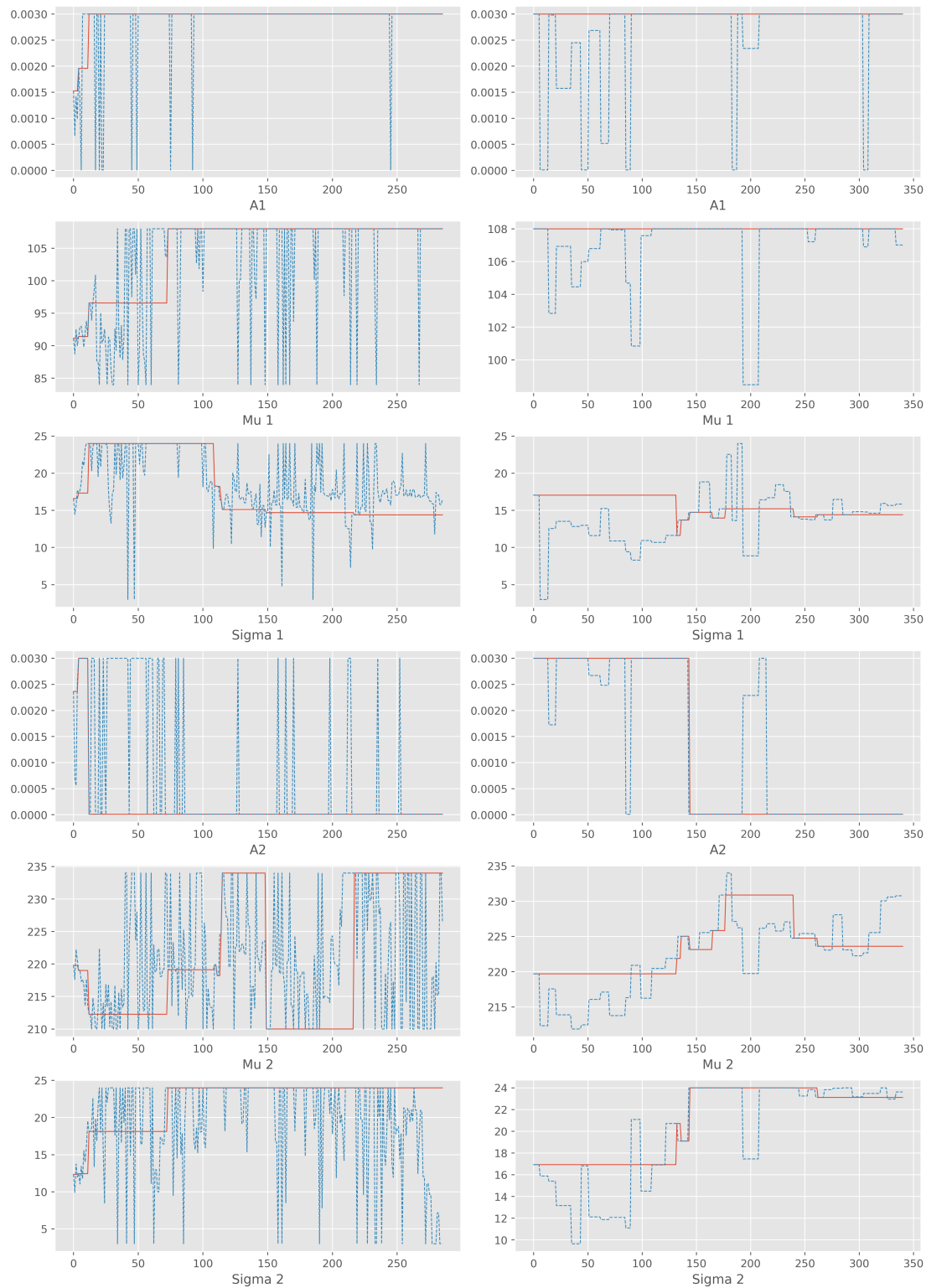


Figure 4: Selected parameters across all iterations for both frameworks. Note: No sampled data included. For framework 1 only the iterations until stationary are included. The red line indicates the parameter for best social welfare. The first column refers to framework 2 and the second to framework 1

Lastly, in another direction, a shortcoming of the toll was also identified, as travellers departing after the peak hour (10:00-11:00 AM) faced a high toll charge and added scheduled delay due to it. That could be mainly reasoned to the system's demand which is comparatively low during this period and thus the applied toll during this period does not affect significantly the SW. That consequently may prevent the frameworks to design a better solution.

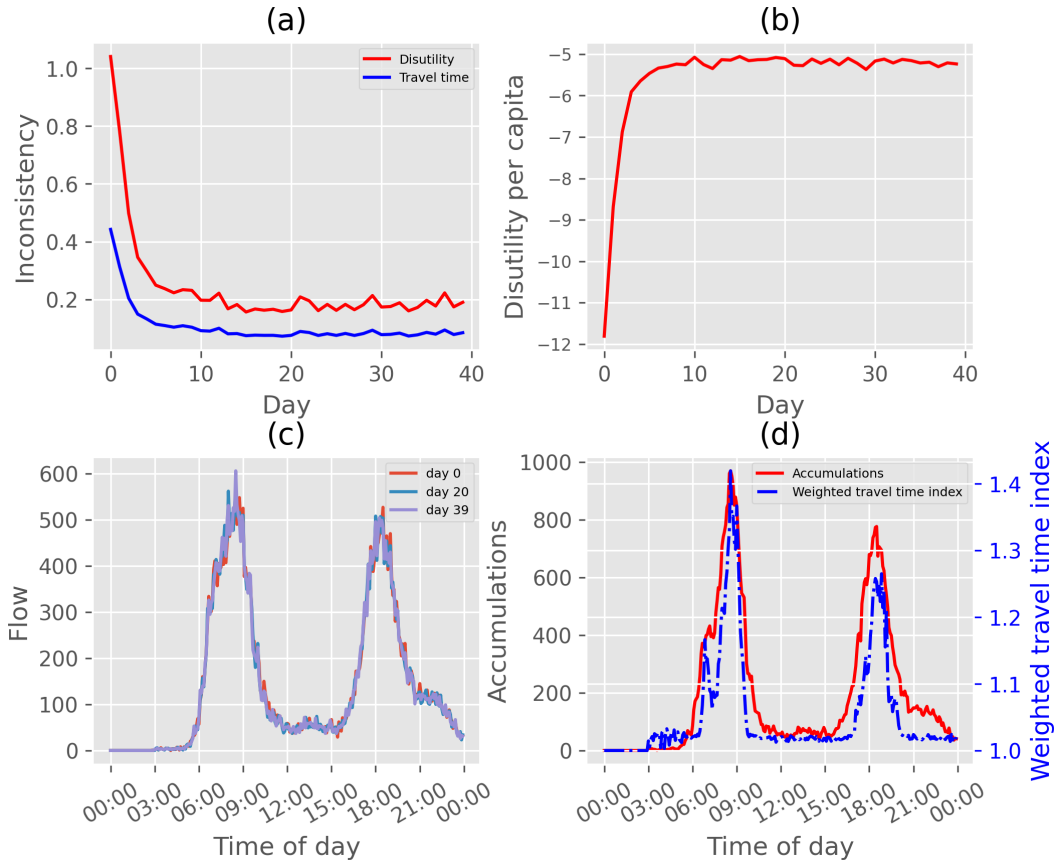


Figure 5: Optimum toll case. Simulation and congestion graphs.
 Note: The accumulation and TTI are based on the last iteration

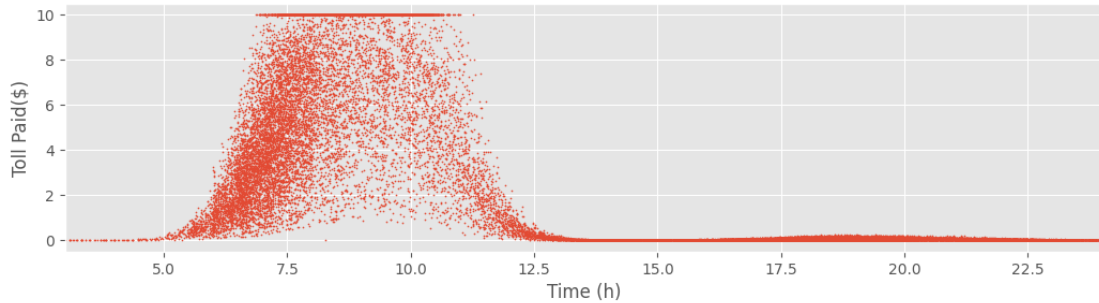


Figure 6: Optimum toll case. Amount of toll paid (\$) per user in equilibrium state
 Note: The graph is based on the last iteration

In other experiments, framework 2 attempted to determine whether there was a genuine distance-based toll. It was found that the amplitudes converged to zero since the framework could not find a toll that enhances the SW. Nonetheless, when the framework was asked to find a very low price that improves the SW, it was not confident of any solution after a sufficient number of iterations. This limitation may arise from the fact that the differences in the SW system shown for such a small toll could not be significantly different from the no-toll case. Also, considering the quite big system variance, with a standard deviation at 3%, this might be misleading. Similarly, framework 2 was set to find a better solution for the SW, while a larger search space in terms of boundaries and dimensions was set. Nonetheless, that was not possible in 30 days. The fact that in these 30 days, it ran only for ~ 90 iterations in total may be one of the causes of this failure. In addition to the number of iterations, the acquisition weight could also have an impact on the solution as the search space was expanded.

4 CONCLUSIONS

During this research, two frameworks combining a detailed agent-based simulator, namely Sim-Mobility, and BO algorithms, were developed and tested. Their main goal was to determine the optimum road pricing policy of a day by optimising the system's SW. It was found that both frameworks were able to pinpoint the optimum distance-based toll design with UB at \$10. Nevertheless, in terms of performance, the second framework proved that needs much fewer iterations.

The advantages of the analysis of a road pricing policy through a detailed agent-based simulation were also showcased. First, the policy was able to improve the SW of the system as well as the TTI during the peak period. Also, by diving into more disaggregated metrics, the identification of improvements in link-level density, during the peak period, was possible. In the same vein, by using a zonal level analysis, the regions with bigger travel time gains and CS losses per traveller, which were the city's southeast and northern parts respectively, were identified.

Moreover, framework 2 found that there is no genuine distance-based toll that improves the SW in this environment setup. However, the framework's inability to identify a solution, when the amplitude's search space significantly decreased, was identified. This may be due to the number of iterations combined with the system's inherent variability. Finally, framework 2 could not locate the same solution or discover a better toll profile when faced with an expansion of the size and dimensions of the search space. Thus, more iterations or a higher acquisition weight may be necessary to resolve this issue.

REFERENCES

- Adnan, M., Pereira, F., Azevedo, C. L., Basak, K., Lovric, M., Raveau, S., ... Ben-Akiva, M. (2016). Simmobility: A multi-scale integrated agent-based simulation platform..
- Börjesson, M., Eliasson, J., Hugosson, M. B., & Brundell-Freij, K. (2012, 3). The stockholm congestion charges-5 years on. effects, acceptability and lessons learnt. *Transport Policy*, *20*, 1-12. doi: 10.1016/j.tranpol.2011.11.001
- de Palma, A., Kilani, M., & Lindsey, R. (2005, 8). Congestion pricing on a road network: A study using the dynamic equilibrium simulator metropolis. *Transportation Research Part A: Policy and Practice*, *39*, 588-611. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0965856405000388> doi: 10.1016/j.tra.2005.02.018
- de Palma, A., & Lindsey, R. (2011, 12). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, *19*, 1377-1399. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0968090X11000362> doi: 10.1016/j.trc.2011.02.010
- Ekstrtopm, J., Engelson, L., & Rydergren, C. (2009). Heuristic algorithms for a second-best congestion pricing problem. *NETNOMICS: Economic Research and Electronic Networking*, *10*, 85-102. doi: 10.1007/s11066-008-9019-9
- Frazier, P. I. (2018, 7). A tutorial on bayesian optimization. Retrieved from <http://arxiv.org/abs/1807.02811>
- Gibson, M., & Carnovale, M. (2015, 9). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics*, *89*, 62-73. doi: 10.1016/j.jue.2015.06.005
- Krause, A., & Ong, C. S. (2011). *Contextual gaussian process bandit optimization*. Retrieved from <https://www.researchgate.net/publication/229310653>
- Liu, R., Jiang, Y., & Azevedo, C. L. (2021, 6). Bayesian optimization of area-based road pricing. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/MT-ITS49943.2021.9529326
- Liu, R., Jiang, Y., Seshadri, R., Azevedo, C. L., & Ben-Akiva, M. (n.d.). *Contextual bayesian optimization of congestion pricing with day-to-day dynamics*. (unpublished)
- Lu, Y., Adnan, M., Basak, K., Pereira, F. C., Ben-Akiva, M., Carrion, C., ... Saber, V. H. (2015). Simmobility mid-term simulator: A state of the art integrated agent based demand and supply model.. Retrieved from <https://www.researchgate.net/publication/279223479>

- Luo, Q. Y., Zhu, J. X., Guan, X. Y., & Jia, H. F. (2019). Optimal cordon-based congestion pricing considering feasible cordons and congestion transfer. *IEEE Access*, 7, 93545-93557. doi: 10.1109/ACCESS.2019.2928368
- Mckay, M. D., Beckman, R. J., & Conover, W. J. (2000, 2). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42, 55-61. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/00401706.2000.10485979> doi: 10.1080/00401706.2000.10485979
- Pigou, A. C. (1920). *The economics of welfare*. Macmillan.
- Rasmussen, C. E., & Williams, C. K. I. (2005). *Gaussian processes for machine learning*. The MIT Press. Retrieved from <https://direct.mit.edu/books/book/2320/gaussian-processes-for-machine-learning> doi: 10.7551/mitpress/3206.001.0001
- Srinivas, N., Krause, A., Kakade, S. M., & Seeger, M. W. (2012, 5). Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58, 3250-3265. doi: 10.1109/TIT.2011.2182033
- Verhoef, E. T. (2005, 11). Second-best congestion pricing schemes in the monocentric city. *Journal of Urban Economics*, 58, 367-388. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0094119005000495> doi: 10.1016/j.jue.2005.06.003
- Wang, Y., Geng, K., May, A. D., & Zhou, H. (2022, 3). The impact of traffic demand management policy mix on commuter travel choices. *Transport Policy*, 117, 74-87. doi: 10.1016/j.tranpol.2022.01.002
- Zhong, S., Gong, Y., Zhou, Z., Cheng, R., & Xiao, F. (2021, 4). Active learning for multi-objective optimal road congestion pricing considering negative land use effect. *Transportation Research Part C: Emerging Technologies*, 125. doi: 10.1016/j.trc.2021.103002

A stated choice experiment to estimate preference for fully automated taxis: comparison between immersive virtual reality and online surveys

Hao Yin*¹, Elisabetta Cherchi²

¹ Research Assistant, School of Engineering, Newcastle University, UK

² Professor, School of Engineering, Newcastle University, UK

SHORT SUMMARY

In this paper, we aim to study the role of the immersive Virtual Reality (VR) experience in the preferences elicited with standard Stated Choice (SC) experiments embedded into a VR environment. For this purpose, a SC was built and implemented both online and within a VR environment and respondents were asked to reply to both surveys. The SC experiment consists of a binary choice between a normal taxi with the driver (NT) and a fully automated taxi (AT). The context is a well-known street in the city centre of Newcastle upon Tyne (UK). Hybrid choice models were estimated and results compared. Results suggest that VR experience indeed has significant effects on some attributes examined (waiting time and good reviews) and on the role of the latent variables in the choice of AT. Trust is significant only online, while injunctive norms and perceived safety only in the VR environment.

Keywords: Immersive Virtual Reality, Stated Choice Experiment, Fully Automated Taxis, Social Conformity, Internal validity

1. INTRODUCTION

Stated Choice (SC) experiments are commonly used to investigate users' acceptance of Autonomous Vehicles (AV), as this is a product not yet available in the market. Hypothetical bias affects all SC experiments, but it is more marked in the case of highly innovative products as respondents have no experience with them and could not have formed a preference for the product (see a discussion in Cherchi and Hensher, 2015). Pictures and videos have been increasingly used to provide more realistic information about the new products and, in particular in the case of AV to show respondents how the system could work (e.g., Haward and Dai, 2014; Kolarova et al., 2018). In this line, recent applications have also used Virtual Reality (VR) environment, mostly with the aim to control for the framing effect, i.e., to improve the preliminary information about the non-markets good respondents were going to evaluate (Bateman et al., 2009; Fiore et al., 2009; Phillips and Marsh 2015; Patterson et al., 2017).

VR experiments represent a new area of research that promises to change fundamentally the way in which consumers' preferences for innovations are measured. Studies have shown that people can develop realistic spatial knowledge in the VR environment that is like actual physical environments (O'Neill 1992; Ruddle et al., 1997; Tlauka and Wilson 1996). VR can generate a sufficiently natural and familiar field, able to provide 'field cues' or 'field hints' that occur in the real world (Fiore et al., 2009). VR allows the sensation of immersion in the activities on the screen and with the virtual elements (Animesh et al. 2011; Faiola et al. 2012; Nah et al., 2011), prompting individuals to act as if they were in the real world (Sanchez-Vives et al. 2005). Based on this

theory the expectation is that VR experiments should perform better in eliciting individual preferences, i.e., should provide more realistic results than a standard online survey. However, very little is known in this area.

Some authors (Farooq et al., 2018; Arellana et al., 2020; Bogacz, et al., 2021; Feng et al., 2022) have used VR technology with stated preference experiments applied to pedestrian or cycling experiments, which involves continuous movement, not a choice among discrete alternatives. In some studies, for modelling purposes, the continuous behaviour has been converted into a choice, but from a neurological point of view, motor actions (like cycling or walking) activate different circuitries in our brain compared to choice-based actions and show a better overlap between brain activities during imagined and real movements (see discussion in Cherchi, 2020). Moreover, the stated preference experiment in these studies is used to control the elements of the VR environment where respondents perform a continuous behaviour. This makes it more difficult to assess internal validity, i.e., the impact of the VR experience in the elicited consumer preferences, compared to a traditional stated choice online survey. Rossetti and Hurtubia (2020) studied the ecological validity of VR experiments (i.e., whether the results can be generalized to real-life settings), but the focus is on the qualitative assessment of aspects of an urban environment.

In this paper, we aim to study the impact of the immersive VR experience in the preferences elicited with standard SC experiments embedded into a VR environment. For this purpose, a SC was built and implemented both online and within a VR environment and respondents were asked to reply both surveys. The SC experiment consists of a binary choice between a normal taxi with the driver (NT) and a fully automated taxi without driver and without steering wheel (NT). A set of attitudinal questions were also included in the survey to investigate if the VR environment has a mediating role on the impact of Trust, Injunctive Norm and Perceived Safety in the choice of the taxi type. Hybrid choice models were estimated with both surveys allowing assessing the impact of the VR experiment in the elicited consumer preferences.

2. METHODOLOGY

The core of the methodology set up in this research consists of a SC experiment built to elicit preferences for automated versus normal (i.e., with driver) taxis. The same SC experiment is used to collect information online and within a VR environment. Differently from the existing literature in the field, respondents go through a standard stated choice experiment while moving in the VR environment and “living” the choice experience. In our immersive VR experiment, respondents found themselves into the street where there is a taxi rank with the two types of taxis picking up customers and a ticket board (as it is the case in reality) where to select and pay for the taxi respondents wish to use. The advantage of this setting is that the SC experiment in the VR environment is perfectly comparable with the standard SC online, allowing to better disentangle the impact of the immersive VR experience in the elicited preferences.

The SC experiment built includes six attributes. Three level of service attributes (waiting time, travel time and fixed journey fare, with three levels each), one attribute to measure the impact of fuel type (with 2 levels: electric or gasoline) and two attributes to measure the impact of social conformity (number of customers who have used AT or NT, with three levels and customer rating with two levels). A heterogeneous Bayesian efficient design was generated in Ngene (ChoiceMetrics 2012). Priors were taken from models estimated in several pilot tests based on orthogonal designs. Three SC experiments were optimised based on three travel distances of 5 km, 10 km and 15 km (the only differences are the attribute level value of ‘travel cost’ and ‘travel

time' among 3 SC designs). 16 choice scenarios were generated and randomly divided in 2 blocks. Each respondent was presented with 8 scenarios.

The above process is not different from a standard screen-based SC. However, building comparable SC experiments, when implemented in VR and online, poses some challenges. The most important and interesting issue is that some of the elements that are typically used in an online SC and are considered perfectly acceptable, look unrealistic when used in a VR-based environment. For example, we had to set the context at a taxi rank to allow respondents to make the choice in the virtual street. Traditional taxi services are still extensively used in Newcastle and there are numerous taxi ranks in the city centre. Some attributes appeared to be unrealistic when included in the SC in the VR environment. For example, the form of payment is a standard attribute often included in online SC experiments, but interestingly when used in the SC experiment within the VR environment it appeared clearly unrealistic. This is because when we choose a transport option in reality, we are only presented with the characteristics of the options, and only after we make the choice the ticket machine asks us how we would like to pay. The realism of the SC experiment embedded in the VR experiment made this problem evident. After several tests, the attribute was removed from the experimental design, and the question "how do you want to pay?" included after each SC scenarios (both in the online and the VR-bases survey). Finally, in standard SC experiments, respondents are typically presented with the destination of their most recent trip once at the beginning of the SC and before each scenarios are asked to assume that they have to do a trip always with the same destination. Interestingly, this standard procedure, that sounds perfectly reasonable in the screen-based SC, appeared extremely unrealistic in the VR environment. This is because in the VR environment respondents "live" the choice process, any assumption that they have to do the same trip felt awkward. Differently from the standard practice, we allowed respondents to choose different destinations in each scenario, and hence the 6 scenarios presented to each respondent can belong to any of the 3 designs (5km, 10km or 15km).

The VR-based SC experiment and its online counterpart were administrated in Newcastle in 2022. The final sample consists of 156 valid responses (1248 pseudo-individuals). These 156 respondents answered first the online SC survey, which according to the standard practice included: general questions about familiarity with automated vehicles, information about a last trip performed by normal taxi, a stated choice experiment, a set of socioeconomic information and nine statements to measure three latent psychological constructs: injunctive norm, perceived safety and trust. Approximately one week after, respondents were invited to the lab to perform the same SC experiment but this time embedded into the VR environment. All respondents interviewed are residents in the northeast of England and satisfy the requirements to be 18 years or older and have used a normal taxi in Newcastle in the last year. Table 1 reported the key characteristics of the sample. Our sample approximates the gender distribution of the Newcastle population but underrepresents young people (20% against 25% in the Newcastle population).

Table 1. Sample Characteristics

Socio-demographic characteristics		%
Gender	Female	52.6
	Male	46.8
Age	Rather not to say	0.6
	Younger than 30 years old	19.9
	30 years old or older	80.1
Education level	Bachelor degree or below	66.0
	Master or Doctorate degree	34.0
Current work status	Employed full-time	61.5
	Others	38.5
Personal monthly disposable income	Less than £500	12.8
	£501-£1500	40.4
	£1501-£2500	21.8
	£2501- £3500	9.0
	£3501-£ 4500	2.6
	More than £4500	0.0
	I do not wish to disclose it	13.5
Travel characteristics		%
Frequency of using taxis	More than once a week	10.3
	Between Once a month & Once a week	42.3
	Between Twice a year & Once a month	39.7
	Less than twice a year	7.7
Frequency of talking with driver	Very infrequently	8.3
	Somewhat infrequently	8.3
	Occasionally	30.1
	Somewhat frequently	36.5
	Very frequently	16.7
Knowledge levels of AVs and ATs		%
Heard of AVs	Yes	69.2
	No	30.8
Familiar with 5 levels of automation	Not at all familiar	39.7
	Slightly familiar	35.3
	Moderately familiar	19.9
	Very familiar	4.5
	Extremely familiar	0.6

3. MODELS ESTIMATED AND RESULTS

Hybrid choice models (HCM) were used to elicit user preferences. The discrete choice component of the HCM is a mixed logit (ML) model that allows estimating the trade-off between the attributes included in the SC experiment, controlling for panel effects (intra-individual correlation). The latent variable component of the HCM allows for estimating the impact of three latent psychological variables.

Model results are reported in Table 2. We note first that in both datasets, all the level of service attributes (travel time, waiting time and travel cost) have the expected sign and are significant at more than 99%. The same for the ‘good reviews’, and to some extent for the type of fuel (‘EV’).

While strong differences are found in the significance of the latent variables. We also note that the HCM estimated with the VR data has overall much better overall fit than the model estimated with online data (lower BIC and AIC).

Table 2. Models estimation results

	HCM Online		HCM VR	
	Estimated Value	Rob. t-test	Estimated Value	Rob. t-test
ASC(AT)	-7.730	-6.42	-5.800	-7.18
SIGMA (AT)	0.931	4.91	1.290	7.40
Level of Services				
Travel cost	-0.535	-9.88	-0.629	-12.47
Travel time	-0.083	-4.68	-0.127	-6.89
Waiting time	-0.130	-7.14	-0.189	-10.71
AT Vehicle Type				
EV	0.331	1.90	0.554	3.22
Social Conformity				
Good review	0.384	3.46	0.754	7.01
Systematic heterogeneity in alternatives				
Bachelor degree or above(AT)	0.614	2.38	-	-
Frequently talking with driver (AT)	-0.385	-1.57	-	-
Male (AT)	0.948	3.50	0.966	3.36
Latent variable				
Injunctive norm (AT)	0.336	1.18	0.485	2.41
Perceived Safety (AT)	0.397	1.50	0.805	4.05
Trust (AT)	1.400	4.07	0.339	1.32
Summary of Statistics				
Number of draws	500		500	
Maximum Log-likelihood	-2220.789		-2212.280	
Akaike Information Criterion	4537.578		4512.569	
Bayesian Information Criterion	4783.784		4738.258	
Number of individuals	156		156	
Number of observations	1248		1248	

Before looking at the differences between the two models, we note that a joint hybrid choice model was estimated to control for possible scale heterogeneity between online and VR-based SC data. The scale between online and VR-based SC data was not significantly different from one. We then compare the results using the models estimated separately. While the scale is not different, results show that the estimated preferences are significantly different between the online and the VR-based dataset for several attributes. The estimated marginal utility of travel cost, travel time, and fuel type is the same in the online and in the VR survey, while the estimated marginal utility of waiting time and good reviews are significantly higher in the VR than online. Willingness to pay (WTP) for saving waiting time is 1.2 times higher in the VR (18.0 £/hour) than in the online survey (14.6 £/hour) and the WTP for good reviews is more than 1.5 times higher in the VR (£ 1.2) than in the online survey (£ 0.72). This result is interesting and expected. Both results can be related to the higher realism provided by the VR environment. In the VR respondents can see other customers queueing, can move in the space where they are going to wait for the taxi, this could prompt a more realistic evaluation of the waiting time. If this argument is correct, then we should conclude that online surveys do underestimate the WTP for waiting time. If the choice context is more realistic, customer reviews provide respondents with a stronger hint or cue to

evaluate and compare the quality of these two types of taxi services, because respondents might have the feeling that they are really going to take the taxi.

Another interesting finding is that the VR environment indeed affects the role of the latent psychological constructs in the choice of automated taxis. Interestingly, trust significantly affects the preference for ATs only in the online survey, which makes sense because participants have not seen these innovative taxis operating, then trust in the automation is more important than in the VR experiment where they can see ATs operating around them. In line with that, when respondents see ATs really on the road, then the perception of safety becomes more important. Finally, injunctive norms are significant only in the VR environment, and this also makes sense, because the realism of the VR might give a feeling of being seen by others, which is related to the social norms.

4. CONCLUSIONS

This paper discussed the impact of the immersive VR experience in the preferences elicited with SC experiments. The same SC experiment was implemented in an online survey and within a VR environment and respondents were asked to reply to both surveys. Hybrid Choice models were estimated and preferences estimated in the two environments compared. Results suggest that the immersive VR environment does have an impact on the preferences elicited with SC experiments. Notably, it seems that the immersive experience has a strong impact on the preference for waiting time and customer reviews, both attributes are related to the experience that respondents can live in the immersive VR. Results also show that the role of the latent psychological factors tested (trust, injunctive norms, and perceived safety) is different (opposite) in the VR and in online. Before knowing or experiencing how the AT service works, trust plays a critical role. After or during the (virtual) experience, perceived safety becomes relevant instead of trust, and injunctive norms also become significant. Even if respondents replied to the two surveys at approximately one-week length one from the other, results could still be affected by order effect. At the same time, since the same respondents replied to both surveys, we are sure that the results are not affected by differences in socio-economic characteristics.

ACKNOWLEDGMENTS

This research was carried out as part of the Veronica project, funded by ESRC, UK.

REFERENCES

- Animesh, A., Pinsonneault, A., Yang, S.B., and Oh., W. 2011. An Odyssey into Virtual Worlds: Exploring the Impacts of Technological and Spatial Environments on Intention to Purchase Virtual Products. *MIS Quarterly: Management Information Systems*, Vol. 35, No. 3, pp.789-810.
- Arellana, J., Garzón, L., Estrada, J. and Cantillo, V., 2020. On the use of virtual immersive reality for discrete choice experiments to modelling pedestrian behaviour. *Journal of choice modelling*, Vol. 37, 100251.

Bateman, I.J., Day, B.H., Jones, A.P. and Jude, S., 2009. Reducing gain–loss asymmetry: a virtual reality choice experiment valuing land use change. *Journal of environmental economics and management*, Vol. 58, No. 1, pp.106-118.

Bogacz, M., Hess, S., Calastri, C., Choudhury, C.F., Mushtaq, F., Awais, M., Awais, M. Nazemi, M., van Eggermond, M.A.B. and Erath, A., 2021. Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: an application to virtual reality cycling. *Transportation research part C: emerging technologies*, Vol. 133, 103435.

Cherchi, E. (2020) Our IATBR: 45 years of contribution to transport behaviour research. In *Mapping the Travel Behavior Genome*. K. Goulias (ed.) Elsevier, Chapter Vol. 1, pp. 17-28.

Cherchi, E. and Hensher, D.A., 2015. Workshop synthesis: Stated preference surveys and experimental design, an audit of the journey so far and future research perspectives. *Transportation Research Procedia*, Vol. 11, pp.154-164.

ChoiceMetrics, N., 2012. *1.2 User Manual & Reference Guide*, Australia.

Faiola, A., Newlon, C., Pfaff, M. and Smyslova, O. 2012. Correlating the Effects of Flow and Telepresence in Virtual World: Enhancing Our Understanding of User Behavior in Game-Based Learning. *Computer in Human Behavior* Vol. 29, pp. 1113–1121.

Farooq, B., Cherchi, E. and Sobhani, A., 2018. Virtual immersive reality for stated preference travel behavior experiments: A case study of autonomous vehicles on urban roads. *Transportation research record*, Vol. 2672, No. 50, pp.35-45.

Feng, Y., Duives, D.C., and Hoogendoorn, S.P., 2022. Development and evaluation of a VR research tool to study wayfinding behaviour in a multi-story building. *Safety science*, Vol. 147, 105573.

Fiore, S.M., Harrison, G.W., Hughes, C.E. and Rutström, E.E., 2009. Virtual experiments and environmental policy. *Journal of Environmental Economics and Management*, Vol. 57, No. 1, pp.65-86.

Howard, D., Dai, D., 2014. Public perceptions of self-driving cars: the case of Berkeley, California. In: *93rd Annual Meeting of the Transportation Research Board*. Washington, D.C.

Kolarova, V., Steck, F., Cyganski, R. and Trommer, S., 2018. Estimation of the value of time for automated driving using revealed and stated preference methods. *Transportation research procedia*, Vol. 31, pp.35-46.

O'Neill, M.J. (1992). Effects of Familiarity and Plan Complexity on Wayfinding in Simulated Buildings. *Journal of Environmental Psychology* Vol. 12, No. 4, pp. 319-327.

Patterson, Z., Darbani, J.M., Rezaei, A., Zacharias, J. and Yazdizadeh, A., 2017. Comparing text-only and virtual reality discrete choice experiments of neighbourhood choice. *Landscape and Urban Planning*, Vol. 157, pp.63-74.

Phillips, Y. and Marsh, D., 2015. Virtual Reality and Scope Sensitivity in a Choice Experiment About Coastal Erosion. 59th Conference Australian Agricultural and Resource Economics Society, February 10-13, 2015, Rotorua, New Zealand.

Rossetti, T. and Hurtubia, R., 2020. An assessment of the ecological validity of immersive videos in stated preference surveys. *Journal of choice modelling*, Vol. 34, 100198.

Ruddle, R.A., Payne, S.J. and Jones, D.M. 1997. Navigating Buildings in Virtual Environments: Experimental Investigations Using Extended Navigational Experience. *Journal of Experimental Psychology: Applied*, Vol. 3, No. 2, pp.143-59.

Sanchez-Vives, M.V. and Slater, M., 2005. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, Vol. 6, No. 4, pp.332-339.

Tlauka, M. and Wilson. P.N. 1996. Orientation-Free Representations from Navigation through a Computer-Simulated Environment. *Environment and Behaviour*, Vol. 28, No. 5, pp.647-664.

Travel mode choice modelling of visually impair people through latent variables

Fernanda Guajardo and Sebastián Raveau

Department of Transport Engineering and Logistics
Pontificia Universidad Católica de Chile, Chile

SHORT SUMMARY

It has been shown that people with disabilities perceive certain travel attributes differently, affecting their behaviour. It is relevant to understand the behaviour of people with disabilities to support public policies that address their needs. The objective of this research is to identify the factors that affect the mobility decisions of blind or visually impaired people, taking Santiago de Chile as a case study. With information from a total of 1,322 trips in Santiago made by people with and without disabilities, hybrid models of modal choice were estimated, including two latent variables: human interactions and use of technology. People who use more technology prefer ride hailing. Modes with direct contact with the driver are perceived more positively by people who assign importance to human interactions. Additionally, there is a significant difference in the perception of walking time. Walking time affects approximately 30% more blind or visually impaired people than people without visual impairment. Based on the results, there is proof of the relevance of having public policies that ensure subsidized taxi trips for people with visual disabilities or people with reduced mobility.

Keywords: hybrid discrete choice models, blind, visually impaired, latent variables, mobility, use of technology, human interactions

1. INTRODUCTION

Transport systems are vital for the development of society. Some people, particularly people with disabilities, may have difficulty entering a transport system that does not take their needs into account (Hallgrimsdottir et al., 2016). Although there have been advances to better support the needs of people with disabilities, there are still barriers that restrict independent travel for this group of people (Park and Chowdhury, 2018). It is relevant to understand more clearly the travel behaviour of people with disabilities to prevent depression, poverty, and other socioeconomic harms (Ermagun et al., 2016). Since social exclusion is often the result of the inability to use or access a public transport system (Park and Chowdhury, 2018).

The planning and design of integrated systems have been predominantly focused on public transport users without disabilities (Park and Chowdhury, 2022). Therefore, transport systems should aim to offer quality service for all users, by providing policies according to their needs, especially for those with disabilities. Certainly, not all people with disabilities have the same needs, specific and different needs respecting disability groups shall be also considered.

Low et al. (2020) state that “*common perceptions tend to focus on the provision of barrier-free access for wheelchair users. This group is of course important, but there are other types of disabilities, including those that are visually impaired*”. In the context of Chile, the second most common disability within the adult population with some kind of disability, after physical difficulties is a visual impairment with 11.9% having difficulties seeing even when wearing glasses (MDS,

2016). Loss of vision or blindness makes it difficult for people to move and affects their independence when traveling (Low et al., 2020).

Therefore, it is essential to understand better the needs of Visually Impaired People (VIP) to develop more effective public policies, especially in developing countries with limited resources. However, disability has been treated in a general way, generating improvements that people with visual disabilities may not necessarily benefit from. The main objective of this research is to quantify aspects of the trip that affect VIP travel mode choices to support public policies that are aimed at the needs of VIPs, taking Santiago de Chile as a case study, using the information of VIPs trips captured by a revealed preferences survey. However, to the authors' knowledge, studies of mode choice decisions specifically among VIPs have not been carried out nor in developing countries.

There are various studies on the choice of mode and trip generation of elderly and disabled people but most of these have not considered the types of disabilities of the individuals. Also, most of the research has been carried out in developed countries, contemplating travel modes that are not necessarily available in underdeveloped countries, such as paratransit systems. In addition, the context of Santiago differs, in terms of social and cultural context, levels of services experimented on public transport, and non-rate reductions in the public system for people with disabilities, among others.

This short paper is structured as follows. Section 2 explains how the data was obtained and an analysis of the collected information is performed, while Section 3 describes the modelling approach, and the results are discussed. Finally, section 4 closes with the conclusions and the possible future research.

2. SURVEY DESIGN

This section describes the survey and how the information was collected. It also mentions how the service levels were obtained. Finally, it explains how the availabilities of travel modes were defined.

Accessibility aspects

To make an accessible instrument for VIPs, various considerations were taken. Firstly, it was relevant to keep the respondent constantly informed of the aspects of the survey, such as: duration of the survey with a screen reader, number of alternatives for drop-down alternative questions, structure of each section and total amount of questions by section. Secondly, the survey avoids the type of drop-down responses, especially if there are a lot of possible alternatives (for example, when asking about the commune of residence). In such cases, we prioritized open answers. To make the information retrieval process easier, we asked for more general aspects of the trip and then more specific aspects. It is important to mention the survey did not require the use of images. Finally, we used *Google Forms* because according to the Google Forms Accessibility Compliance Report (Google, 2019) it is compatible for people with vision limitations and partially compatible for blind people by providing important accessibility information for most interface elements.

Survey description

In the first section of the survey, we obtained information on the individual characteristics, such as gender, age, residence address, socioeconomic levels, visual condition, knowledge of braille, mobility support, and whether the person has reduced mobility. In the second section of the survey, we gathered information on the last trips made from the home of the respondent, respondents could declare at most three trips and at least one trip. We gathered information about each trip, which consisted of motive, day, time and travel mode of the trip, and destination address. The third section captures indicators of the individual's attitude related to possible latent variables, which can be seen in Table 1. The individuals had to declare through a *likert* scale from 1 to 7 if they agreed or disagreed with each statement.

Table 1- Attitude indicators

Potential latent variable	Attitude indicators	Abbreviation
Use of technology	"If I go to a place I don't know, I use technological Apps"	<i>UT₁</i>
	"I use technological Apps to have information about my trip"	<i>UT₂</i>
Human interactions	"It makes me feel safe knowing that there are people around me"	<i>HI₁</i>
	"I care about being treated cordially by people I don't know"	<i>HI₂</i>

Level of Services

We used the *Google API* to obtain the level of services. To have a more accurate representation of the level of services experimented on a trip we used the average travel times between the reported routes of the *API* (between 2 and 4 routes). In the case of taxis and Apps mobility services the same average time as the car was considered, and an extra time of 10 minutes was added for the case of taxis and 5 minutes for Apps mobility services. For public transport trips, in addition to the average time in the vehicle, we considered the average waiting times, average walking distance, and amount of transfer by mode (Metro/Bus).

For the cost of the alternative public transport, only bus and only metro we considered the current costs determined by the Metropolitan Public Transport Directory. For the cost of the car trip, we considered a cost by kilometre with a mixed performance of 14.6 [Km/l], the cost per kilometre was used because a priori it cannot be assumed whether the person had to pay for parking or pay a toll. For the price of gasoline, we used the average price of the different service stations in Santiago. For taxi costs we used the actual costs considering a base rate and a charge for every 200 [m] travelled. On the other hand, in the case of mobility apps services we considered the tariff system of the company Cabify. Finally, for the bicycle mode, we used a cost based on distance assuming a monthly subscription of a Chilean bike share company and a daily use of 8 [Km].

Mode availability

All individuals had the following modes of transportation available: walking, taxi, and application mobility services. VIP had the mode bicycle unavailable and if the individual did not have any visual impairment the bicycle was considered available in case if in other trips reported the travel mode was bicycle was declared or if there was a shared bicycle station in a radius less than or equal to 500 [m] from their home and destination. Similarly, in the case of a car (whether a driver or a passenger car), was considered available for all the trips of people who had used the car in one of the three declared trips and for those who did not declare having used a car were assigned availability by replicating the distribution of the socioeconomic groups of the people who used a

car. Finally, in the case of public transport, the alternative to all trips was assigned to them. For the only bus and only Metro modes, availability was assigned in case there were routes available only with those modes.

Data Collection

We sampled 484 individuals with valid trips, 25 of those declared information on one trip, 80 on two trips, and 379 on three trips. Therefore, we obtained information on a total of 1,322 trips. We were able to collect trips through all sectors of Santiago, however, to replicate the distribution of gender, age, and socioeconomic group of the population more precisely in the sample, correction factors were used through Furness Method (1965).

After applying the correction factors, 22.5% of the trips were made by visually impaired people, while 77.5% were made by people without visual impairment. 80.6% of the total trips in the survey were made on a weekday, while 19.4% were made on the weekend. A low percentage of trips (6.8%) were made in active modes such as walking and cycling, the rest of the trips, were made by through motorized modes. 60.2% of the trips were made in public transport modes: Metro and bus, or the use of both. 24.6% of the trips were made by car. Among the modes with a lower use percentage are walking, taxi, Apps mobility services, and taxi, with 5.5%, 4.5% 4.0% and 1.3%, respectively. Most of the trips are in the range of 4 [Km] to 6 [Km] and the average of travel distances is 11.82 [Km] explaining why only 6.8% of the trips are made on active modes.

3. MODE CHOICE MODELLING

This section details the estimated hybrid model, the hierarchical structure of the discrete choice model, and then the values of the estimated parameters are presented. It is important to consider that the time is in hours and the costs are in euros (€).

Hybrid model

For the estimation of hybrid models, we considered the simultaneous method that considers estimating the MIMIC model and the discrete choice model simultaneously (Raveau et al., 2010). Two of the four potential latent variables were significant for the modelling, the other two did not explain the discrete decision of travel mode. The number of observations used to estimate the 51 parameters of the hybrid model was 1,322 observations. The final log-likelihood was -7,404 as shown in Table 2.

Table 2- Main model indicators

Number of observations	1,322
Number of estimated parameters	51
Final log-likelihood	-7,404

The structure of the MIMIC model incorporated into the hybrid model (shown in Figure 1) considers the characteristics of the individuals corresponding to the structural equations. In the case of the latent variable, “human interactions”, three binary variables associated with the characteristics of the respondents were considered, indicating if the person has: higher education, reduced mobility, or visual impairment. With respect to the measurement equations, the latent variable is composed by HI_1 indicator and HI_2 indicator. The specification of the structural equation of this

latent variable can be seen in Equation (1). On the other hand, for the structural equation of the latent variable, “use of technology”, the following characteristics were considered: high income and older than 60 years. For the measurement equations we used the UT_1 and UT_2 indicators for the latent variable “use of technology”. The specification of the structural equation of the latent variable is found in Equation (2).

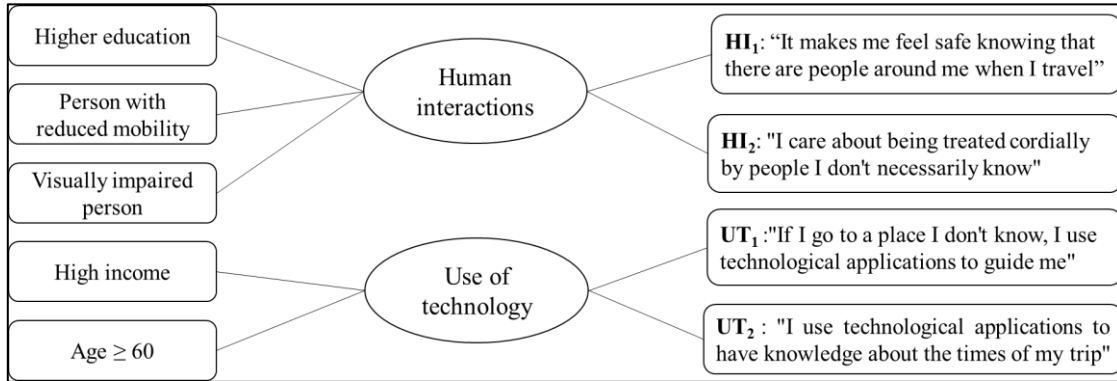


Figure 1 - MIMIC structure

$$HI = s_{HE} \cdot HigherEducation + s_{RM} \cdot ReducedMobility + s_{VIP} \cdot VisuallyImpaired + \sigma_{HI} \quad (1)$$

$$UT = s_{HI} * HighIncome + s_E * Eldery + \sigma_{UT} \quad (2)$$

The estimated parameters from the MIMIC model of the latent variable human interactions, presented in Table 3, are consistent with what is expected (only the structural equations' results are presented). Parameter s_{HE} is positive, and therefore people with higher education value more human interactions when travelling. Parameters s_{RM} (significant with a confidence level of over 80%) and s_{VIP} are positive, so people with reduced mobility and VIP perceive human interactions as important, with VIP being the ones who most assign importance to human interactions.

Table 3- Estimated parameters of the latent variable human interactions

	Parameter	Description	Value	t-test
Structural equation	s_{HE}	Higher education	0.42	3.68
	s_{RM}	Reduced mobility	0.37	1.60
	s_{VIP}	Visually impaired people	1.17	5.93
	σ_{HI}	Standard deviation	0.78	5.81

Table 4 presents the estimated parameters for the structural equation of the latent variable use of technology. People with high income have a positive sign and people over 60 years of age have a negative sign, which indicate opposite effect. People with high income have a greater use of technology than people who are not part of this category and people over 60 years of age have a lower use of technology.

Table 4- Estimated parameters of the latent variable use of technology

	Parameter	Description	Value	t-test
Structural equation	S_{HI}	Higher income	2.11	7.13
	S_E	Elderly	-4.34	-9.02
	σ_{UT}	Standard deviation	-2.96	-10.70

Given that the Public Transport System in Santiago is fully integrated both physically and fare-wise (DTPM, 2022), it was decided to use a hierarchical structure in the discrete choice model. This structure will allow capture the correlation between the bus, Metro and Metro-bus alternatives, which will be grouped in a nest of public transport as shown in Figure 2.

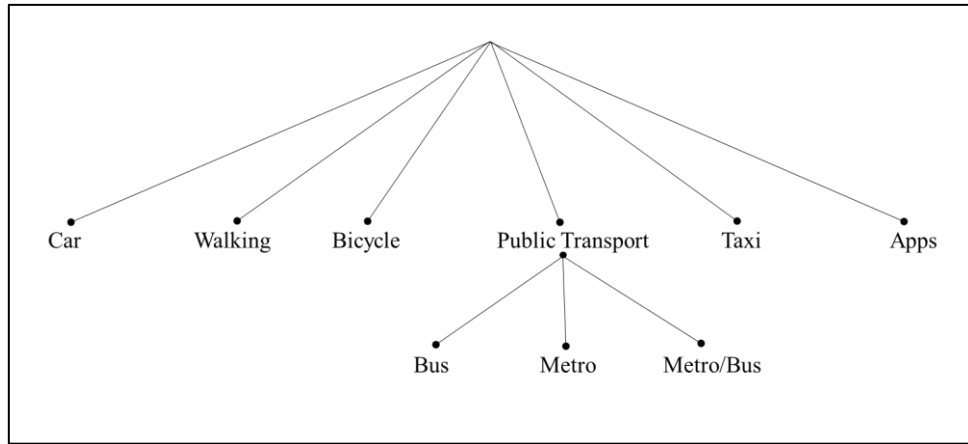


Figure 2- Hierarchical structure of the discrete choice model

Regarding utility functions by mode, the specification's contingency table is presented in Table 5. All modes have an associated modal constant, in the case of walking, the modal constant was set to zero. All modes except walk have a cost parameter associated with them. A generic parameter for time in the vehicle was considered, except for the bicycle mode, which has a specific parameter, since this mode involves an effort that the cyclist must do. We used a systematic variation of tastes in relation to the binary variable of VIP to analyse how walking time affects VIP. This interaction was included in the modes of walking and those related to public transport. For modes associated with public transport, a waiting time variable was also included to understand how this variable affects people. In addition, a variable for the number of transfers from bus to another bus was included in the bus and Metro/bus travel mode alternatives. For the mode only Metro, this variable is not included since, by design, it will always take the value zero. The latent variable of technology use was included in the utility function of the Apps mobility services where, in addition, a parameter associated with the variable latent variable of human interactions for this same alternative, finally, there is a specific parameter for the latent variable of human interactions for taxi.

Table 6 shows the estimated parameters from the mode choice model. Regarding significance, all variables (ignoring modal constants) are significant at a 95% confidence level. The value of the t-test with respect to 1 of the parameter $\lambda_{PublicTransport}$ is 4.95, validating the structure presented in Figure 2. Therefore, the correlation of the alternatives within the public transport nest is 87.8%.

Table 5- Utility function of the different travel modes

Travel mode	ASC	Bi-cycle time	Walking time with VIP interaction	Vehicle time	Waiting time	Cost	Bus-Bus transfer	Latent variable <i>HI</i>	Latent variable <i>UT</i>
Car	S			G		G			
Walking	*		G						
Bicycle	S	S				G			
Bus	S		G	G	G	G	G		
Metro	S		G	G	G	G			
Metro-Bus	S		G	G	G	G	G		
Taxi	S			G		G		S	
Apps	S			G		G		S	S

S: specific parameter / G: generic parameter / *: fixed parameter

Table 6- Estimated parameters for the choice model

Parameter	Value	t-test
ASC_{Car}	-0.13	-0.572
$ASC_{Walking}$	0.00	-
$ASC_{Bicycle}$	-1.32	-2.83
ASC_{Bus}	0.12	0.554
ASC_{Metro}	0.76	4.16
$ASC_{Metro-Bus}$	0.28	1.30
ASC_{Taxi}	-4.94	-4.41
ASC_{Apps}	-4.77	-4.14
$\lambda_{PublicTransport}$	2.86	7.63
$\beta_{transfersBusBus}$	-0.108	-2.75
β_{cost}	-0.753	-7.52
$\beta_{TimeBicycle}$	-5.18	-3.28
$\beta_{TimeWalking}$	-3.4	-10.30
$\beta_{TimeVehicle}$	-0.898	-3.68
$\beta_{TimeWaiting}$	-1.60	-4.46
$\alpha_{WalkingVIP}$	-1.12	-2.37
θ_{UT}	0.389	3.66
θ_{IH_Apps}	2.64	2.74
θ_{IH_Taxi}	3.82	3.79

The parameters θ_{IH_Taxi} and θ_{IH_Apps} are statistically different with 90% confidence. The value of the latent variable, human interactions, is positive both for the taxi mode and for Apps, meaning that people who perceive human interactions as relevant tend to choose modes such as taxi or Apps, this may be because these modes are more personalized services where you must interact with a person (driver). Also, people who have a greater use of technologies will tend to use modes of transport that are requested through mobile applications.

The time parameters follow the relationship:

$$|\beta_{TimeVehicle}| < |\beta_{TimeWaiting}| < |\beta_{TimeWalking}| < |\beta_{TimeBicycle}|$$

In several practical studies the values of the walking and waiting time parameters are two or three times the value of the time in the vehicle (Ortúzar and Willumsen, 2011). In this case, the value of the waiting time parameter is approximately twice the value of the vehicle time parameter while the value of the walking time parameter is approximately four times the value of the vehicle time.

4. CONCLUSIONS

In this study, it was verified through discrete choice hybrid models that there is heterogeneity in the perception of tangible and non-tangible attributes in the choice of modes of transport by individuals. The main factors that affect the travel experience were identified according to the mode used by VIP or reduced mobility.

From the modelling, we obtained that people with visual disabilities perceive human interactions as relevant, which could influence their modal choice, preferring modes with direct contact with the driver. This same effect can be distinguished in people with reduced mobility, so they could prefer modes such as: taxi, Uber, Cabify or Didi due to the social relationships generated. This study has shown that the attitude we have and how we relate to each other as a society can affect people with visual disabilities and reduced mobility, leading them to prefer more expensive modes such as those mentioned above. There is heterogeneity in the perception of walking times, VIP are affected by approximately 30% more walking time than people without visual impairment. This implies that people who are blind or have low vision are willing to pay 30% more to save the same amount of walking time as a person without visual impairment.

The results of this research are a contribution since they can be used for the social evaluation of projects in Santiago as we quantify how much walking time affects people with visual impairments. Based on the results of the qualitative and quantitative study, the need to create transportation subsidies in Chile for people with visual disabilities is validated and we must also consider that in most countries various subsidised transport services are provided to service these population groups (elderly and disabled people) (Schmöcker et al., 2008). State subsidies are relevant to be applied in Chile, as mentioned above, people with visual disabilities are considerably more affected by walking time and, furthermore, both people with reduced mobility and people with visual disabilities tend to prefer modes of more costly transport like taxi.

It was expected in the modelling that the latent variable of human interaction would generate a positive impact on the metro mode given that within the functions of the metro attendants is assisting people with visual disabilities on the trip, however, this result was not obtained. A possible hypothesis to understand these results is the reduction of more than 1,500 Metro workers announced at the beginning of 2022. On the other hand, it was also expected that people with visual disabilities would be decisive in the latent variable of technology use. The Spanish disability observatory estimates that the population with disabilities is at a 33% disadvantage on the economic

ambit in relation to the general population (OED, 2022). Therefore, there could be a correlation between the economic income variable that would not allow the visual disability variable to be included at the same time.

Finally, as studies to be carried out in the future, two areas of interest are proposed to deepen the research presented. In the first place, it would be interesting to study how the variability of travel and waiting times affects people with disabilities as it was mentioned in the qualitative analysis but not explored in the quantitative models. And, on the other hand, it would be relevant consider vehicle crowding in the estimation of the models, which were not considered in this analysis since these service levels could not be obtained from a reliable data source for the different travel modes.

REFERENCES

- DTPM (2022). Sistema Integrado de Transporte. Directorio de Transporte Público Metropolitano. <https://www.tarjetabip.cl/tarifas.php>
- Ermagun, A., Hajivosough, S., Samimi, A. and Rashidi, T.H. (2016). A joint model for trip purpose and escorting patterns of the disabled. *Travel Behaviour and Society* 3, 51-58.
- Furness, K.P. (1965). Time function iteration. *Traffic Engineering and Control* 7, 458-460.
- Google. (2019). Google Forms Accessibility Conformance Report. <https://static.googleusercontent.com/media/www.google.com/es//accessibility/static/pdf/google-forms-vpat.pdf>
- Hallgrimsdottir, B., Wennberg, H., Svensson and H., Ståhl, A. (2016). Implementation of accessibility policy in municipal transport planning – Progression and regression in Sweden between 2004 and 2014. *Transport Policy* 49, 196-205.
- Low, W., Cao, M., De Vos, J., and Hickman, R. (2020). The journey experience of visually impaired people on public transport in London. *Transport Policy* 97, 137-148.
- OED (2022). La discapacidad como forma de desigualdad económica. Observatorio Estatal de la Discapacidad. <https://www.observatoriodeladiscapacidad.info/informe-breve-la-discapacidad-como-forma-de-desigualdad-economica/>
- Ortúzar, J. de D. and Willumsen, L.G. (2011). *Modelling Transport*. 4a edición, John Wiley & Sons, Chichester.
- Park, J. and Chowdhury, S. (2018). Investigating the barriers in a typical journey by public transport users with disabilities. *Journal of Transport & Health* 10, 361-368.
- Park, J., and Chowdhury, S. (2022). Investigating the needs of people with disabilities to ride public transport routes involving transfers. *Journal of Public Transportation* 24, 100010.
- Raveau, S., Álvarez-Daziano, R., Yáñez, M. F., Bolduc, D., and de Dios Ortúzar, J. (2010). Sequential and Simultaneous Estimation of Hybrid Discrete Choice Models: Some New Findings. *Transportation Research Record* 2156, 131-139.
- Schmoker, J., Quddus, M., Noland, R., and Bell, M. (2008). Mode choice of older and disabled people: a case study of shopping trips in London. *Journal of Transport Geography* 16, 257-267.

Microscopic Simulation-based Testing of Internal Boundary Control of Lane-free Automated Vehicle Traffic

Milad Malekzadeh*¹, Dimitrios Troullinos¹ Ioannis Papamichail¹,
Markos Papageorgiou¹

¹Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece

SHORT SUMMARY

The TrafficFluid concept allows vehicles to move in a lane-free environment and enables capacity sharing between two directions without lane restrictions. In this context, Internal Boundary Control (IBC) has recently been introduced by the authors and its application has been investigated using macroscopic models. This paper presents a microscopic simulation-based validation of IBC using SUMO TrafficFluid-Sim, i.e. a simulation tool able to implement lane-free traffic. A Linear Quadratic Regulator (LQR), that is a feedback control scheme, is employed for IBC. An extension of the well-known Cell Transmission Model (CTM) is utilized for the design of the controller. Simulation investigations confirm the effectiveness of the proposed scheme.

Keywords: internal boundary control, lane-free traffic, feedback control, SUMO.

1. INTRODUCTION

Recently, a new concept for vehicular traffic called TrafficFluid, which is applicable for high penetration rates of vehicles equipped with high-level automation and communication systems, was introduced by Papageorgiou et al. (2021). The TrafficFluid concept proposes: (1) lane-free traffic, whereby vehicles are not bound to fixed traffic lanes, as in conventional traffic; (2) vehicle nudging, whereby vehicles may exert a "nudging" effect on, i.e. influence the movement of, vehicles in front of them.

In this context, it turns out to be feasible to utilize Internal Boundary Control (IBC), a control strategy that aims to maximize the use of the road infrastructure (Malekzadeh et al., 2021a). IBC in lane-free traffic takes advantage of the fact that the traffic flow and capacity demonstrate incremental changes in reaction to corresponding incremental changes of the road width. Thus, on a highway or arterial with two opposite traffic directions, the total cross-road capacity (for both directions) may be shared between the two directions in real-time, according to the prevailing demand per direction, by virtually moving the internal boundary that separates the two traffic directions and communicating this decision to Connected automated Vehicles (CAVs), so that they respect the changed road boundary. The characteristics of IBC are analyzed by Malekzadeh et al. (2021a), where its high improvement potential is demonstrated by formulating and solving an open-loop optimal control problem. Additionally, a feedback-based Linear-Quadratic Regulator (LQR) for IBC was developed by Malekzadeh et al. (2021b), aiming at balancing the relative densities in the two directions; and has been demonstrated to be robust and similarly efficient as the open-loop optimal control solution, while avoiding the need for accurate modelling and external demand prediction.

So far, IBC has been investigated utilizing macroscopic models. This work demonstrates that the concept and the proposed control scheme can be similarly efficient even in a more realistic environment, as a microscopic simulator. We make use of the SUMO TrafficFluid-Sim, a simulation tool able to implement lane-free traffic (Troullinos et al., 2021).

2. INTERNAL BOUNDARY CONTROL (IBC)

To implement IBC, a feedback control strategy is considered. The proposed control scheme is designed based on an extension and linearization of the well-known Cell Transmission Model (CTM) (Daganzo, 1994). Let us assume two opposite traffic directions subdivided into n sections. Then, the total section capacity q_{cap} , the total critical density ρ_{cr} and the total jam density ρ_{max} , are shared between the two directions in each section based on the sharing factor $0 < \varepsilon_{i,min} \leq \varepsilon_i \leq \varepsilon_{i,max} < 1$, that is applied per section and which is known in IBC as the control input (Malekzadeh et al., 2021b).

In conventional traffic management, traffic densities (in veh/km) characterize clearly the state of traffic, depending on their value versus the critical density: free traffic (when density is lower than critical density), critical traffic (when density is around critical density) or congested traffic (when density is higher than critical density). However, in the proposed IBC concept, the critical density for each direction and section is not constant, but a linear function of the sharing factor and is changing according to the applied control action. Therefore, the density value by itself is not sufficient, in the IBC context, to characterize the traffic situation in a section. To address this issue, the following relations define the relative densities (dimensionless) per section and per direction. The relative density of section i and direction a or b is given by dividing the corresponding traffic density with the corresponding critical density, which, on its turn, depends on the sharing factor prevailing during the last time-step, as follows

$$\tilde{\rho}_i^a(k) = \frac{\rho_i^a(k)}{\varepsilon_i(k-1)\rho_{cr}}, \quad \tilde{\rho}_i^b(k) = \frac{\rho_i^b(k)}{(1-\varepsilon_i(k-1))\rho_{cr}}. \quad (1)$$

The relative densities reflect clearly the state of the traffic in the IBC context. Specifically, if the relative density of a section and direction is less than 1, it reflects under-critical (free-flow) traffic conditions; if it is around 1, it reflects capacity flow; and if it is greater than 1, it reflects over-critical (congested) traffic conditions (Malekzadeh et al., 2021a).

Linear Quadratic Regulator (LQR)

Linearization of the CTM dynamic equations around a nominal point was presented analytically by Malekzadeh (2021b). To achieve this, the one-step retarded control input was defined as a new state variable according to $\gamma_i(k+1) = \varepsilon_i(k)$, $i = 1, 2, \dots, n$. Following the linearization procedure by Malekzadeh (2021b), the linearized state-space model is

$$\mathbf{x}(k+1) = \hat{\mathbf{A}}\mathbf{x}(k) + \hat{\mathbf{B}}\mathbf{u}(k) \quad (2)$$

where $\mathbf{x}(k) = [\Delta\tilde{\rho}_1^a(k), \Delta\tilde{\rho}_1^b(k), \Delta\gamma_1(k), \dots, \Delta\tilde{\rho}_n^a(k), \Delta\tilde{\rho}_n^b(k), \Delta\gamma_n(k)]^T$ is the state vector and $\mathbf{u}(k) = \Delta\boldsymbol{\varepsilon}(k)$ is the control vector, whereby $\Delta\boldsymbol{\varepsilon}(k) = [\Delta\varepsilon_1(k), \dots, \Delta\varepsilon_n(k)]^T$. Also, $\Delta(\cdot)(k) = (\cdot)(k) - (\cdot)^N$, the superscript N denoting the nominal values, while it has been assumed that $\Delta(\cdot)(k) = 0$ for all disturbances (upstream mainstream inflows, as well as the on-ramp

flows, of each direction). $\hat{\mathbf{A}} \in \mathbb{R}^{3n \times 3n}$ and $\hat{\mathbf{B}} \in \mathbb{R}^{3n \times n}$ are the time-invariant state and input matrices, respectively, while $\mathbf{x} \in \mathbb{R}^{3n}$ and $\mathbf{u} \in \mathbb{R}^n$. If the control time-step is defined as a multiple of the model time-step, i.e. $T_c = MT$, where M is an integer, then the discrete control time index is $k_c = \lfloor kT/T_c \rfloor$. Thus, the linear state-space equation may be changed as follows, in order to be based on the control time-step T_c ,

$$\mathbf{x}(k_c + 1) = \mathbf{A}\mathbf{x}(k_c) + \mathbf{B}\mathbf{u}(k_c) \quad (3)$$

where $\mathbf{A} = \hat{\mathbf{A}}^M$, and $\mathbf{B} = (\hat{\mathbf{A}}^{M-1} + \hat{\mathbf{A}}^{M-2} + \dots + \mathbf{I})\hat{\mathbf{B}}$.

When employing the LQR methodology, as done by Malekzadeh (2021b), the control goal is the minimization of the quadratic criterion.

$$J = \frac{1}{2} \sum_{k_c=0}^{\infty} [\mathbf{x}^T(k_c)\mathbf{Q}\mathbf{x}(k_c) + \mathbf{u}^T(k_c)\mathbf{R}\mathbf{u}(k_c)] \quad (4)$$

where $\mathbf{Q} \in \mathbb{R}^{3n \times 3n}$ is a diagonal positive semidefinite matrix and $\mathbf{R} \in \mathbb{R}^{n \times n}$ is a diagonal positive definite matrix. The first term penalizes deviations of the state variables from zero, i.e. deviations of $\tilde{\rho}_i^a(k_c)$, $\tilde{\rho}_i^b(k_c)$, $\gamma_i(k_c)$, $i = 1, 2, \dots, n$, from their respective desired nominal values. The second term penalizes deviations of the control inputs from the nominal values.

The nominal value for relative densities on both directions is set equal to 1 so that the controller is motivated to operate the system near capacity, which is good for traffic efficiency. In particular, due to the quadratic penalty terms, the controller tends to mitigate strong density departures from the critical density at specific sections, i.e., mitigate traffic congestion. In addition, if capacity flow is not feasible (e.g. due to lack of demand), then minimizing a sum of squares has the tendency to balance deviations from the nominal values at different sections and directions, something that is conform with the secondary operational sub-objective of balancing the margin to capacity across sections and directions. On the other hand, the nominal value for the sharing factors is set to 0.5, so as to have smooth and moderate internal boundary changes. Thus, minimization of the second term in (4) mitigates deviations of the sharing factors from 0.5 and balances these deviations in space and time, which is a secondary operational sub-objective, as unnecessarily strong internal boundary changes over space and time should be avoided. The optimal controller minimizing criterion (4) subject to the model (3) is given by a linear state-feedback control law of the form $\mathbf{u}(k_c) = \mathbf{K}\mathbf{x}(k_c)$, where $\mathbf{K} \in \mathbb{R}^{n \times 3n}$ is a constant gain matrix given by

$$\mathbf{K} = (\mathbf{R} + \mathbf{B}^T\mathbf{P}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{P}\mathbf{A} \quad (5)$$

and \mathbf{P} is a unique positive semidefinite solution of the discrete-time algebraic Riccati equation.

Lane-Free Microscopic Simulation for IBC

Experimental evaluation is performed using TrafficFluid-Sim (Troullinos et al., 2021), a lane-free extension of the well-known microscopic simulator SUMO (Lopes et al., 2018). For IBC, additional functionalities had to be developed in order to facilitate the use of moving boundaries, some of which are outlined as part of future work in (Troullinos et al., 2022). Consider a network with multiple routes, e.g., a highway with on-ramps and off-ramps. Each vehicle entering the network is assigned a specific route. In conventional lane-based traffic, the vehicle would need to do appropriate lane-changing operations in order to follow its routing scheme, whereas in lane-free

environments, this would be translated to the vehicle operating according to left and right boundaries that guide its lateral placement appropriately, e.g., a vehicle entering from an on-ramp would need to merge appropriately in the main highway. Consequently, the boundaries are designed with the use of sigmoid functions that specify the lateral availability of the associated route and allow for smooth lateral movement.

In TrafficFluid-Sim, monitoring and control of vehicles is feasible through an API, which now includes additional functionality that provides density information according to the sectioning of the network, and controls the moving boundaries' lateral level in an online manner. As such, the user can now specify the sectioning of a bidirectional highway containing on-ramps and off-ramps, rendering the lane-free microscopic environment compatible with the proposed IBC problem formulation. Density information per section and direction of movement can be directly requested from the updated API for the calculation of relative densities $\tilde{\rho}_i^a(k)$, $\tilde{\rho}_i^b(k)$. The control vector stemming from LQR is parsed appropriately to the control function of the API, updating the lateral level of the left boundary (from the perspective of each direction) for each section. The left boundaries of each direction are essentially “synced” according to the control input, while the right boundaries corresponding to the exterior part of the road remain unaffected.

For practical purposes, the microscopic simulator includes a lateral distance margin between the two directions so that vehicles traveling near the boundary retain a safety distance from the other direction. Additionally, a moderate delay time is specified in every section at the direction that the road widens, so that vehicles travelling on the other direction have the necessary time to comply with the reduced lateral space. Figure 1 demonstrates a snapshot of the employed simulator. This snapshot includes an on-ramp and its acceleration area before merging with direction a as well as direction b that is separated from direction a using internal boundaries (red and green).



Fig. 1: A snapshot of the simulation environment TrafficFluid-Sim

3. RESULTS AND DISCUSSION

The hypothetical highway for the simulation is depicted in Fig. 2. Its length is 3 km and, when no control action is applied, the width for each direction is 10.2 m (corresponding to the width of 3 lanes). The highway is subdivided to 6 equal virtual sections in order to design and apply IBC. The vehicles are moving following the Ad-hoc strategy presented by Malekzadeh et al. (2022). The dimensions for each one of the vehicles are determined by choosing randomly (with uniform distribution) one of the six "dimension classes" reported in Table 1.

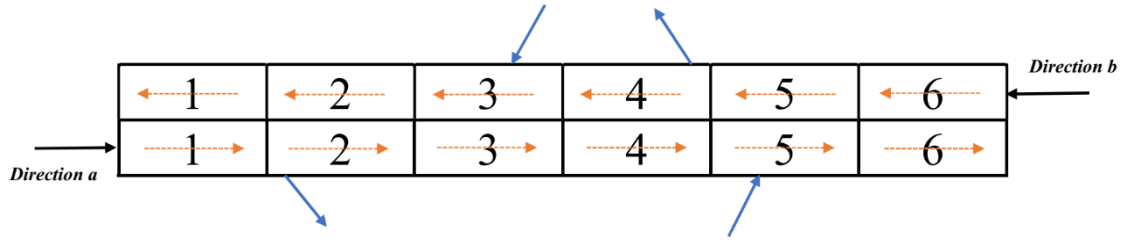


Fig. 2: The considered highway stretch

Table 1. The different dimension classes for the vehicles used in the simulation

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Length (m)	3.20	3.90	4.25	4.55	4.60	5.15
Width (m)	1.60	1.70	1.80	1.82	1.77	1.84

The nominal values for the control design are defined based on observations as $q_{cap}|_N = 28000$ veh/h and $\rho_{cr}|_N = 360$ veh/km, while the CTM time-step and the control time-step used are $T = 10$ s and $T_c = 60$ s, respectively. The weighing matrices used in the objective function are selected to be $\mathbf{Q} = \text{diag}(\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_n)$ where $\mathbf{S}_i = [\mathbf{I}_2, \mathbf{0}_{2 \times 1}; \mathbf{0}_{1 \times 3}]$ and $\mathbf{R} = 10^{-2} \mathbf{I}_{n \times n}$. The upper and lower bounds for the sharing factors, used to avoid utter blocking of any of the two directions, are equal for all sections $i = 1, 2, \dots, 6$ and are given the values $\varepsilon_{i, \min} = 0.1$ and $\varepsilon_{i, \max} = 0.9$. These values are used to truncate the LQR outcome.

The mainstream and on-ramp demand flows per direction are presented in Fig. 3 while the exit rate used for the off-ramps is on average 5%. If no control is used and the mainstream capacity is shared evenly between the two directions, the highway faces congestion. Figure 4 displays the corresponding spatio-temporal evolution of the relative density defined in (1). It can be observed that congestion is formed around time 23 min for direction *a* and time 53 min for direction *b* due to on-ramp merging. In both cases, congestion spills back and covers upstream sections.

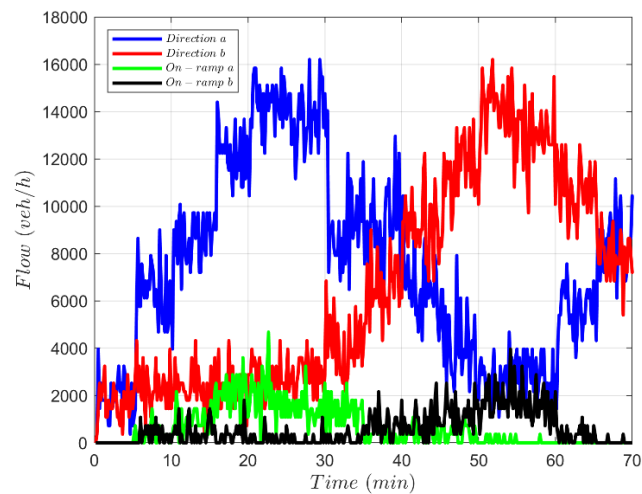


Fig. 3: Demand flows per direction and on-ramps

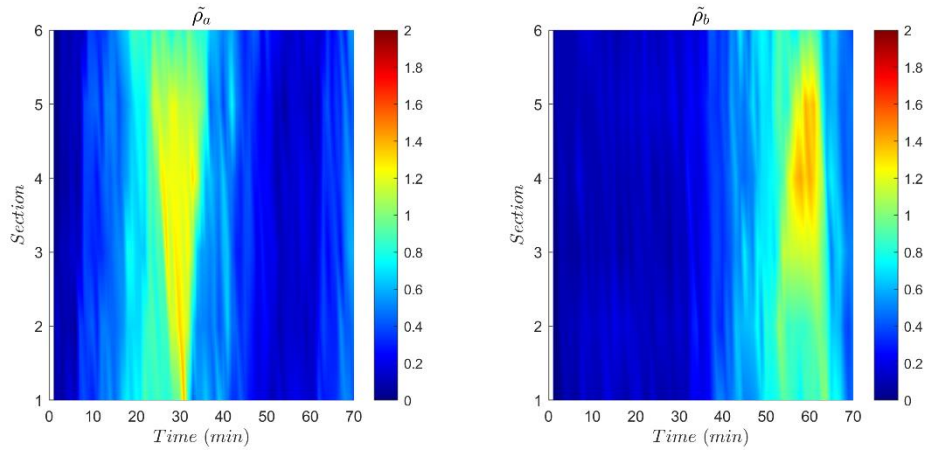


Fig. 4: Relative density for the two directions in the no-control case

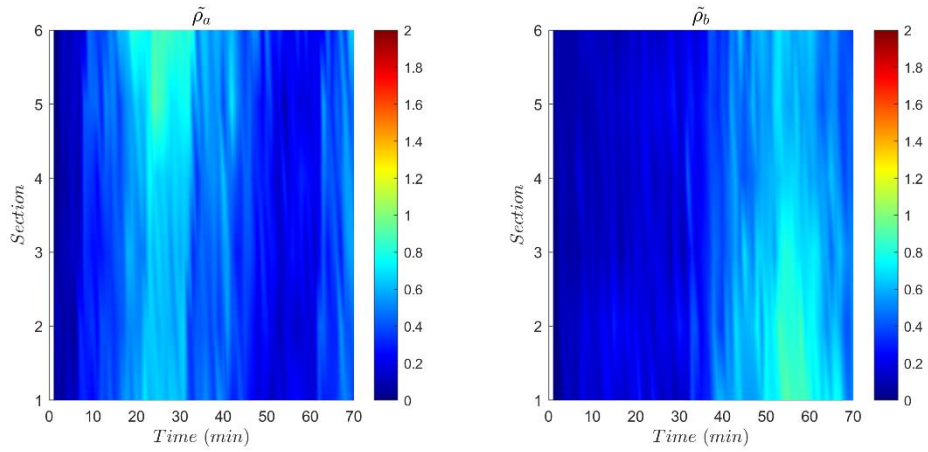


Fig. 5: Relative density for the two directions in the IBC case

On the other hand, congestion is solved utterly when IBC is activated. The corresponding spatio-temporal evolution of the relative density in the presence of IBC is presented in Fig. 5. It is evident that the relative densities in both directions are under 1. More insights can be gained by the trajectories for flow, relative density and the sharing factor per direction that are presented in Figs 6-8 over the whole simulation period. It can be seen that flow and relative density increase in section 5 for direction a due to the presence of inflow from the corresponding on-ramp. Likewise, this happens in section 3 for direction b . However, the controller is able to avoid the congestion by proper capacity sharing between the two directions.

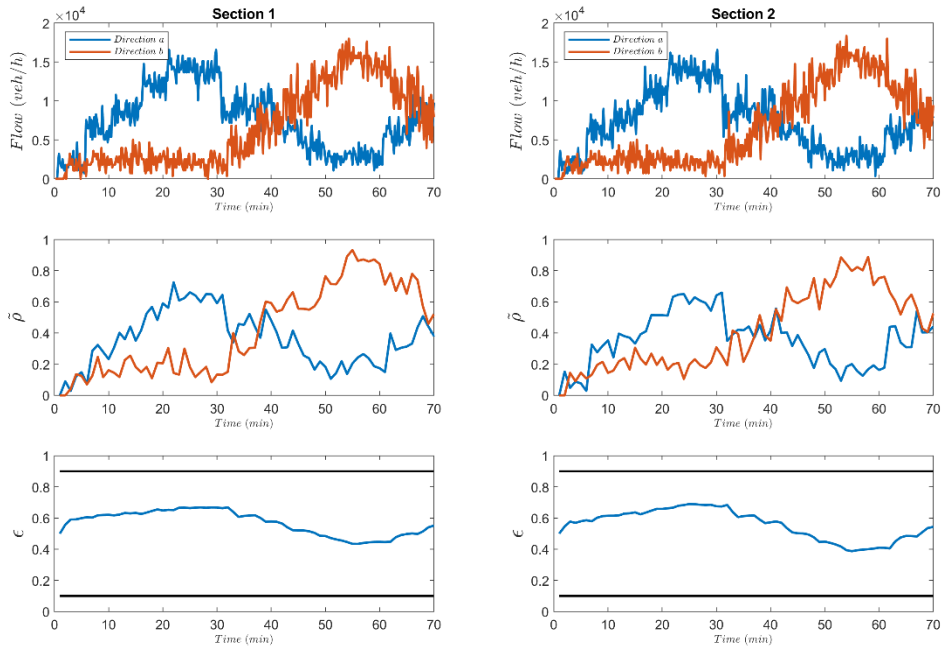


Fig. 6: Flow, relative density and control input for the IBC case (Section 1 & 2)

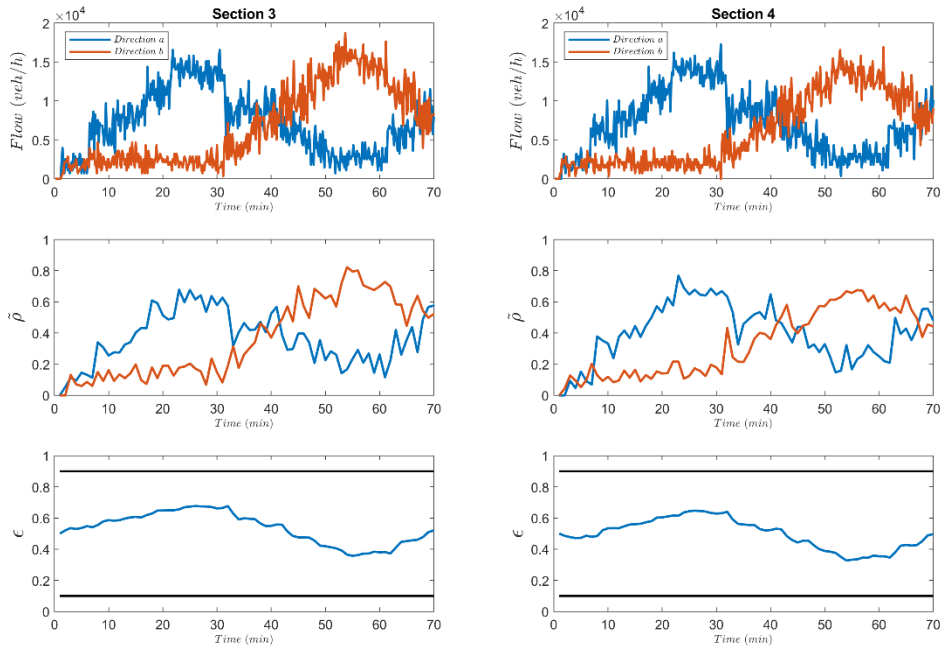


Fig. 7: Flow, relative density and control input for the IBC case (Section 3 & 4)



Fig. 8: Flow, relative density and control input for the IBC case (Section 5 & 6)

4. CONCLUSIONS

In this study, applicability of internal boundary control for lane-free traffic has been investigated via microscopic simulation. In order to implement the IBC scheme, the LQR controller has been employed. The controller was designed based on the well-known Cell Transmission Model. Then, a feedback control scheme that uses the relative densities in each section as input has been implemented. The simulation results validated the effectiveness of the proposed IBC scheme. Future work includes testing of the same and other control schemes, e.g. those developed by Malekzadeh et al. (2023), on longer highways.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 Research and Innovation programme/ ERC Grant Agreement n. [833915], project TrafficFluid.

REFERENCES

Daganzo, C.F. 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, Vol. 28, No. 4, pp. 269-287.

- Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P. and Wießner, E. 2018. Microscopic traffic simulation using sumo. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2575-2582.
- Malekzadeh, M., Yanumula, V.K., Papamichail, I. and Papageorgiou, M., 2023. Overlapping internal boundary control of lane-free automated vehicle traffic. *Control Engineering Practice*, Vol. 133, Article 105435.
- Malekzadeh, M., Manolis, D., Papamichail, I. and Papageorgiou, M. 2022. Empirical investigation of properties of lane-free automated vehicle traffic. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2393-2400.
- Malekzadeh, M., Papamichail, I., Papageorgiou, M. 2021b. Linear-Quadratic regulators for internal boundary control of lane-free automated vehicle traffic. *Control Engineering Practice*, Vol. 115, Article 104912.
- Malekzadeh, M., Papamichail, I., Papageorgiou, M., Bogenberger, K. 2021a. Optimal internal boundary control of lane-free automated vehicle traffic. *Transportation Research Part C: Emerging Technologies*, Vol. 126, Article 103060.
- Papageorgiou, M., Mountakis, K.S., Karafyllis, I., Papamichail, I., Wang, Y. 2021. Lane-free artificial-fluid concept for vehicular traffic. *Proceedings of the IEEE*, Vol. 109, No. 2, pp. 114-121.
- Troullinos, D., Chalkiadakis, G., Manolis, D., Papamichail, I. and Papageorgiou, M. 2021. Lane-free microscopic simulation for connected and automated vehicles. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 3292-3299.
- Troullinos, D., Chalkiadakis, G., Manolis, D., Papamichail, I. and Papageorgiou, M. 2022. Extending SUMO for Lane-Free Microscopic Simulation of Connected and Automated Vehicles. In *SUMO Conference Proceedings*, Vol. 3, pp. 95-103.

Changes in Car Ownership due to Life Events: Insights from the UK Longitudinal Household Survey

Cristian Domarchi*¹, Elisabetta Cherchi²

¹ PhD Student, School of Engineering, Newcastle University, United Kingdom

² Professor of Transport, School of Engineering, Newcastle University, United Kingdom

SHORT SUMMARY

The decision of how many vehicles should a household have –if any– is likely to depend on life events that change the transport requirements of the household, such as the birth of a child, the change of employment status, a significant income variation, or a child moving out of the household to start living in another one. In this paper, we model changes in car ownership level as a function of socioeconomic individual and household attributes, as well as significant life events using a large sample of UK households sourced from the Understanding Society survey. We estimate a discrete choice model with specific parameters for increasing or decreasing car ownership levels and considering panel and dynamic effects. Results show that life events play a significant role in predicting car ownership levels, and that households are relatively stable over time in terms of car holdings.

Keywords: Car ownership, Discrete choice models, Life events, Longitudinal survey, Panel data

1. INTRODUCTION

The decision of how many vehicles should a household have –if any– is likely to depend on socioeconomic, environmental, and accessibility attributes (de Jong *et al.*, 2004 and Anowar *et al.*, 2014), and also on life events that change transport requirements, such as the birth of a child, the change of employment status, a significant income variation, or a child moving out of the household to start living in another one. A large part of the car ownership literature has adopted a cross-sectional approach, analysing the number of vehicles in a household at a specific point in time. Despite its relevance, the key effect of life events has been much less frequently studied, likely because of the lack of suitable longitudinal datasets that allow following household decisions over time.

In this paper, we model changes in car ownership level as a function of socioeconomic individual and household attributes, as well as significant life events, using a large sample of UK households sourced from the Understanding Society survey (University of Essex *et al.*, 2020). This survey has previously been used in the transport context. Whittle *et al.* (2022) use this data to analyse transport mode frequency changes triggered by life events, focusing on individual behaviour. They do not study car ownership at the household level. Clark *et al.* (2016), on the other hand, model car ownership changes in UK households, but they only make use of the first two waves of the Understanding Society survey. Additionally, they study the effects of the two possible outcomes (increasing or decreasing car holdings) using independently estimated models, as opposed to a single specification, which allows for a proper representation of the dynamics. This approach

has surprisingly been followed in other studies in different contexts (Prillwitz *et al.*, 2006; Oakil *et al.*, 2014).

We take advantage of the richness of our dataset and estimate a single discrete choice model with specific parameters for increasing or decreasing car ownership levels. Furthermore, we estimate additional parameters to analyse potential differences due to the current car ownership level, as it would be expected that, for example, the increased utility drawn from buying an additional vehicle differs from what the household would get from purchasing their first car. To our knowledge, this is the first study to estimate a single discrete choice model that simultaneously accounts for all these effects.

We estimate an error component model with systematic heterogeneity in the preferences, which accounts for the panel effects, i.e., correlations due to the dataset containing repeated observations from the same unit of analysis. In addition, we investigate possible dynamic effects in consecutive choices, incorporating lagged choice variables as explanatory attributes in the utility functions, and correcting for possible endogeneity.

2. METHODOLOGY

We rely on data sourced from the Understanding Society survey, a UK-based household longitudinal survey that collects information about social, economic, and behavioural variables, including some transport behaviour questions (University of Essex *et al.*, 2020). The survey follows a large sample of individuals over time, with each observation point defined as a “wave”. We define respondents of wave 9 (2018) as the initial household set and work backwards to identify their corresponding household in every previous wave. Our processed dataset contains 10,067 households. The main household attributes remain relatively stable over time, with mean car ownership almost invariant (1.32 cars per HH in the last wave), and the average household size showing a slight 3.3% decrease between waves 1 and 8. Household attributes from wave 1 were used as baseline variables in the discrete choice models.

Inter-wave “life events” were identified by comparing each wave with the previous one. The main household-based inter-wave life events identifiable from the sample are:

- Car ownership changes.
- Household structure changes: Variations in the number of adults and children in the household.
- Residential relocation: Change of address between waves. Some relocations can be further identified as long-distance moves (switching regions), urban to rural moves, and rural to urban moves.
- Income changes between waves.
- Household splits: Household members appear in different households in the next wave.

Similarly, the main personal inter-wave life events are:

- Partner gains and losses.
- New-born child.
- Employment changes: Employment start, Employment exit, Employment switch, Retirement.
- Driving licence acquisition.

On average, 16.2% of households change their level of car ownership between any two waves. The most frequent life events are changes in employment (12.1%) and household composition

(11.3%), while residential relocations (4.3%), childbirths (2.3%) and partner gains and losses (2.1%) are scarcer.

To model the decision of changing the number of vehicles on each household we assumed that, during inter-wave period t , each household h faces three alternatives j :

- Keeping the number of vehicles in the household constant ($j = 0$).
- Increasing the number of vehicles in the household ($j = 1$).
- Decreasing the number of vehicles in the household ($j = 2$).

We defined the systematic utility that household h derives from choosing alternative j during period t as follows:

$$\begin{aligned}
 V_{0ht} &= \alpha_0 \\
 V_{1ht} &= \alpha_1 + \sum_{k=1}^K \beta_{1k} \cdot X_{kht} + \sum_{l=1}^L \theta_{1l} \cdot Y_{lh} \\
 V_{2ht} &= \alpha_2 + \sum_{k=1}^K \beta_{2k} \cdot X_{kht} + \sum_{l=1}^L \theta_{2l} \cdot Y_{lh}
 \end{aligned} \tag{1}$$

Here, X_{kht} is the value of attribute k for household h during period t . These K attributes are inter-wave changes. Analogously, Y_{lh} is the value of attribute l , one of the L baseline attributes that characterise the initial condition of household h in the dataset. Since, for each household and period, these attributes have the same value across all alternatives, we estimate $2(K + L)$ alternative-specific parameters β_{jk} and θ_{jl} , as well as the alternative-specific constants α_j . Sensitivities to the attributes might not be constant across the population and, furthermore, the current car ownership level might influence their relevance. We investigate this effect using systematic taste variations (Ortúzar and Willumsen, 2011) in the inter-wave parameters β_{jk} . When estimating these effects, we use households with one car as the reference level.

Next, to address the correlation between observations from the same household in different waves, we define the net utility of this model as:

$$U_{jht} = V_{jht} + \lambda_{jh} + \varepsilon_{jht} \tag{2}$$

Here, the error term has two components: ε_{jht} are i.i.d. extreme value (EV) type 1 error terms and $\lambda_{jh} = \sigma_j \eta_h$, known as the *panel effect*, varies across individuals but not across waves. We assume that the η_h terms are Normal (0,1) distributed error components for each household h , which capture the correlation between observations from the same household. σ_j are alternative-specific parameters to be estimated. The described model (**Model 1**) is an error component model with systematic heterogeneity in the preferences. Following Walker *et al.* (2007), the panel data structure allows estimating the three alternative-specific variances.

In a second specification (**Model 2**) we deal with the time sequence of household choices in a dynamic model. First, we assume that the choice at time t partly depends on choice at $t - 1$ only (dynamic process of order 1), that this dependence is household-specific (i.e., it only depends on previous choice of the same household), and that the weight of this dependence, ρ is the same for every household (Danalet *et al.*, 2016). The net utility of this model becomes:

$$U_{jht} = V_{jht} + \lambda_{jh} + \rho \gamma_{jh,t-1} + \varepsilon_{jht} \tag{3}$$

Here, $y_{jh,t-1}$ is a dummy variable that takes the value 1 when household h chooses alternative j at time $t - 1$. As explained by Woolridge (2005), this modelling structure introduces endogeneity due to correlation between the lagged variable $y_{jh,t-1}$ and the unobserved factors ε_{jht} . This phenomenon, called the initial conditions problem, must be corrected to obtain consistent parameters. Following the method proposed by Woolridge (2005) and implemented in Danalet *et al.* (2016), we model the panel effect term as follows:

$$\lambda_{jn} = \lambda_2 + \gamma \cdot y_{jh2} + \tau \cdot y_{jht}^{count} + \xi_{jn} \quad (4)$$

Here, y_{jh2} is the choice that household h makes in the first inter-wave period (for most households, $t = 2$, or the period after wave 1), while y_{jht}^{count} is the count of previous choices of alternative j up to time t (but not including the choice at time t). The inclusion of these terms addresses the endogeneity issue. Finally, λ_2 , γ and τ_m are coefficients to be estimated and ξ_{jn} is the panel effect term.

Since the initial choice is used in the panel effect and this specification considers the trajectory of consecutive choices by the household, we can only model choices from inter-wave period $t = 1$ onwards. In our case, this means that the decision of changing car ownership level between waves 1 and 2 is not modelled.

3. RESULTS AND DISCUSSION

For clarity, we split the modelling results in two tables. Table 1 presents the main life events coefficients for both estimated models and Table 2 details the goodness-of-fit indicators and the main validation results. For space reasons, we omit the baseline coefficients.

The main life event effects have the expected signs in both specifications. Increasing the number of adults in the household has a positive effect on the probability of acquiring an additional car, although this effect is more relevant for 1-car households than for carless households or those with 2 or more vehicles. Conversely, a higher number of adults tends to reduce the likelihood of discarding a car, with the effect being more significant in households with 2 or more cars. An increase in the number of kids in the household between two waves tends to increase the probability of buying an additional car and reduce the likelihood of discarding one.

As previously found in several studies, the effect of residential relocation in car holdings is mixed. The probability of buying an additional car increases after a house move for carless households and those with just one car and decreases for households with two or more cars. Car dependency has been shown to exist in rural areas where the access of alternatives to the private car is limited (e.g., Zhao and Bai, 2019; Carroll *et al.*, 2021), and our results show that moving from an urban to a rural setting significantly increases the probability of adding a vehicle to the household. Conversely, relocating from an urban to a rural area appears to have the opposite effect, making it more likely that the household discards one of their vehicles. A long-distance move positively influences a reduction in car holdings. As expected, a household split is significantly tied to a reduced car ownership level. This is likely explained by the splitting households sharing the vehicles that originally belonged to the “parent” household, or at least one of them becoming a carless household as a result of the split.

The effects of employment change depend strongly on the current ownership level in the household. Entering employment has a clear and significant positive effect in the likelihood of buying an additional car only for carless households and those with 1 car, but the effect is not significant for households with 2 or more vehicles. Similarly, an employment switch seems to both increase the likelihood of buying an additional car *and* reduce the probability of discarding one; however, the effect is opposite for households with 2 or more cars. On the other hand, as expected, job losses and retirement seem to be positively correlated with car disposals.

Table 1. Model results I – Life events coefficients

Attribute	Unit	Model 1: Panel effect, static				Model 2: Panel effect, dynamic			
		1: Increase		2: Decrease		1: Increase		2: Decrease	
		Coef.	t-test (0)	Coef.	t-test (0)	Coef.	t-test (0)	Coef.	t-test (0)
Alternative specific constant	–	-3.541	-22.90	-1.866	-12.17	-4.297	-31.73	-2.352	-13.68
<i>Household size</i>									
Adult number increase	<i>Reference</i>	0.992	9.88	-0.691	-8.02	0.983	8.45	-0.626	-6.38
	<i>0 car HH</i>	-0.613	-3.71	–	–	-0.521	-2.93	–	–
	<i>2+ car HH</i>	-0.616	-4.73	-1.326	-11.76	-0.581	-3.66	-1.366	-10.50
Children number increase	<i>Reference</i>	0.447	5.05	-0.540	-5.78	0.563	5.77	-0.489	-4.51
	<i>2+ car HH</i>	-0.310	-2.65	–	–	-0.305	-2.17	–	–
<i>Residential relocation</i>									
Residential relocation	<i>Reference</i>	0.503	3.76	–	–	0.303	1.96	–	–
	<i>2+ car HH</i>	-0.973	-4.14	–	–	-0.719	-2.65	–	–
Urban to rural move	–	0.947	4.42	–	–	1.321	5.41	–	–
Rural to urban move	–	–	–	1.052	4.01	–	–	0.851	2.67
Long distance move	–	0.271	1.41	0.705	3.64	0.444	2.00	0.865	3.85
Household split	–	–	–	0.389	1.60	–	–	0.398	1.38
<i>Personal life</i>									
Partner gain	<i>Reference</i>	1.744	12.01	–	–	1.926	11.02	–	–
	<i>2+ car HH</i>	-1.720	-5.55	–	–	-2.044	-5.29	–	–
Partner loss	–	–	–	0.507	5.95	–	–	0.613	6.16
<i>Employment status</i>									
Enter employment	<i>Reference</i>	0.643	6.71	–	–	0.590	5.35	–	–
	<i>2+ car HH</i>	-0.932	-5.96	–	–	-0.721	-3.96	–	–
Exit employment	–	–	–	0.392	4.05	–	–	0.334	3.01
Retired	–	–	–	0.330	3.20	–	–	0.371	3.25
Switch employment	<i>Reference</i>	0.627	7.26	-1.247	-5.72	0.524	5.07	-1.323	-5.02
	<i>2+ car HH</i>	-0.751	-6.10	1.597	6.90	-0.564	-3.78	1.634	5.87
<i>Transport</i>									
Licence acquisition	<i>Reference</i>	1.093	6.42	–	–	1.284	5.46	–	–
	<i>0 car HH</i>	1.316	4.80	–	–	1.217	3.31	–	–
<i>Income level</i>									
Income increase (x £1,000)	–	0.101	8.63	–	–	0.078	8.38	–	–
Income decrease (x £1,000)	–	–	–	0.046	5.12	–	–	0.055	5.53

Acquiring a driving licence has been widely acknowledged as a major transport milestone, and an important predictor of car ownership increase (Clark *et al.*, 2014; Clark *et al.*, 2016; Rau and Manton, 2016). Our models show that this effect is significant for all households, but more than doubled for those without a car. Partner loss is associated with a decrease in car ownership and, conversely, gaining a partner contributes to increase the probability of acquiring an additional car, but only in households with less than 2 vehicles. Finally, family income changes have a significant effect on car ownership variations, but the effect is not symmetric. We found that an increase in income influences the likelihood of buying an additional car, while a reduction in the same amount has a *lower* effect on the probability of discarding a vehicle.

Table 2. Model results II – Panel/dynamic coefficients, model fit, and validation

Attribute	Model 1		Model 2	
	Coef.	T-Test	Coef.	T-Test
<i>Panel and dynamic coefficients</i>				
Panel effect – Base alternative	1.062	41.69	1.576	41.04
Panel effect – Increase	-0.012	-1.82	-0.028	-1.03
Panel effect – Decrease	-0.003	-0.54	-0.018	-1.12
Previous choice (ρ)	–	–	-0.895	-13.46
First choice (γ)	–	–	-0.204	-3.88
Choice frequency (τ)	–	–	-0.708	-15.78
<i>Model fit</i>				
Number of households	7,080	–	7,051	–
Number of observations	48,930	–	41,838	–
Log-likelihood (0)	-50,380	–	-43,063	–
Log-likelihood (k)	-26,183	–	-22,044	–
Log-likelihood (*)	-22,001	–	-18,385	–
Adjusted rho index (0)	0.562	–	0.571	–
Adjusted rho index (k)	0.157	–	0.163	–

Table 2 shows that there is a significant correlation between observations by the same household, as shown by the significant variances of the error term for the base alternative in both models. However, only Model 2 considers the trajectory of household choices over time (e.g., the dynamic effect). The three dynamic parameters are significant and have the expected signs as they all point to a stability of car ownership levels. In particular, the probability of purchasing an additional car in wave t is strongly *reduced* if the household already bought an additional vehicle in time $t - 1$. Similarly, the household is much less likely to discard a car if they already did so in the previous wave. Households are also less likely to change their car ownership status if they have done so in the past, which is the main reason why both the choice frequency and the first-choice parameters are negative and statistically significant in explaining choice over time. The dynamic effects appear to reflect the fact that the number of car households tends to be very stable over time, that changes are mostly induced by significant life events, and that they are unlikely to be repeated.

4. CONCLUSIONS

Using a highly detailed panel dataset from a nationally representative longitudinal survey, which followed a large sample over a period of 9 years, we estimated choice models to explain changes in car holdings in UK households. The results show that important life events related to household size and structure, employment status, income level changes, transport milestones, and other personal events, can help understanding the decision to increase or decrease vehicle holdings in households.

Our models account for correlation between observations from the same household over time, and our preferred specification (Model 2), also considers dynamic effects. All these are significant in explaining car ownership level changes, and their sign confirms that households are relatively stable over time in terms of car holdings, and they are unlikely to change their car ownership status if they have previously done so. Although both modelling frameworks allow obtaining robust parameters to understand car ownership level change, the dynamic specification shows a slightly improved explanatory power.

It could be argued that life events can also trigger vehicle replacements, which involve substitutions of fuel type, vehicle segment, and even make and model. In addition, a specific life event can also disrupt car use. However, the available dataset is not transport-specific and does not include these variables. In addition, the modelled events are likely not the only causes of changes in car ownership levels. Health-related issues, personal circumstances, and school relocations are examples of variables that cannot be sourced from the sample. Similarly, the decision of buying a car might also be influenced by personal beliefs, attitudes, and social norms, aspects that are absent from the survey.

Our results confirm that plans directed at tackling the increase of the number of cars must not only consider economic measures, but also how to provide access to transport alternatives. The effect of residential location is telling in this respect, as the provision of better transport links might reduce the need of buying a car when moving to a rural setting. Increases in car ownership should not be considered as the only response to changes in the life cycle, or an inevitable consequence of economic growth. The focus should be on offering more sustainable transport opportunities that allow satisfying these needs without relying on additional vehicles, especially considering their adverse environmental and social effects.

ACKNOWLEDGEMENTS

This work was supported by a Leverhulme Trust Doctoral Scholarship in Behaviour Informatics and the multimodal study of behaviour, and by the Newcastle University Overseas Scholarship (NUORS).

REFERENCES

- Anowar, S., Eluru, N. and Miranda-Moreno, L.F. 2014. Alternative modeling approaches used for examining automobile ownership: A comprehensive review. *Transport Reviews*, Vol. 4, No. 4, pp. 441-473.
- Carroll, P., Benevenuto, R. and Caulfield, B. 2021. Identifying hotspots of transport disadvantage and car dependency in rural Ireland. *Transport Policy*, Vol. 101, pp. 46-56.
- Clark, B., Chatterjee, K. and Melia, S. 2016. Changes in level of household car ownership: the role of life events and spatial context. *Transportation*, Vol. 43, pp. 565-599.
- Clark, B., Chatterjee, K., Melia, S., Knies, G. and Laurie, H. 2014. Life events and travel behavior: Exploring the interrelationship using UK household longitudinal study data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2413, pp. 54-64.
- Danalet, A., Tinguelli, L., de Lapparent, M. and Bierlaire, M. 2016. Location choice with longitudinal WiFi data. *Journal of Choice Modelling*, Vol. 18, pp. 1-17.
- de Jong, G.C., Fox, J., Daly, A., Pieters, M. and Smit, R. 2004. Comparison of car ownership models. *Transport Reviews*, Vol. 24, No. 4, pp. 379-408.
- Oakil, A.T.M., Ettema, D., Arentze, T. and Timmermans, H. 2014. Changing household car ownership level and life cycle events: an action in anticipation or an action on occurrence. *Transportation*, Vol. 41, pp. 889-904.
- Ortúzar, J.de D. and Willumsen, L. 2011. *Modelling Transport*. 4th edn. Chichester, UK: Wiley and Sons.

Prillwitz, J., Harms, S. and Lanzendorf, M. 2006. Impact of life-course events on car ownership, *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1985, pp. 71-77.

Rau, H. and Manton, R. 2016. Life events and mobility milestones: Advances in mobility biography theory and research. *Journal of Transport Geography*, Vol. 52, pp. 51-60.

University of Essex, Institute for Social and Economic Research, NatCen Social Research and Kantar Public 2020. *Understanding Society: Waves 1-10, 2009-2019 and Harmonised BHPS: Waves 1-18, 1991-2009* [data collection], 13th edition, UK Data Service.

Walker, J.L., Ben-Akiva, M. and Bolduc, D. 2007. Identification of parameters in Normal error component logit-mixture (NECLM) models. *Journal of Applied Econometrics*, Vol. 22, pp. 1095-1125.

Whittle, C., Whitmarsh, L., Nash, N. and Poortinga, W. 2022. Life events and their association with changes in the frequency of transport use in a large UK sample. *Travel Behaviour and Society*, Vol. 28, pp. 273-287.

Woolridge, J.M. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, Vol. 20, pp. 39-54.

Zhao, P. and Bai, Y. 2019. The gap between and determinants of growth in car ownership in urban and rural areas of China: A longitudinal data case study, *Journal of Transport Geography*, Vol. 79, p. 102487.

Providing a Revenue-forecasting Scheme to Relocate Groups of Ride-Sourcing Drivers

Caio Vitor Beojone* and Nikolas Geroliminis

Urban Transport Systems Laboratory (LUTS), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, CH-1015, Switzerland

SHORT SUMMARY

Proper positioning of ride-sourcing drivers may improve vacant travel times, waiting times, and matching opportunities. Herein, we evaluate the potential repositioning response of drivers when provided a guidance based on estimates of their earnings in a system offering ride-hailing (solo) and ridesplitting (shared) rides. We develop a strategy that enumerates the best regional repositioning destination based on the expected number of requests. A mixed continuous-discrete time Markov Chain (MDCTMC) is developed to predict drivers activities and the revenues associated with them. Our main findings indicate that the proposed approach is likely to retain drivers confidence by improving their earnings compared to other drivers. We also show that it manages to decrease the number of unserved requests compared to several state-of-art benchmarks and decreased the deadheading.

Keywords: Macroscopic Fundamental Diagram; Markov chain; Shared mobility; Urban mobility.

1 INTRODUCTION

In a daily basis, geographical variations on the demand can create an imbalance between the ride-sourcing service demand and supply of drivers to serve it, requiring actions to maintain a satisfactory service quality. However, the fleets of this service are formed by drivers that are free to make a series of decisions.

Therefore, the operator requires persuasion to relocate the available pool of drivers. Sadeghi & Smith (2019), and Powell et al. (2011) incentivized drivers to decide on the location for the next assignment. However, these strategies suffer from a reactive nature, mainly accounting for past events. For instance, if an area faces recurrent losses of requests, customers will likely change their travel option.

Other strategies emerge from the optimization of passenger-driver matching algorithms. These strategies include Alonso-Mora et al. (2017) who sent empty vehicles to the location of recently unsatisfied customers. Other examples can be found in Wang & Yang (2019) and references therein. More recent studies, try to take actions before losing these passengers. Zhu et al. (2022) uses coverage control to proactively position idle drivers. Although proactive, these approaches assume full compliance, ignoring the individual objectives of human drivers.

From all the above, multiple challenges arise when positioning ride-sourcing currently available drivers. The first challenge is to take the burden of identifying the most profitable options from drivers with limited information, which only have access to limited information while accounting for drivers' future activities. A challenge remains in giving positions that minimizes unnecessary overlapping in demand coverage among drivers. Finally, the strategy must ensure that compliant drivers have an improved outcome to ensure compliance.

Herein, we evaluate the potential repositioning response of drivers when provided an estimate of their earnings. The operator uses a mixed discrete-continuous time Markov chain (MDCTMC) to estimate individual earnings for a given decision in the short-term. A microscopic process identifies the positions and paths with the highest chances of matching. In a simulated study, we compared the performance of guided drivers and unguided ones. We show that guided drivers have increased revenues and are likely to follow the provided guidance in the long term. Finally, the proposed strategy is compared to state-of-art strategies, achieving superior results in terms of lost requests and deadheading.

2 IDENTIFYING THE BEST REGIONAL REPOSITIONING DECISIONS

Assume that the objective of the repositioning strategy is to maximize the driver's revenues by maximizing the number of requests served during a short prediction horizon τ . Therefore, the driver moves from the current position i to the potential position j for a significant portion of the prediction. Then, we can summarize the expected number of requests a driver can cover $P_{ij}(t)$ in this period as shown in Equations [1] and [2]. Finally, Equation [3] indicates the chosen path for the driver.

$$P_{ij}(t) = \sum_{o \in \mathcal{R}} \sum_{d \in \mathcal{R}} P_{ij}^{od}(t) \quad (1)$$

$$P_{ij}^{od}(t) = \int_t^{t+\tau} p_{ij}^{od}(s) ds \quad (2)$$

$$\arg \max_j P_i = P_{ij}(t) \quad (3)$$

Where $P_{ij}^{od}(t)$ and $p_{ij}^{od}(t)$ represent the number of requests and the demand rate with a regional OD-pair od covered in the path between driver's current location i and potential repositioning destination j , respectively.

For the computation of the instantaneous covered demand rates $p_{ij}^{od}(t)$ we consider: (i) geographical distribution of demand, (ii) the associated path, and (iii) other drivers' demand coverage. Figure 1 depicts these elements.

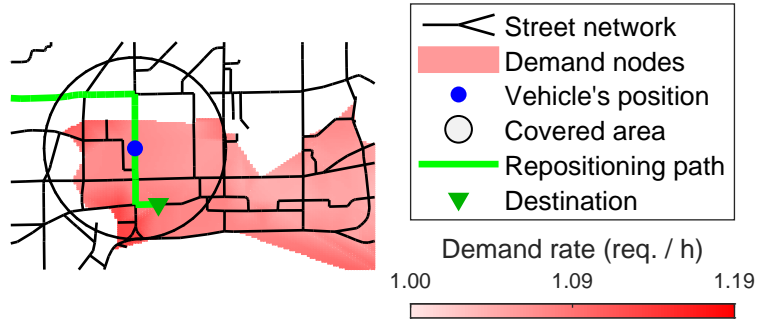


Figure 1: Illustration of the computation of $p_{ij}^{od}(t)$.

3 PROVIDING DRIVERS WITH REPOSITIONING GUIDANCE BASED ON REVENUE FORECASTING

While drivers decide about repositioning by themselves, they have incomplete information about the system conditions. Therefore, it is more likely that the service provider, who concentrates information about all drivers and demand can use the mobile application to supply drivers with repositioning information. The idea is to present the driver with the expected best decisions whenever he becomes available for relocation.

Predicting drivers' activities

Since ride-sourcing drivers offer rides for a profit, one should expect their decisions aiming to maximize it. Therefore, to convince drivers about the best decisions they should be presented with the consequence of their choices on their earnings.

To predict the earnings, it must identify the driver's activities and actions in the near-future. Inspired by Beojone & Geroliminis (2022), consider a list of activities \mathbb{A} , such that $A \in \mathbb{A}$ indicates a driver's current activity. A set \mathcal{R} with R heterogeneous regions, i.e., $\mathcal{R} = \{1, 2, \dots, R\}$ illustrate the urban network area, while the pair $od \in \mathcal{R}^2$ depicts a driver's current and destination regions. Therefore, $A_{od} \in \mathcal{K}$ describes a driver current state in the set of all possible states. Setting the list of activities $\mathbb{A} = \{I, RH, S1, S2\}$ assumes that a driver can execute the following four activities completing the state-space with a size $|\mathcal{K}| = |\mathbb{A}| \cdot |\mathcal{R}|^2$.

- Vacant (I): a vacant driver available for passengers.
- Ride-hailing (RH): a busy driver assigned to a ride-hailing passenger.
- Single ridesplitting ($S1$): a driver assigned to a single ridesplitting passenger.
- Shared ridesplitting ($S2$): a driver assigned to two ridesplitting passengers.

Assuming that the time spent in each activity is random, a Markov chain can depict a driver's movements. If the operator has detailed information on the urban area, then it can accurately tell the time a driver needs to reach different areas in their repositioning activities.

Based on the previous, we can break a driver's activity predictions into three phases. In the first phase (a), a continuous time Markov chain (CTMC) represents the movements of a driver before reaching the subsequent region in the repositioning path. In the second phase (b), a discrete time Markov chain (DTMC) represents the driver reaching the boundary of the current region. Lastly, in the third phase (c), after a driver reached the destination region, he is once again free to get any assignments in the network. Therefore, obtaining a construct referred as a mixed discrete-continuous time Markov chain (MDCTMC) model (Ingolfsson, 2005).

Repositioning movements: phase (a)

The provided shortest path is summarized in the sequence of regions $\mathbf{r} = (r_1, r_2, \dots, r_n)$ the driver will cover on the movement between regions o and d ($r_1 = o$ and $r_n = d$). Therefore, the transitions in phase (a) include those from the starting region o until the current region l . Equation [4] further details the dynamics for state probabilities $\pi_{ij}^K(t)$.

$$\dot{\pi}_{ij}^I(t) = -\pi_{ij}^I(t) \sum_{s \in \mathcal{S}} \lambda_{ij}^s(t) \quad (4a)$$

$$\dot{\pi}_{ij}^{RH}(t) = -\pi_{ij}^{RH}(t) \mu_{ij}^{RH}(t) + \pi_{ij}^I(t) \lambda_{ij}^H(t) \quad (4b)$$

$$\dot{\pi}_{ij}^{S1}(t) = -\pi_{ij}^{S1}(t) \left(\mu_{ij}^{S1}(t) + \sum_{h \in \mathbf{r}} \beta_{ij}^h \lambda_{ih}^S(t) \right) + \pi_{ij}^I(t) \lambda_{ij}^{S1}(t) + \pi_{ij}^{S2}(t) \vartheta_{ij}(t) \mu_{ij}^{S2} \quad (4c)$$

$$\dot{\pi}_{ij}^{S2}(t) = -\pi_{ij}^{S2}(t) \mu_{ij}^{S2}(t) + \sum_{h \in \mathcal{R}} \hat{\lambda}_{ihj}^{S2}(t) \quad (4d)$$

Where the base transitions can be described through the arrival process $\lambda_{od}^s(t)$ for a service $s \in \mathcal{S} = \{H, S\}$ (ride-hailing and ridesplitting), and the service process $\mu_{od}^K(t)$, in which driver completes a ride or transfers to a neighboring region. Note that any $\mu_{ij}^K(t)$ and $\lambda_{ij}^s(t)$ is only defined for $j = d$ and $i \in r_1, \dots, r_{n-1}$, otherwise they have a value of 0. β_{od}^h represents the ratio of od trips that will pass through region h . ϑ_{ij} indicates the probability of the driver in $S2_{ij}$ having a passenger to deliver in i before proceeding to j with the other one.

Repositioning boundary: phase (b)

The DTMC of phase (b) is depicted in Equation [5], which is supported by Equations [6] and [7] detailing the transition matrix. Note that the only change from the end of phase (a) occurs from state I_{id} to I_{ld} , whereas the remaining states keep the same probabilities. Since the travel time to reach region l is assumed deterministic, the transition occurs with certainty.

$$\pi(t^+) = B(i, d, l) \pi(t^-) \quad (5)$$

$$B(i, d, l) = [b_{ij}^K(i, d, l)] \in \mathbb{B}^{|\mathcal{K}| \times |\mathcal{K}|} \quad (6)$$

$$b_{ij}^K(i, d, l) = \begin{cases} 1, & \text{for } K_{ij} = I_{ld}, RH_{id}, S1_{id}, S2_{id}, S2_{ih} \\ 0, & \text{otherwise} \end{cases}, \quad (7)$$

Where t^+ and t^- refer to the instant right after and right before time t . $B(i, d, l)$ is the transition matrix representing the DTMC. \mathcal{K} is the state space of the model, and $|\mathcal{K}|$ is its cardinality.

After repositioning: phase (c)

Once the driver reaches the area intended in the repositioning decision, it becomes free to answer any requests and a CTMC (different from phase(a)) depicts its activities. Figure 2 illustrates the state space of a single region and the possible transitions. In the figure, regions ‘ k ’ and ‘ h ’ can be a set of regions immediately before or after region ‘ o ,’ respectively.

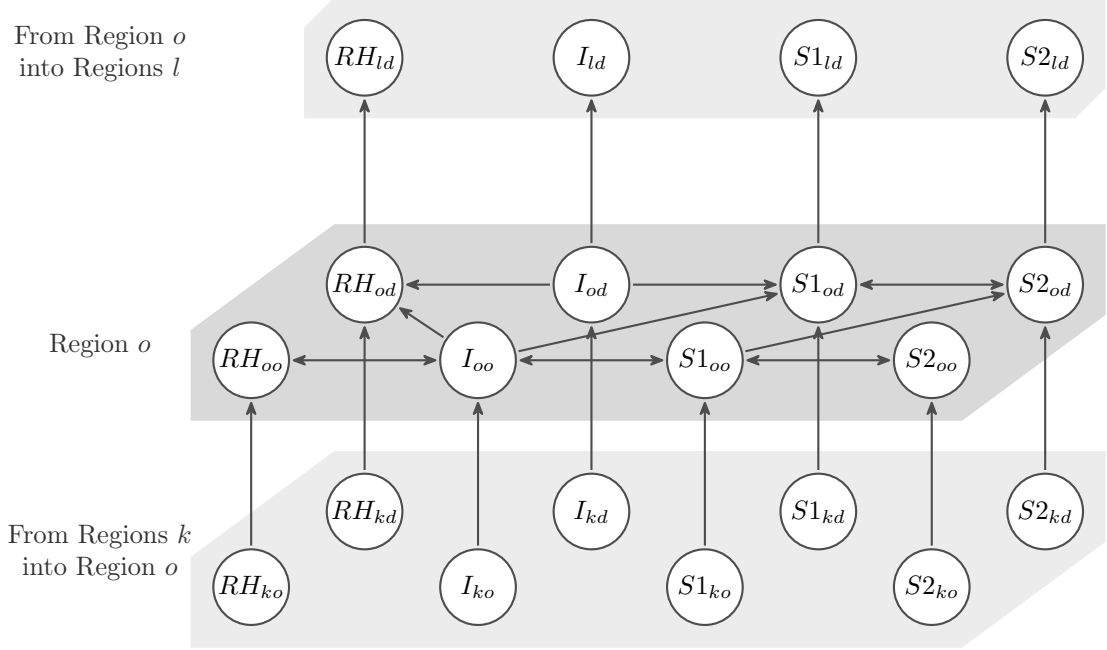


Figure 2: General state transition structure focusing on a region o and the inflows and outflows related to this region.

Therefore, we can describe the CTMC with the respective Equation 8, where we estimate the state probability $\pi_{od}^K(t)$. Table 1 provides the entries for each state in the CTMC.

$$\dot{\pi}_{od}^K(t) = - \text{Exits} + \text{Entrances} \quad (8)$$

State	Exits	Entrances
I_{oo}	$\pi_{od}^I(t) \sum_{s \in \mathcal{S}} \sum_{h \in \mathcal{R}} \lambda_{oh}^s(t)$	$\pi_{oo}^{RH} \mu_{oo}^{RH}(t) + \pi_{oo}^{S1}(t) \mu_{oo}^{S1}(t)$
I_{od}	$\pi_{od}^I(t) \sum_{s \in \mathcal{S}} \lambda_{od}^s(t)$	$\sum_{h \in \mathcal{R}} \hat{\mu}_{hd}^I(t)$
RH_{od}	$\pi_{od}^{RH}(t) \mu_{od}^{RH}(t)$	$\sum_{h \in \mathcal{R}} \pi_{oh}^I(t) \lambda_{od}^H(t) + \sum_{h \in \mathcal{R}_o} \hat{\mu}_{hod}^{RH}(t)$
$S1_{od}$	$\pi_{od}^{S1}(t) \mu_{od}^{S1}(t) + \hat{\lambda}_{ohd}^{S2}(t) + \hat{\lambda}_{odh}^{S2}(t)$	$\sum_{h \in \mathcal{R}} \pi_{oh}^I(t) \lambda_{od}^S(t) + \sum_{h \in \mathcal{R}_o} \hat{\mu}_{hod}^{S1}(t) + \hat{\mu}_{od}^{S2}(t)$
$S2_{od}$	$\pi_{od}^{S2}(t) \mu_{od}^{S2}(t)$	$\sum_{h \in \mathcal{R}} \hat{\lambda}_{ohd}^{S2}(t) + \sum_{h \in \mathcal{R}_o} \hat{\mu}_{hod}^{S2}(t)$

Table 1: Summary of state transitions in the Markov Chain model.

We have to detail the entries of the coefficients for ‘Exits’ and ‘Entrances,’ in Equation [8]. To shorten the description in Table 1, we aggregated some particular transitions explained in Equations [9]–[13]. In particular, Equation [9] illustrates that a driver can use different paths on his way to the destination. Equations [10] and [11] illustrate that shared ridesplitting drivers might complete a ride (Equation [10]) before transferring (Equation [11]). Finally, Equations [12] and [13] illustrate

that new shared ridesplitting rides can have different delivery order, such as a *last-in-first-out* (LIFO) order (Equation [12]) or a *first-in-first-out* (FIFO) order (Equation [13]).

$$\hat{\mu}_{hod}^K(t) = \pi_{hd}^K(t) \cdot \theta_{hod} \cdot \mu_{hd}^K(t) \quad K \neq S2 \quad (9)$$

$$\hat{\mu}_{od}^{S2}(t) = \pi_{od}^{S2}(t) \cdot \vartheta_{ood}(t) \cdot \mu_{od}^{S2}(t) \quad (10)$$

$$\hat{\mu}_{hod}^{S2}(t) = \pi_{hd}^{S2}(t) \cdot (1 - \vartheta_{ood}(t)) \cdot \theta_{hod} \cdot \mu_{hd}^{S2}(t) \quad (11)$$

$$\hat{\lambda}_{ohd}^{S2}(t) = \pi_{od}^{S1}(t) \cdot \beta_{od}^h \cdot \lambda_{oh}^S(t) \quad (12)$$

$$\hat{\lambda}_{odh}^{S2}(t) = \pi_{od}^{S1}(t) \cdot \beta_{oh}^d \cdot \lambda_{oh}^S(t) \quad h \neq d \quad (13)$$

Where, $\theta_{hod} \in [0, 1]$ distributes transfer flows over its neighboring regions such that the equality $\sum_{h \in \mathcal{R}_o} \theta_{ohd} = 1$ holds; ϑ_{ood} becomes the fraction of shared trips passing through o that will deliver a passenger before continuing to d ; and β_{od}^h (β_{oh}^d) represents the ratio of od (oh) trips that will pass through region h (d).

Note that there are two forms of transition processes, generally represented by $\lambda_{od}^s(t)$ and $\mu_{od}^K(t)$ for the passenger assignment rate and trip completion/transfer rates, respectively. These are translations of the macroscopic transitions in Beojone & Geroliminis (2022) to the individual level.

Estimating drivers' expected revenues

Drivers' earnings come from the fares passengers pay when booking rides, which are composed by a couple of elements. Firstly, there is a fixed booking fee $f_{od}^{s,B}$ relative to the reservation of a ride. Secondly, there is a travel fee relative to the trip distance $f_{od}^{s,T}$. The platform keeps a commission κ for this fare and returns to the drivers the remaining part.

From the MDCTMC, one can approximate the revenue generation by means of a continuous rate. Therefore, a driver receives a part of the fees proportionally to the number of received assignments during the evaluation and to the kilometers traveled in busy states. Equation [14] summarizes the expected revenue after commission κ . However, it requires us to break the gross revenue into its minor components throughout Equations [15] – [21].

$$E[R^{\text{net}}] = (1 - \kappa)E[R] \quad (14)$$

$$E[R] = E[R^B + R^T] = E[R^B] + E[R^T] \quad (15)$$

$$E[R^B] = E \left[\sum_s R^{s,B} \right] = \sum_s E [R^{s,B}] \quad (16)$$

$$E[R^{s,B}] = E \left[\sum_{od \in \mathcal{R}} R_{od}^{s,B} \right] = \sum_{od \in \mathcal{R}^2} E [R_{od}^{s,B}] \quad (17)$$

$$E[R_{od}^{s,B}] = E \left[f_{od}^{s,B} n_{od}^{s,B} \right] = f_{od}^{s,B} E [n_{od}^{s,B}] \quad (18)$$

$$E[R^T] = E \left[\sum_s R^{s,T} \right] = \sum_s E [R^{s,T}] \quad (19)$$

$$E[R^{s,T}] = E \left[\sum_{od \in \mathcal{R}^2} R_{od}^{s,T} \right] = \sum_{od \in \mathcal{R}^2} E [R_{od}^{s,T}] \quad (20)$$

$$E[R_{od}^{s,T}] = E \left[f_{od}^{s,T} d_{od}^{s,T} \right] = f_{od}^{s,T} E [d_{od}^{s,T}] \quad (21)$$

Where $n_{od}^{s,B}$ is the number of booked rides for service s from region o to region d ; and $d_{od}^{s,T}$ is the passenger-distance traveled for a service s from region o to region d .

Finally, Equations [22] and [23] estimate the remaining expected number of booked rides and passenger-distance travelled based on the instantaneous probabilities from the MDCTMC. In summary, these estimates are functions of a starting time t_0 and an evaluation period τ for a particular choice γ .

$$E[n_{od}^{s,B}(t_0, \tau|\gamma)] = \int_{t_0}^{t_0+\tau} \lambda_{od}^s(t) \sum_{K \in \mathcal{K}_B^s} \pi_{od}^K(t) dt \quad \mathcal{K}_B^s = \begin{cases} \{I, RP\}, & \text{if } s = H, \\ \{I, RP, S1\}, & \text{if } s = S \end{cases} \quad (22)$$

$$E[d_{od}^{s,T}(t_0, \tau|\gamma)] = \int_{t_0}^{t_0+\tau} v_o(t) \sum_{K \in \mathcal{K}_T^s} n_p^K \pi_{od}^K(t) dt \quad \mathcal{K}_T^s = \begin{cases} \{RH\}, & \text{if } s = H, \\ \{S1, S2\}, & \text{if } s = S \end{cases} \quad (23)$$

Where n_p^K is the number of assigned passengers to a driver in activity K . For states RH and $S1$, $n_p^K = 1$, while for $S2$ have $n_p^{S2} = 2$.

4 COMPUTATIONAL RESULTS

In this prototype application, we represented the central business district of Shenzhen, including parts of the Luohu and Futian Districts. The considered network consists of 1'858 intersections connected by 2'013 road segments. The experiment used a simulator based on Bejone & Geroliminis (2021, 2022), using Floy-Warshall algorithm to compute shortest paths and a Speed-MFD to estimate average traveling speeds. A network-weighted k -mean algorithm separated the area into three distinct regions and the Speed-MFD data (Figure 3).

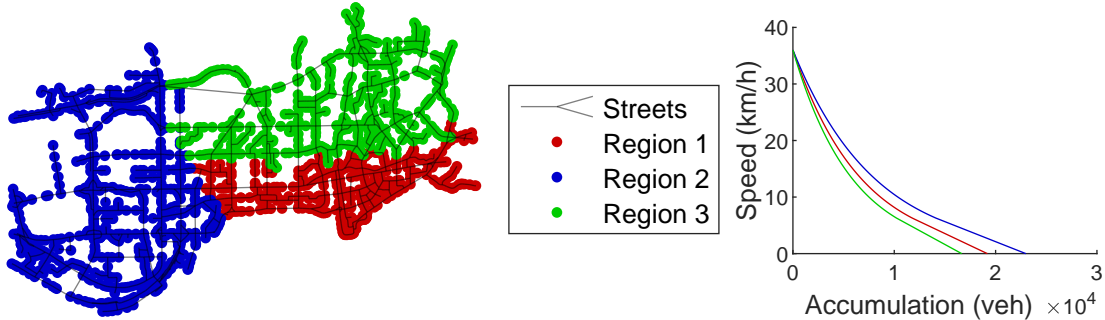


Figure 3: Settings of the experiments. (Left) urban area map and regions with respective centroids of the k -mean problems. (Right) Regional speed-MFD.

We assume a non-homogeneous Poisson arrival process for all travelers in the area to highlight the effects of imbalanced demand. Note that around 85% of arriving travelers use private vehicles, while the remaining use one of the ride-sourcing service options.

Sensitivity analysis

The first analysis of the repositioning strategy focused on the sensitivity to the fraction of drivers receiving the repositioning guidance. One can correlate such scenarios with the operator selecting a groups of drivers for a ‘loyalty program’, and as part of the benefits, these drivers receive improved guidance in their search for assignments. In the direction of this parallel, we refer to drivers which receive guidance as ‘guided’ ones, whereas the others are referred as ‘unguided’.

We assume a logit decision process, where the utility of each option is depicted exclusively by the revenue it generates to a ‘guided’ driver. Finally, we define that ‘unguided’ drivers look for the region with the highest demand per driver rate (similar to a ‘high demand’ flag in current ride-sourcing operations).

Additionally, we evaluated the results for fleet sizes of 2000, 2500 and 3000 active drivers in ride-sourcing services. We ran cases with 0%, 25%, 50%, 75% and 100% of drivers covered in the ‘loyalty program’. As final parameters, we considered booking fares of US\$2.20 and US\$2.00, and traveling fares of US\$1.00 and US\$0.80 for ride-hailing and ridesplitting, respectively.

Since the proposed repositioning framework must persuade ‘guided’ drivers, their outcomes should be higher than those of ‘unguided’ drivers. Figure 4 shows the average revenues of both groups of drivers at different guidance rates (fraction of the fleet that receives repositioning guidance information). Firstly, average revenues increased compared to a scenario where drivers never relocate in all cases with guidance (the exception occurs at 0% guidance rate). We must point out

that guided drivers consistently had higher revenues than non-compliant ones. It highlights that the proposed framework captured the possibility of areas with lower demand being more profitable. However, it is interesting to observe that as the operator expands the number of ‘guided’ drivers (more than 75% of the fleet), the combined average revenue slightly decreases.

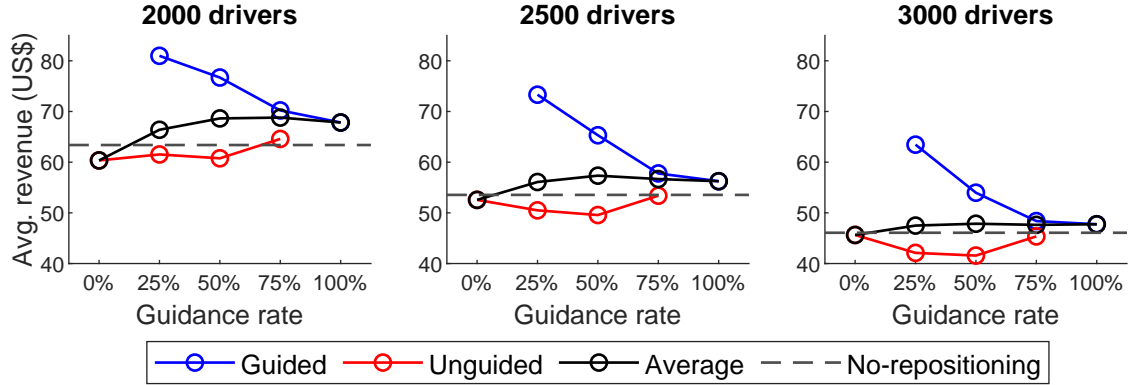


Figure 4: Average revenues of ‘guided’ and ‘unguided’ drivers following by a combination of these in scenarios with different penetration rates of the loyalty program, compared to the base ‘No-repositioning’ case.

It is interesting to take a closer look at individual revenues and understand their distribution. Figure 5 shows the histograms of the revenues for ‘unguided’ and ‘guided’ drivers in the base ‘No-repositioning’ scenario and one with 25% guidance ratio for a service fleet of 2000 drivers. With a left skewed distribution, no repositioning scenario had lower average revenue average. In the scenario with 25% of ‘guided drivers, the distribution of revenues was unimodal with the average close to the mode, creating a clear distinction between the groups. The average revenue of compliant drivers was higher than the revenue of 97% of non-compliant ones.

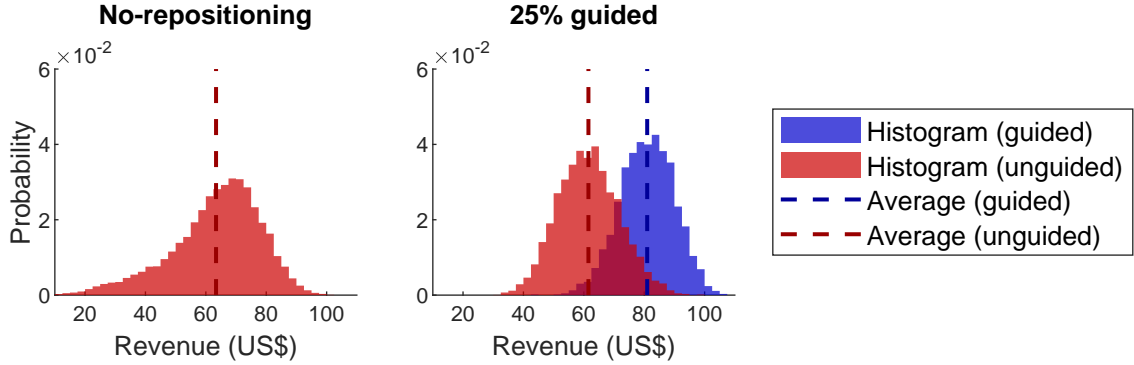


Figure 5: Histograms of revenues in two separated instances with 2000 drivers.

Benchmark comparisons

As argued earlier, other repositioning strategies exist to improve service quality but they can fall short on their reactive nature and/or full compliance assumptions. Here, we position the proposed strategy in comparison with other benchmark strategies. In addition to the ‘No-repositioning’ base case and the ‘Proposed’ strategy, we evaluate the strategies named below:

1. ‘Past-revenue’: Provides a portion of ‘guided’ drivers with an average revenue estimation accounting exclusively for past events in each area;
2. ‘Past-loss’: Dispatches the closest idle driver to the area of a recently lost request (Alonso-Mora et al., 2017);
3. ‘Coverage’: Performs optimal coverage control, distributing idle drivers according to the demand distribution in the area (Zhu et al., 2022);

It is important to highlight that the ‘Past-revenue’ strategy has a reactive nature but allows drivers to decide, whereas ‘Past-loss’ strategy is reactive and assumes full compliance with the instructions. Foremost, the objective of repositioning vehicles is to improve the service quality, especially by making the service available in previously uncovered areas. Figure 6 compares the number of unattended service requests (abandonments) for all evaluated strategies. Firstly, when no drivers receive guidance and base their decisions on ‘high-demand areas’ information (0% guidance) it, actually, worsened the service increasing abandonments compared to the base case. Reactive strategies (‘Past revenues’ and ‘Past losses’) performed poorly, with little to no improvement compared to the base case. The most interesting point goes to comparison between the ‘Proposed’ and the ‘Coverage’ approaches. Although optimized, the ‘Coverage’ approach was outperformed by the ‘Proposed’ approach with a 50% guidance ratio, decreasing abandonments by 61% for a fleet of 2000 vehicles. Nevertheless, there are small increase in abandonments if guidance ratio is higher than 50%, which highlights the limitations the individualized decision-making towards service quality.

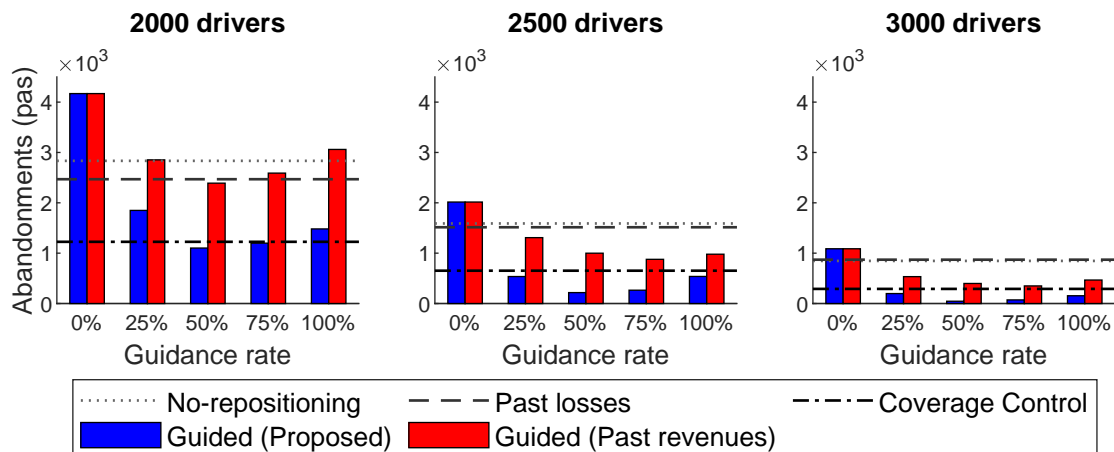


Figure 6: Summary of passenger abandonment results for all compared strategies.

In a different direction, most attention to ride-sourcing effect over congestion goes to their dead-heading. In Figure 7, we summarize the deadhead accumulated for each strategy for the same tested cases with 2000 vehicles and 50% guidance ratio. The only strategy to increase it, both in the unassigned and pick-up activities, is the ‘Past-revenues’ strategy. In the ‘Proposed’ approach, the unassigned deadhead is minimized but the ‘Coverage’ approach minimized the ‘Pick-up’ deadhead. It highlights two distinct points about of these strategies. First, while the ‘Proposed’ approach maximizes the chances of a drivers being assigned, the driver only needs to be as close as passengers’ waiting time tolerance accepts. Second, ‘Coverage’ approach had minimal ‘Pick-up’ deadhead because it “mimics” the expected demand distribution.

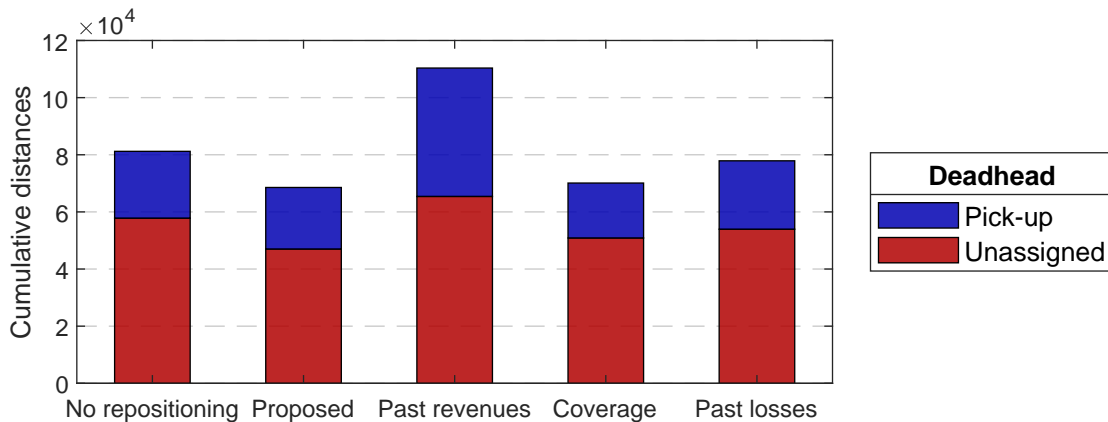


Figure 7: Summary of the deadhead associated with each repositioning strategy, separated into unassigned and pick-up kilometrages. Scenario with 2000 drivers and 50% of guidance rate in the ‘Proposed’ and ‘Past revenues’ methods.

5 CONCLUSIONS

In this paper we proposed a relocation strategy for ride-sourcing drivers by providing them with an estimate of their earnings. Therefore, we do not assume drivers unrestricted compliance to the provided guidance and, thus, they are free to make the decision that they expect will maximize their earnings. The first step in the proposed approach identifies the locations that are expected to maximize the chances of a driver getting a match in the forecast horizon. Then, a MDCTMC model is developed to capture the activities a driver will perform depending on his/her decision, which is later translated into an estimate of the driver's earnings. We showed that the proposed approach is likely to retain drivers confidence by improving their earnings compared to other drivers if the operator selects only a fraction of active drivers to provide guidance. Besides improving earnings, we show that the proposed approach manages to decrease the number of unserved requests in the system compared to several state-of-art benchmarks. It increased vehicle occupancy, and decreased the deadheading.

The findings provide a path for testing the impacts of different regulatory schemes in such systems. The provided guidance could further benefit drivers and unserved passengers, if it comes paired with other mechanisms to foster movements to poorly covered areas. It could include lower commissions in these areas, or other price changes to make it more attractive to drivers. Other research directions include developing optimal control to reposition without the decision-making process by drivers, which would be more realistic in cases with autonomous vehicles but it would also serve as upper (lower) bound for performance measurements and evaluation.

REFERENCES

- Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., & Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences USA*, 114(3), 462–467. doi: 10.1073/pnas.1611675114
- Beojone, C. V., & Geroliminis, N. (2021). On the inefficiency of ride-sourcing services towards urban congestion. *Transportation Research Part C: Emerging Technologies*, 124, 102890. doi: 10.1016/j.trc.2020.102890
- Beojone, C. V., & Geroliminis, N. (2022). A dynamic multi-region mfd model for ride-sourcing systems with ridesplitting. *arXiv*, 2211.14560. doi: 10.48550/arXiv.2211.14560
- Ingolfsson, A. (2005). Modeling the M(t)/M/s(t) queue with an exhaustive discipline. *Working paper*. Retrieved from http://www.bus.ualberta.ca/aingolfsson/working_papers.htm
- Powell, J. W., Huang, Y., Bastani, F., & Ji, M. (2011). Towards reducing taxicab cruising time using spatio-temporal profitability maps. In D. Pfoser et al. (Eds.), *Advances in spatial and temporal databases* (pp. 242–260). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Sadeghi, A., & Smith, S. L. (2019). On re-balancing self-interested agents in ride-sourcing transportation networks. In *2019 IEEE 58th Conference on Decision and Control (CDC)* (pp. 5119–5125). doi: 10.1109/CDC40024.2019.9030043
- Wang, H., & Yang, H. (2019). Ridesourcing systems: A framework and review. *Transportation Research Part B: Methodological*, 129, 122–155. doi: 10.1016/j.trb.2019.07.009
- Zhu, P., Sirmatel, I. I., Trecate, G., & Geroliminis, N. (2022). Distributed coverage control for vehicle rebalancing in mobility-on-demand systems. *TRB Annual Meeting*, 22-03340.

Exploring the effectiveness of different modeling approaches for predicting travel mode choice: An analysis of Machine Learning versus Econometric methods

Mohammadjavad Javadinasr ¹, Sina Asgharpour ², Motahare Mohammadi ³, Abolfazl (kouros) Mohammadian ⁴, Joshua Auld ⁵

1. Ph.D. Candidate, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor St, Chicago, IL 60607. Email: Mjavad2@uic.edu
2. Ph.D. Candidate, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor St, Chicago, IL 60607. Email: sasgha3@uic.edu
3. Ph.D. Candidate, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor St, Chicago, IL 60607. Email: mmoham73@uic.edu
4. Professor, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor St, Chicago, IL 60607. Email: Kouros@uic.edu
5. Ph.D. Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439. Email: jauld@anl.gov

Abstract

The selection of transportation modes has been a popular topic in transportation planning, and it has been studied for several decades using random utility optimization. However, recent developments in machine learning techniques have opened up new opportunities for accurate prediction. In this study, we used data from the Chicago including activity and travel records for 12,000 households in a 24-hour period. Our objective was to develop travel mode choice models for commute trips through estimating a multinomial logit model to identify the factors that influence people's choices of commute travel mode. Notably, our study is the first in Chicago to consider seven travel modes, including walking, biking, walking to transit, driving to transit, auto driver, auto passenger, and TNC. Additionally, we employed a machine learning classifier to model the mode choice problem and compared its performance with the econometric model.

Introduction

Investigating the contributing factors in forming people's decision to select a travel mode and ultimately finding accurate and reliable predictions for the share of each mode in the network is critical to transportation planning and development (Ortúzar and Willumsen 2011). Traditionally, the most widely employed approach to address the mode choice problem has been based on the concept of Random Utility Maximization (RUM) which is usually formulated through the multinomial logit models (MNL) (McFadden 1973). The simple parameter estimation procedure offered by the closed-form mathematical structure of the MNL models has resulted in widespread adoption of them in travel behavior studies (Hagenauer and Helbich 2017). Moreover,

the MNL models can shed light on the contribution of each explanatory variable on the dependent variable (i.e., the outcome choice) by providing the marginal effect values which are very useful, especially for policy implications and analysis. More recently and especially in the past decade, the emergence of big data and the computational power and insights brought by machine learning (ML) techniques have led to incorporating data-driven approaches into travel demand modeling. ML models act as powerful classifiers with high accuracy power. Instead of using a prespecified utility function, ML techniques rely on the relationships between the explanatory variables and the outcome class (i.e., the mode choice) and are able to capture complex and nonlinear relationships (Hillel et al. 2021). However, ML techniques are mostly considered black box tools, meaning they provide their predictions without giving enough information on how they have reached such results.

In this study, we develop travel mode choice models for commute trips in Chicago by employing a dataset from the Chicago Metropolitan Agency for Planning (CMAP) travel tracker survey conducted in 2018-19. The motivation for this study was to estimate a mode choice module for POLARIS, which is an agent-based modeling framework for integrated travel demand and network simulations in Chicago metropolitan area (Auld et al. 2016). One of our contributions is to estimate a mode choice model through the inclusion of emerging modes such as TNC by using the most updated travel survey available in Chicago. To do so, we use a multinomial logit model as well as a machine learning classifier to develop models that represent the choice between several modal options, including Walk, Bike, Auto driver, Auto passenger, Walk to public transit, Drive to public transit (i.e., park & ride) and TNCs (e.g., Uber, and Lyft). Moreover, using the MNL model, we calculate the value of time, which is a critical policy variable for transportation-related decisions for auto, transit, and TNC modes. Finally, we will compare the results of our econometric and machine-learning mode choice models.

- **Methodology**

In this study, we utilized two approaches to address the mode choice problem. Firstly, we employed the well-established multinomial logit (MNL) model (i.e., econometric analysis) to characterize the underlying factors that affect the decision of people in selecting commute travel modes in Chicago. Secondly, we utilized Gradient Boosting as a classification tools.

The Gradient Boosting method is an advanced version of decision-tree-based models, firstly developed by Friedman (2001). This method *boosts* the accuracy of a given learning algorithm by fitting a set of models. As a result, the ensemble of models outperforms the single model. The Gradient Boosting model, firstly, fits a simple decision tree to the train data, which usually has a poor fitting performance. Subsequently, it trains another decision tree to improve the error term (i.e., the difference between prediction and observation) in the previous tree. This process will continue to iterate until it gets to the minimized error term.

- **Results and Discussion**

- *Data*

The main data source for this study was obtained from a comprehensive travel survey conducted by Chicago metropolitan Agency for Planning (CMAP). This activity-based survey was implemented between August 2018 and April 2019 and includes the activity and travel information of 12,391 households on an assigned travel day. Considering the significant impact of work trips on the transportation system and traffic flow, in this study, we are only focused on mandatory commute trips to model the mode choice of individuals. After the cleaning process, the final dataset for commute trips includes 11442 trips that are used for the modeling steps of this study. We considered seven travel modes including, walking, biking, transit, park and ride (i.e., drive to transit), auto driver, auto passenger, and TNC (e.g., Uber, and Lyft) in the choice set. With almost 36% share, the auto driver had the highest rank among all other modes in commute trips in Chicago.

- *Multinomial Logit Model*

Table 1 presents the estimation results for the multinomial logit model for seven commute travel modes. In order to validate our econometric mode choice approach, we calculated the scalar measures of fit, i.e., McFadden's R2, for the estimated multinomial logit model which turned out to be 0.22. As can be seen in Table 2, the travel times corresponding to each mode are significantly negative in all utility functions which is consistent with the literature on mode choice modeling. People who possess a graduate degree are more likely to use active modes (i.e., walking or biking) for their commute trips. In the auto driver mode, the travel cost (e.g., the cost of gasoline) and the parking cost are two significant factors that negatively affect the utility of this mode. Moreover, people who own their homes and people with a higher number of vehicles in the household are more likely to use auto driver mode for commuting. This finding makes sense and can act as a proxy for the wealth of individuals.

Table 1. the results of the estimation of the MNL mode choice model for commute trips.

Variable	Walk	Bike	Auto driver	Auto Passenger	Transit	Park & ride	TNC
Constant	4.018***	REF ¹	3.229***	2.086***	4.174***	2.056***	-0.230
Walk Time	-0.140**						
Graduate Degree	2.433***	2.100***					
Flexible Work	1.045***						
Bike Time		-0.073*					
# Household Vehicle		-0.179**	0.385***				-0.503**
Employment		1.094***					2.624***
Auto Travel Time			-.0332**	-.0332***			-0.033**
Auto Cost			-.126***	-.126***			
Homeownership			.477***				
Parking Cost			-.00101*				
Transit Time					-0.0188*		
Transit Cost					-0.126**		

Transit Access Time						-0.076**		
Transit Egress time						-.077***		
Under 18						-2.586**		
Park & Ride Time							-0.018**	
Park & Ride Cost							-0.126**	
TNC Wait Time								-0.190**
TNC Cost								-0.101**

Note: 1) REF means the reference point.

* Significant at 90%, ** Significant at 95%, *** Significant at 99%.

Regarding the transit mode, the time and cost of the travel as well as the access/egress time revealed to be the significant determining factors. Moreover, travel cost and time are the only significant variables in the utility of the park & ride mode. One important finding is the significance of the TNC wait time in our MNL model suggesting in-vehicle travel time and travel cost are not the only alternative-specific features that influence the choice of TNCs. This finding can have policy implications, especially for companies such as Uber and Lyft to take measures that can lead to lower waiting times and increase the share of TNC mode in people's commute trips.

Another piece of information that can be extracted from table 2 is the value of times. The value of time (VOT) is the marginal substitution of the time associated with the cost of each mode (Abdel-Aal 2017). According to table 1, the VOT for auto modes (both auto driver and auto passenger) is \$ 15.8 per hr., for transit modes (both transit and park & ride) is \$ 8.9 per hr., and for TNC is \$ 19.7 per hr. These findings are consistent with the literature (Lam and Small 2001) given that it is expected for TNC to have higher VOT compared to auto and transit modes, and for transit mode to have the lowest among the specified modes.

- ***Machine learning model***

The entire dataset has 11,441 records with 68 explanatory variables and a target variable (i.e., mode choice class). Since some key variables contain a high percent of missing values (e.g., *transit time* with 44% missing), complete case analysis leads to loss of information. This issue is not the case for MNL model since it only needs features to be available in the corresponding utility functions. For instance, *access time to transit* is only available for trips that have transit as an available option in the choice set (please note that the available choice set is not the same for all observations). To tackle this issue in Machine Learning methods, we employed imputation by K-nearest Neighbors. Among different K values, K=10 (i.e., 10 neighbors) led to better imputation results in terms of bias reduction.

To train the models, we split the dataset into 80% train and 20% test datasets. Moreover, we employed 5-fold cross validation to calibrate the hyperparameters of the models.

2 shows the accuracy of the individual and final models over 5-fold train, and test datasets, respectively. As shown, the fluctuation (i.e., coefficient of variation) of the accuracy of the Gradient Boosting model on train data is marginal implying a consistent performance. The model has a mean accuracy of 0.877 on train data. Although the training datasets of the MNL and Machine Learning model are different, the R^2 statistics can reflect the train accuracy of the MNL model. According to Table , the R^2 of the MNL is 0.22 implying the poor performance of the MNL compared to the Gradient Boosting model. This difference highlights the advantage of the machine learning models over econometric models in terms of performance.

Table 2. Accuracy of Employed Methods on Train (5-fold Cross Validation) and Test Data

		Gradient Boosting
Train Data Accuracy (Cross Validation)	Fold 1	0.894
	Fold 2	0.898
	Fold 3	0.891
	Fold 4	0.898
	Fold 5	0.898
	Mean (Coeff. of Variation)	0.896 (0.0036)
Test Data Accuracy	Accuracy	0.897

To understand the pattern of misclassification among classes, we used confusion matrix. Confusion matrix constitutes the frequency table showing the distribution of the predicted and observed (true) classes for the test data instances. Therefore, it provides a validation of a classification model with respect to ground-truth information. 1 illustrates the confusion matrices for the Gradient Boosting model. In this figure, the cell numbers represent the percentage of each observed class falling into a predicted class. Therefore, the summation of the cells in each row is 1. Besides, this percentage is also represented by color intensity. Based on 1, in general, our ML model has a good performance in class prediction.

In our employed dataset, *bike*, *park & ride*, and *TNC* classes are the minority classes with the associated frequencies between 1% to 3% from whole data while the other classes have a frequency over 17%. Predicting minority classes is one of the challenges of ML models. According to Figure 1, the Gradient Boosting model can correctly predict 69% of instances with *bike*, 46% of *park & ride*, and 22% of *TNCs*. This pattern pertains to the low accuracy of the machine learning models in predicting minority classes in unbalanced data. In these cases, the MNL models can be a good alternative to predict the minority classes, especially noting the benefits of the MNL model in interpretation and policy evaluation.

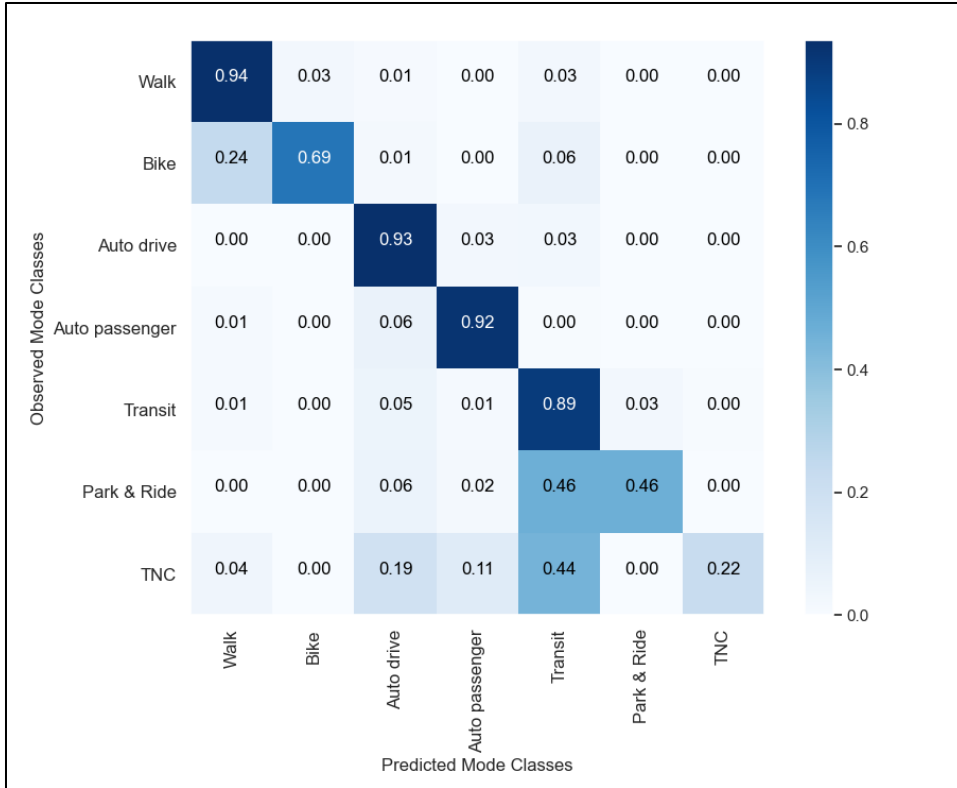


Figure 1: Confusion Matrix of Gradient Boosting Model

As a remarkable advantage, Gradient Boosting provides the *feature importance* facilitating to understand the individual contribution of the explanatory variables. In figure 2, feature importance of top 16 variables is illustrated. As shown, among all variables, *Auto Travel Time* has the most significant effect on the target variable (i.e., mode choice) with the feature importance of 0.164. Furthermore, *Auto Availability*, *Walk Time*, *Number of Passengers*, and *Transit Cost* are the other significant variables affecting individuals' mode choice. As shown, mainly time and cost attributes constitute the important factors affecting individuals' mode choice in the Gradient Boosting model which is consistent with the results of the MNL. Most of the significant variables that appeared in the feature importance are also identified by the MNL model except for *Auto Availability*, *# of Auto Passengers*, *Destination in CBD*. This trend confirms that although MNL has a poor prediction performance, it can detect the most significant attributes affecting the target variable.

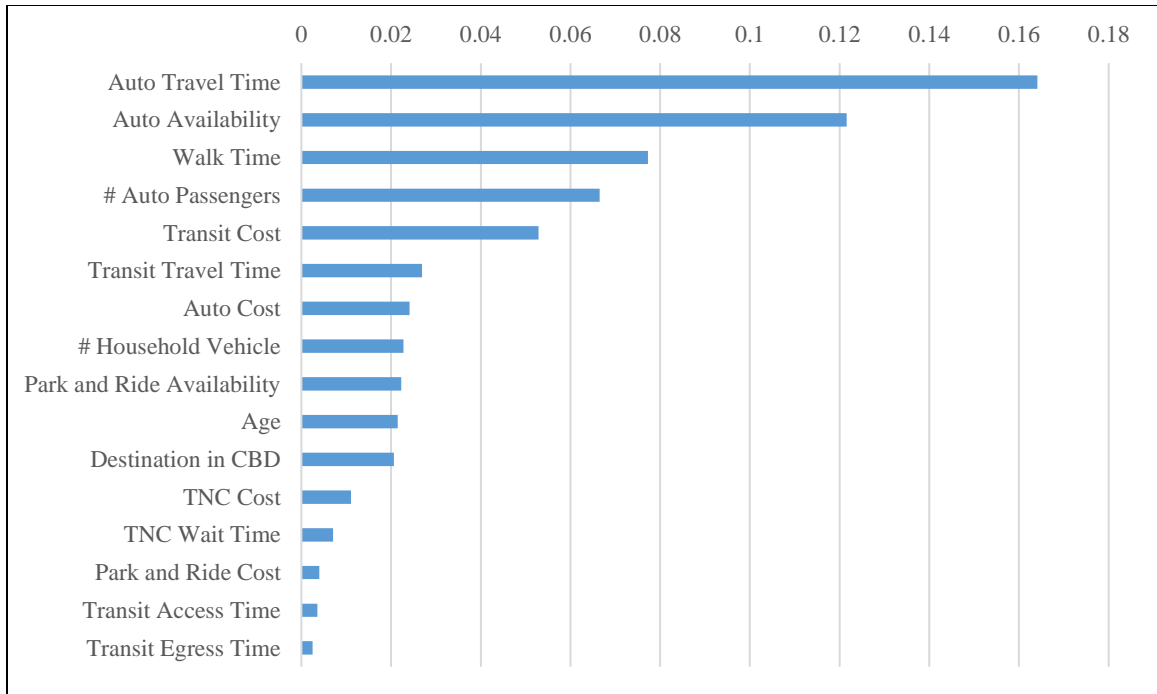


Figure 1: Feature Importance of Gradient Boosting Model

- **Discussion**

In this study, we developed travel mode choice models for commute trips in Chicago by employing a dataset from CMAP travel survey conducted in 2018-19. We considered seven travel modes including walking, biking, walking to transit, driving to transit (i.e., park and ride), auto driver, auto passenger, and TNC in the choice set. We employed econometric (i.e., Multinomial Logit Model) and machine learning (i.e., Gradient Boosting) techniques to address the mode choice problem. According to our results, travel times and costs are the most contributing factors in forming individuals’ decision to select a mode. In comparison, the MNL framework did not require imputing all the missing values due to the different utilities associated with the different choice sets defined for each observation. Moreover, consistent with the previous literature, we also corroborated the higher accuracy of ML techniques compared to the MNL model. Nonetheless, the accuracy power of the ML model was not equal in predicting all outcome classes. In general, in the presence of unbalanced data, the precision of our ML models to predict the minority classes decreased. The MNL model provided useful and interpretable alternative-specific variables related to the minority classes (such as the TNC wait time) that are very valuable for policy implications.

ACKNOWLEDGMENT

“The work described in this paper is sponsored by the U.S. Department of Energy (DOE) Vehicle Technologies Office (VTO) under the Systems and Modeling for Accelerated Research in Transportation (SMART) Mobility Laboratory Consortium, an initiative of the Energy Efficient Mobility Systems (EEMS) Program. The submitted manuscript has been created by the UChicago Argonne, LLC, Operator of Argonne National Laboratory (Argonne). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357.

The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.”

REFERENCES

- Abdel-Aal, M. M. M. (2017). “Value of time determination for the city of Alexandria based on a disaggregate binary mode choice model.” *Alexandria Engineering Journal*, Elsevier, 56(4), 567–578.
- Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., and Zhang, K. (2016). “POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations.” *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, 64, 101–116.
- Friedman, J. H. (2001). “Greedy Function Approximation: A Gradient Boosting Machine.” *The Annals of Statistics*, Institute of Mathematical Statistics, 29(5), 1189–1232.
- Hagenauer, J., and Helbich, M. (2017). “A comparative study of machine learning classifiers for modeling travel mode choice.” *Expert Systems with Applications*, Pergamon, 78, 273–282.
- Hillel, T., Bierlaire, M., Elshafie, M. Z. E. B., and Jin, Y. (2021). “A systematic review of machine learning classification methodologies for modelling passenger mode choice.” *Journal of Choice Modelling*, Elsevier, 38, 100221.
- Lam, T. C., and Small, K. A. (2001). “The value of time and reliability: measurement from a value pricing experiment.” *Transportation Research Part E: Logistics and Transportation Review*, Pergamon, 37(2–3), 231–251.
- McFadden, D. (1973). “Conditional logit analysis of qualitative choice behavior.”
- Ortúzar, J. de D., and Willumsen, L. (2011). *Modelling transport*.

On the Computation of Accessibility Provided by Shared Mobility

Severin Diepolder^{*1}, Andrea Araldo², Tarek Chouaki³, Santa Maiti⁴, Sebastian Hörl⁵, and Costantinos Antoniou⁶

¹Student, Technical University of Munich, Germany

²Associate Professor, SAMOVAR, Institut Polytechnique de Paris, France

³PhD Candidate, SystemX, France

⁴Postdoc, Technical University of Munich, Germany

⁵PhD, SystemX, France

⁶Professor, Technical University of Munich, Germany

SHORT SUMMARY

Shared Mobility Services (SMS), e.g., Demand-Responsive Transit (DRT) or ride-sharing, can improve mobility in low-density areas, often poorly served by conventional Public Transport (PT). Such improvement is mostly quantified via basic performance indicators, like wait or travel time. However, accessibility indicators, measuring the ease of reaching surrounding opportunities (e.g., jobs, schools, shops, ...), would be a more comprehensive indicator. To date, no method exists to quantify the accessibility of SMS based on empirical measurements. Indeed, accessibility is generally computed on graph representations of PT networks, but SMS are dynamic and do not follow a predefined network. We propose a spatial-temporal statistical method that summarizes observed SMS trips in a graph on which accessibility can be computed. We apply our method to a MATSim simulation study concerning DRT in Paris-Saclay.

Keywords: Accessibility; Public Transport; Shared Mobility;

1 INTRODUCTION

Location-based accessibility measures the ease of reaching surrounding opportunities via transport (Miller (2020)). Accessibility provided by conventional PT is generally poor in low-demand areas, e.g., suburbs (Badeanlou et al. (2022)), because a high frequency and high coverage service in such areas would imply an unaffordable cost per passenger. Poor PT accessibility in the suburbs makes them car-dependent, which prevents urban regions from being sustainable ((Saeidizand et al., 2022, Section 2.2)). SMS, e.g., Demand-Responsive Transit (DRT), ride-sharing, carpooling, car-sharing, are potentially more efficient than conventional PT in the suburbs (Calabrò (2023)). However, their current deployment is commonly led by private companies targeting profit maximization. This may turn SMS into additional source of congestion and pollution (Henao & Marshall (2019); Erhardt et al. (2019)).

We believe that SMS deployment should be overseen by transport authorities under the logic of accessibility improvement. To this aim, a method is needed, able to compute impact of SMS on accessibility, based on empirically observed trips. To the best of our knowledge, this paper is the first to propose such a method. Chandra et al. (2013) study how DRT improves connection to conventional PT stops, without considering the impact on accessing opportunities. Nahmias-Biran et al. (2021) and Zhou et al. (2021) calculate based accessibility from Autonomous Mobility on Demand, based on utilities perceived by agents within simulation. By contrast, our method computes accessibility solely based on observed SMS trip times, either from the real world or simulation. A first attempt of integrating SMS into the graph-based description of PT is done by Le Hasif et al. (2022). However, they use analytic models to model SMS performance and thus fail to give real insights adapted to the areas under study. Our effort consists instead of estimating accessibility from empirical observations via spatial-temporal statistics. General Transit Feed System (GTFS) is the standard data format for PT schedules. Recently, the GTFS-Flex extension allows also describing SMS (Craig & Shippy (2020)). Although our estimates could thus be fed into GTFS-Flex data, for the sake of simplicity, we use plain GTFS instead.

Our contribution consists in developing a spatial-temporal statistical pipeline to transform SMS trip data observations in a graph representation, on top of which well-established accessibility

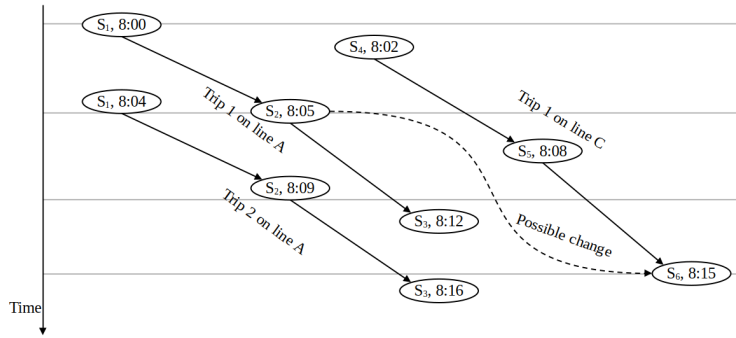


Figure 1: Time-expanded graph, representing two trips on line A and one trip on line B, as well as a potential change.

computation can be performed. The observations that can be taken as input might come from real measurements or from simulation. This paper’s observations come from a MATSim simulation study of DRT deployment in Paris Saclay, from Chouaki et al. (2023).

By providing a first method to compute the accessibility of SMS on empirical observations, this work can contribute to a better understanding of the potential of SMS and guide their future deployment.

2 METHODOLOGY

Accessibility

As in (Biazzo et al. (2019)), the study area is tessellated in hexagons with a grid step of 1km, whose centers $\mathbf{u} \in \mathbb{R}^2$ are called centroids and denoted with set $\mathcal{C} \subseteq \mathbb{R}^2$. Each hexagon contains a certain quantity of *opportunities*, e.g., jobs, places at school, people. With $O_{\mathbf{u}}$ we denote the opportunities in the hexagon around \mathbf{u} and with $T(\mathbf{u}, \mathbf{u}', t)$ the time it takes to arrive in \mathbf{u}' , when departing from \mathbf{u} at time t . As in Miller (2020), accessibility is the amount of opportunities that one can reach departing from \mathbf{u} at time of day t within time τ :

$$acc(\mathbf{u}) \equiv \sum_{\mathbf{u}' \in \mathcal{C}(\mathbf{u}, t)} O_{\mathbf{u}'}. \quad (1)$$

$\mathcal{C}(\mathbf{u}, t) = \{\mathbf{u}' \in \mathcal{C} | T(\mathbf{u}, \mathbf{u}', t) \leq \tau\}$ is the set of centroids reachable within τ . By improving PT, such set can be enlarged such as to consent to reach more opportunities. In this work, the opportunities are the number of people (residents) that can be reached. $T(\mathbf{u}, \mathbf{u}', t)$ is always computed on a graph representation of the transport network. However, SMS are not based on any network. Our effort is thus to build a graph representation of SMS, despite the absence of a network model.

Time-Expanded Graph Model of conventional PT

Inspired by Fortin et al. (2016) and Le Hasif et al. (2022), we model PT as a time-expanded graph \mathcal{G} , compatible with the GTFS format. The nodes of \mathcal{G} are *stoptimes*. Stoptime (\mathbf{s}, t) indicates the arrival of a PT vehicle at stop $\mathbf{s} \in \mathbb{R}^2$ (modeled as a point in the plane) at time $t \in \mathbb{R}$. Different trips on a certain line are represented as sequences of different stoptimes, as in Figure 1, as well as potential line change, within 15 minutes walk, assuming 5 Km/h walk speed, if it is possible to arrive at the new line on time. When a user departs at time t_0 from location \mathbf{x} for location \mathbf{x}' , they can simply walk (but no more than the maximum walk time). Or they can walk to \mathbf{s} , board a PT vehicle at t (corresponding to a stoptime (\mathbf{s}, t) , use PT up to a stoptime (\mathbf{s}', t') and from there walk to \mathbf{x}' . The arrival time at \mathbf{x}' will be t' plus the time for walking. Users are assumed to always choose the path with the earliest arrival time. Path computation is performed within CityChrone (Biazzo et al. (2019)). No capacity constraints are considered.

Integration of shared mobility into the time-expanded graph

SMS is assumed to provide a feeder service to traditional PT. In a feeder area $\mathcal{F}(\mathbf{s}) \subseteq \mathbb{R}^2$ around some selected stops \mathbf{s} (which we also call *hubs*), SMS provide connection to and from \mathbf{s} . The set

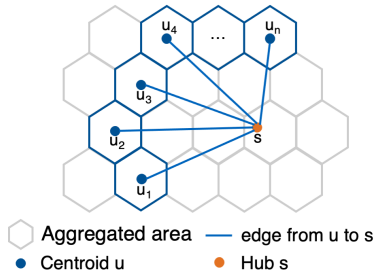


Figure 2: Hub and virtual trips provided by SMS

of centroids in such an area is $\mathcal{C}(\mathbf{s}) = \mathcal{C} \cap \mathcal{F}(\mathbf{s})$. In this section, we will focus on access trips (from a location to a PT stop) performed via SMS. The same reasoning [applies to](#) egress trips, *mutatis mutandis*. We assume to have a set \mathcal{O} of observations. Each observation $i \in \mathcal{O}$ corresponds to an access trip and contains:

- Time of day $t_i \in \mathbb{R}$ when the user requested a trip to the flexible service
- Location $\mathbf{x}_i \in \mathbb{R}^2$ where the user is at time t_i
- Station s_i where the user wants to arrive via the SMS feeder service
- Duration w_i indicating the wait time before the user is served: it can be the time passed between the time of request and the time of pickup from a vehicle, in case of ride-sharing, DRT or carpooling; it can be the time to wait until a vehicle is available at the docks in a car-sharing or bike-sharing system.
- Travel time y_i : time spent in the SMS vehicle to arrive at \mathbf{s} .

We interpret y_i and w_i as realizations of spatial-temporal random fields (Handcock & Wallis (1994)): for any time of day $t \in \mathbb{R}$ and physical location $\mathbf{x} \in \mathcal{F}(\mathbf{s})$, random variables $W^{\mathbf{s}}(\mathbf{x}, t), Y^{\mathbf{s}}(\mathbf{x}, t)$ represent the times experienced by a user appearing in t and \mathbf{x} , for any stop \mathbf{s} . In the following subsection we will compute estimations $\hat{w}^{\mathbf{s}}(\mathbf{u}, t), \hat{y}^{\mathbf{s}}(\mathbf{u}, t)$ of expected values $\mathbb{E}[W^{\mathbf{s}}(\mathbf{u}, t)], \mathbb{E}[Y^{\mathbf{s}}(\mathbf{u}, t)]$ at centroids $\mathbf{u} \in \mathcal{C}(\mathbf{s})$.

To integrate SMS into PT graph \mathcal{G} , SMS are represented as a set of “virtual” trips, running between centroid $\mathbf{u} \in \mathcal{C}(\mathbf{s})$ and hub \mathbf{s} (Figure 2). Each trip has travel time $\hat{y}^{\mathbf{s}}(\mathbf{u}, t)$. The access connection between centroid \mathbf{u} and hub \mathbf{s} is modeled as a sequence of trips, corresponding to stop times (\mathbf{u}, t_j) , for different values of departure time. We thus have to compute the list of such departure times. To do so, we interpret the inter-departure time between such trips as a random field $H^{\mathbf{s}}(\mathbf{x}, t)$, which represent a “virtual” headway. The value of such an interval in \mathbf{x} and t is also a spatial-temporal random field. We use the common approximation $H^{\mathbf{s}}(\mathbf{x}, t) = 2 \cdot W^{\mathbf{s}}(\mathbf{x}, t)$, idealizing the SMS headway as regular so as to apply (2.4.28) from Cascetta (2009). Therefore, we separate stoptimes by $2 \cdot w^{\mathbf{s}}(\mathbf{u}, t)$. More precisely, the stoptimes corresponding to access trips departing from centroid \mathbf{u} to hub \mathbf{s} are:

$$\begin{aligned}
 &(\mathbf{u}, t_0), \\
 &(\mathbf{u}, t_j) \quad \text{where } t_j = t_{j-1} + 2 \cdot \hat{w}^{\mathbf{s}}(\mathbf{u}, t_{j-1}) \text{ for } j = 1, 2, \quad \text{until 11:59 pm,} \quad (2) \\
 &(\mathbf{u}, t_j) \quad \text{where } t_j = t_{j+1} - 2 \cdot \hat{w}^{\mathbf{s}}(\mathbf{u}, t_{j+1}) \text{ for } j = -1, -2, \quad \text{until 00:00 am.}
 \end{aligned}$$

Correspondent stoptimes are added to represent the arrival of access trips $(\mathbf{s}, t_j + \hat{y}^{\mathbf{s}}(\mathbf{u}, t_j))$ and an edge between each departure stoptime and the respective arrival stoptime is added. A similar process is applied for egress trips. At the end of the described process, time-expanded graph \mathcal{G} is enriched with stoptimes and edges representing SMS trips. Having done so, it is possible to reuse accessibility calculation methods for time-expanded graphs, such as CityChrono Biazzo et al. (2019), with no modifications required.

Estimation of Waiting and Travel Times

We now explain how we construct estimation $\hat{w}^{\mathbf{s}}(\mathbf{u}, t)$ used in the previous subsection, for access SMS trips only. Similar reasoning can be applied to $\hat{y}^{\mathbf{s}}(\mathbf{u}, t)$ and egress trips. We assume random field $W^{\mathbf{s}}(\mathbf{x}, t)$ is approximately temporally stationary within each timeslot:

$$W^{\mathbf{s}}(\mathbf{x}, t) = W^{\mathbf{s}}(\mathbf{x}, t_k), \quad \forall \mathbf{x} \in \mathbb{R}^2, \forall t \in [t_k, t_{k+1}[, \forall \text{ station } \mathbf{s} \quad (3)$$

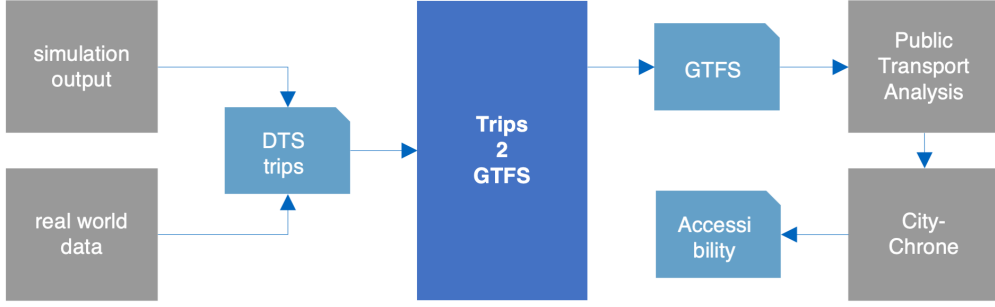


Figure 3: Implementation pipeline. CityChrono is from Biazzo et al. (2019).

For any timeslot, we thus just need to find estimation $\hat{w}_{t_k}^s(\mathbf{u})$ of the expected values of random field $W_{t_k}^s(\mathbf{x}) \equiv W^s(\mathbf{x}, t_k)$. First the observations \mathcal{O} are projected onto time-slot $[t_k, t_{k+1}[$:

$$\mathcal{O}_{t_k}^s \equiv \{\text{observation } i = (\mathbf{x}_i, w_i, y_i) | i \in \mathcal{O}, t \in [t_k, t_{k+1}[, i \text{ is related to an access trip to } \mathbf{s}\} \quad (4)$$

Estimation $\hat{w}_{t_k}^s(\mathbf{u})$ is computed by Ordinary Kriging (AA.VV. (2018)) on the observations $\mathcal{O}_{t_k}^s$ as a convex combination of observations w_i :

$$\hat{w}_{t_k}^s(\mathbf{x}) = \sum_{i \in \mathcal{O}_{t_k}^s} \lambda_i \cdot w_i \quad (5)$$

In short (details can be found in Section 19.4 of Chilès & Desassis (2018)), coefficients λ_i are computed based on a *semivariogram* function $\gamma_{t_k}^s(d)$, which obtained as a linear regression model, with predictors $d_{i,j}$ (distances between all pair of observations) and labels $\gamma_{i,j}$, which are called *experimental semivariances*:

$$\gamma_{i,j} \equiv \frac{1}{2} \cdot (w_i - w_j)^2 \quad (6)$$

The underlying assumption here is that correlation between wait times in different locations vanishes with the distance between such locations. The semivariogram gives the “shape” of this vanishing slope. In estimation (5), closer observations will have a higher weight. Under hypothesis on spatial stationarity and uniformity in all directions (*Kriging Interpolation* (2023)), Theorem 2.3 of Yakowitz & Szidarovsky (1985) proves that Kriging gives an asymptotically biased estimator: as the number of observations tends to infinite, $\hat{w}_{t_k}^s(\mathbf{x})$ tends to the “true” $\mathbb{E}[W_{t_k}^s(\mathbf{x})]$.

3 IMPLEMENTATION

The methodology of Section 2 is implemented in a Python pipeline, which we release as open source (Diepolder (2023)) and is depicted in Figure 3.

1. We first get centroids and cells performing the tessellation via CityChrono.
2. We read the file containing the observations (SMS trips). Such a file can be a simulation output or measurements of real SMS. Each observation includes the same information as in page 3. Observations are stored in a dataframe.
3. We assume SMS is deployed as feeder (as it is the case for the MATSim simulation on which we perform our analysis). Therefore, we can classify every SMS trip as either access or egress, depending on whether the origin or the destination is a PT stop.
4. To establish the feeder area $\mathcal{F}(\mathbf{s})$, we find among the observations \mathcal{O} the furthest cell from \mathbf{s} in which a trip to/from \mathbf{s} has occurred. All cells within such a distance, are assumed to be in $\mathcal{F}(\mathbf{s})$. Observe that feeder areas of different hubs may overlap.
5. We group observations in timeslots (Figure 9).
6. In each time slot $[t_k, t_{k+1}[$ and each centroid \mathbf{u} around each stop \mathbf{s} , we perform Kriging via library `pyInterpolate` (Moliński (2022)) to obtain estimations $\hat{w}^s(\mathbf{u})$ and $\hat{y}_{t_k}^s(\mathbf{u})$.

Table 1: Parameters used for the numerical results.

Parameter	Value	Reference
Side of a hexagon (tessellation)	1km	Badeanlou et al. (2022)
τ (Equation (1))	1 hour	Badeanlou et al. (2022)
Total number of DRT trips	14700	
- access trips	5289	
- egress trips	9412	
Total number of hubs	16	
Walk times		Computed via OpenStreetMap
Population Distribution		from the simulation scenario from Chouaki et al. (2023)

7. We obtain stoptimes and edges using the estimations above, as specified in (2). We add stoptimes and edges to the GTFS data of conventional PT, following the specifications in *GTFS Reference Document* (2023).
8. We give the obtained graph to CityChrono, which will give us accessibility scores in all the centroids.

4 RESULTS AND DISCUSSION

Data Source of the observations

The observation dataset in this study comes from a MATSim simulation, from Chouaki & Puchinger (2021); Chouaki et al. (2023), of door-to-door Demand-Responsive Transit (DRT), deployed as a feeder to and from conventional PT, in Paris-Saclay. The area in which DRT is deployed is depicted in Figure 4, but the entire Paris Region is simulated. Scenario parameters are in Table 1.

Analysis of Temporal and Spatial Patterns of DRT trips

Figure 5 clearly shows morning peak [7 : 00, 10 : 00[, evening peak [16 : 00, 19 : 00[and off-peak (all the other intervals).

In the following figures, the measures DRT trips toward/from all hubs, without distinguishing between hubs. Figures 6 and 7 are negative results: wait and travel times (figures on the right) do not appear to be spatially stationary (the distribution of values measured close to the related PT stops is different than further). Therefore, our estimations are not guaranteed to be asymptotically unbiased (page 4). In our future work, we will explore indirect estimation of wait and travel times through other indicators, e.g., the detour factor of DRT, which respect the requirements for the unbiasedness of Kriging.

Figure 8 shows that wait time follows expected peak/off-peak patterns. Values are generally slow since the simulation is configured so that a DRT trip is accepted only if the dispatcher predicts it is possible to serve it within 10 minutes. All wait times exceeding this limits might be due to the dispatcher not taking traffic correctly into account.

Estimation of Waiting and Travel Times

Figure 9 shows that timeslots of 1h preserve the temporal pattern of trips, so 1h should be preferred to smaller timeslots, so as to perform Kriging with as many observations as possible.

Within each timeslot, estimation of wait and travel times is based on Kriging, which exploits spatial correlation. First, we note in Figure 10 that travel times close to hubs are shorter than further away. Then, we note that the experimental semivariance in Figure 11, i.e., the $\gamma_{i,j}$ between pair of observation i, j (Equation (6)), increases with the spatial distance between the observations: the closer the observations, the more similar are the respective travel times measured therein.

Such trends are not as evident for wait times (Figure 12) although similarity between observations still decay with distance (Figure 13).

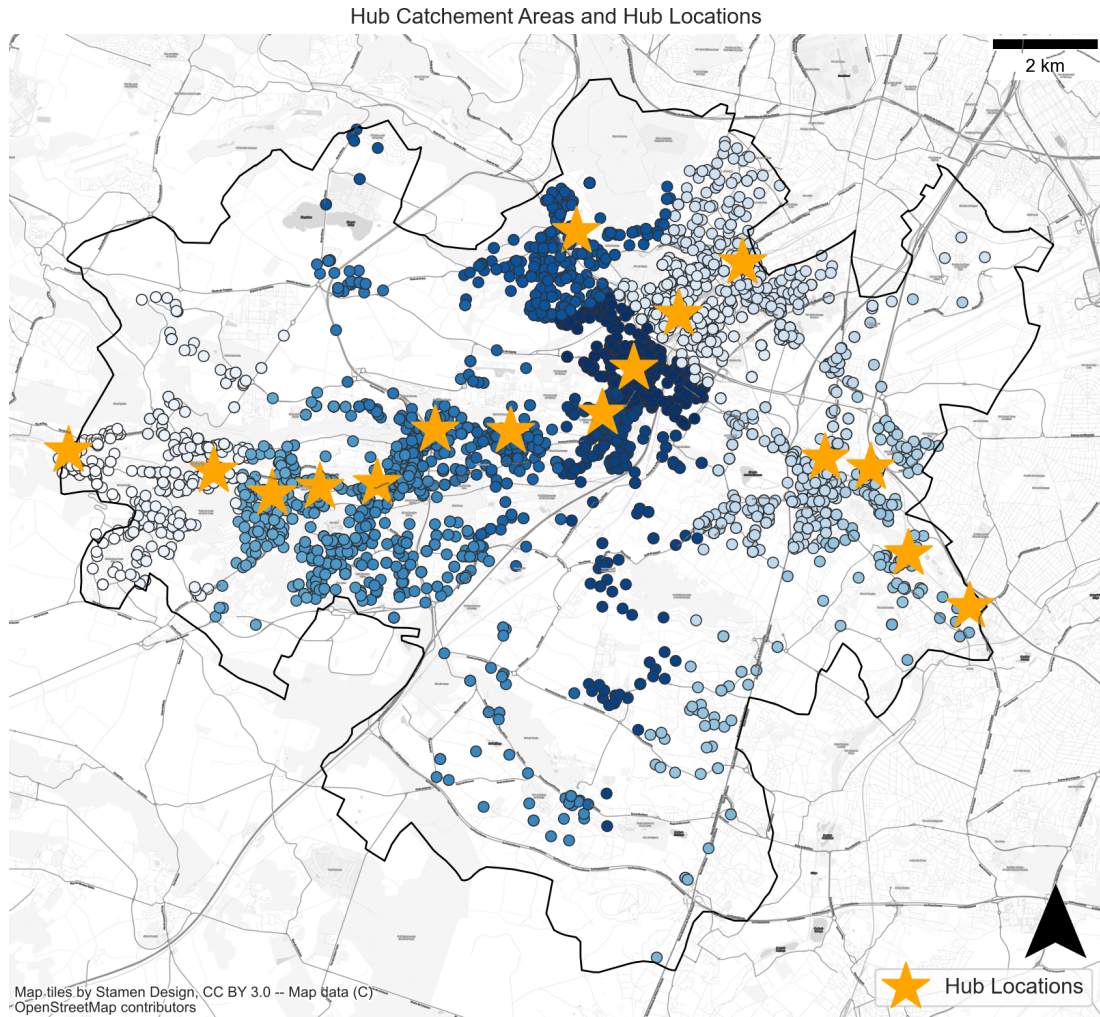


Figure 4: Hub Catchement Areas and Hub Locations. Each dot corresponds to the origin of one trip observed during the simulation. The differentiation in color of the observed trip origins indicates the catchment by different hubs

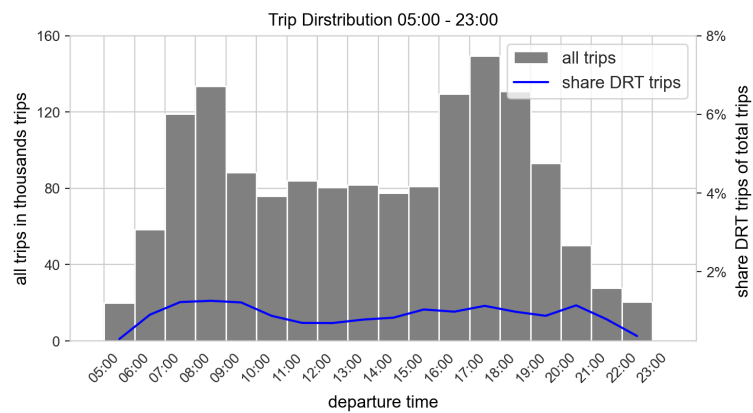


Figure 5: Trips over time. One trip is defined by the departure within the study area of Paris Saclay. A trip consisting out of multiple legs (e.g. walk + drt + PT is considered as one trip)

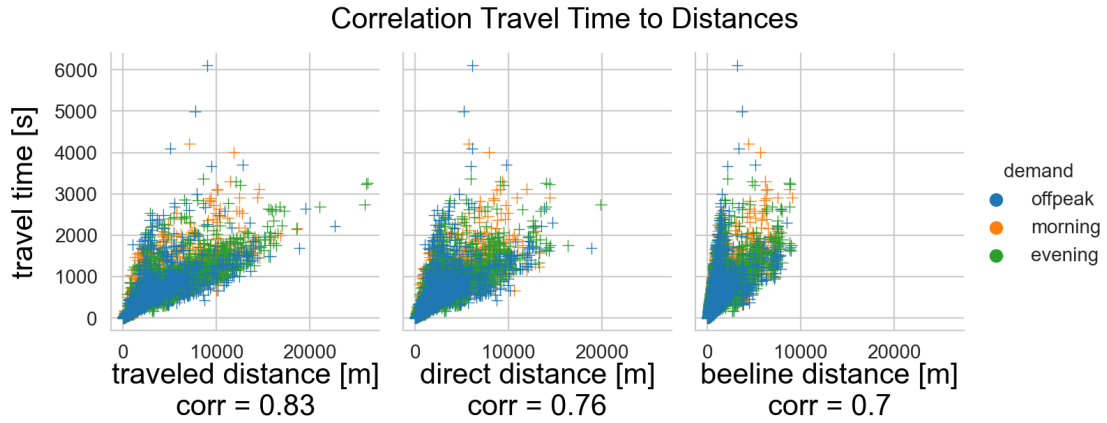


Figure 6: Relation between travel time and distance measures. Traveled distance is the actual Km traveled by the user inside the DRT vehicle. Direct distance is the one from the shortest road network road from the origin centroid to the hub. Beeline is the Euclidean distance.

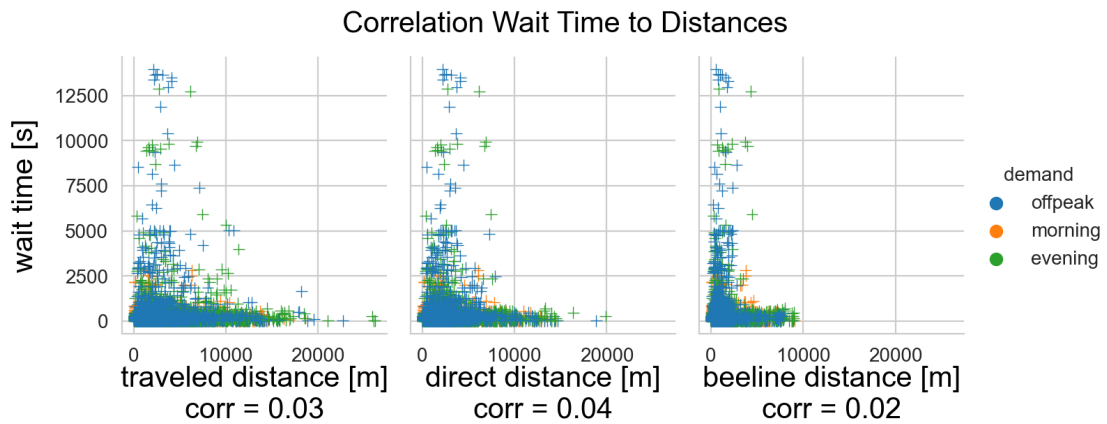


Figure 7: Relation between wait time and distance measures.

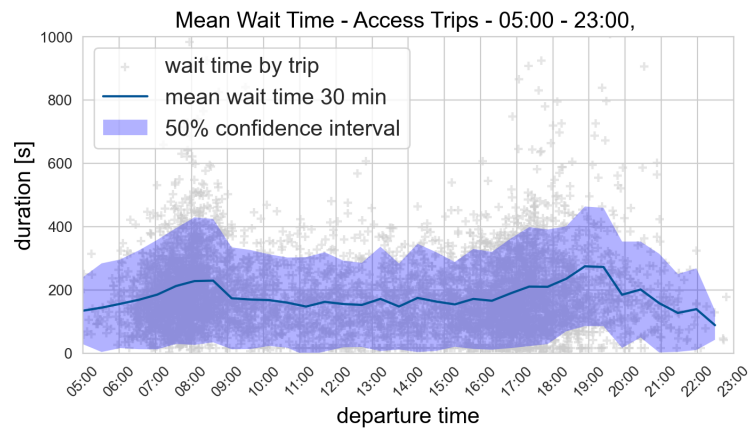


Figure 8: Mean Wait Time - A moving average of wait time during one day

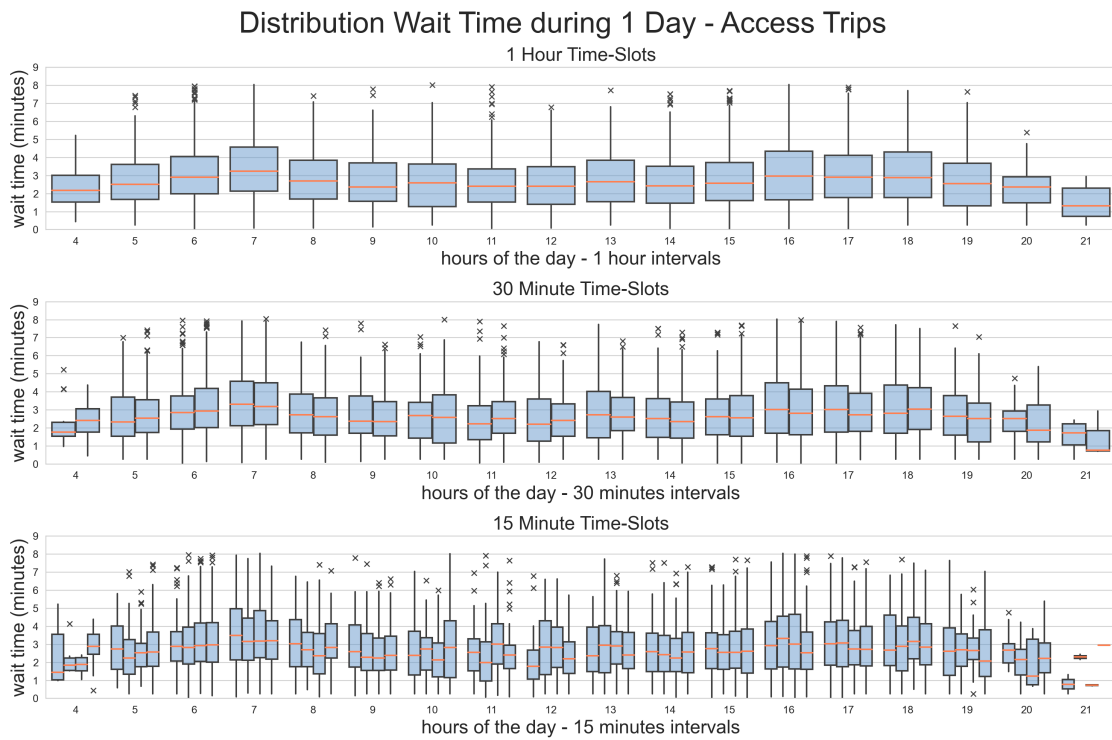


Figure 9: Comparison of different timeslot sizes. Values exceeding 10 minutes are not depicted, as they are due to simulation events unpredictable for the SMS dispatcher

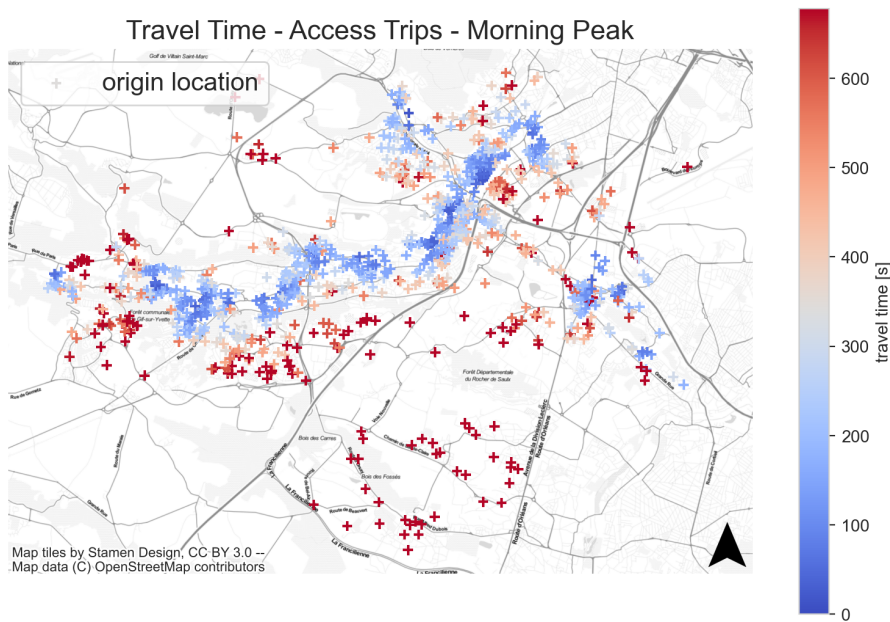


Figure 10: Spatial Trend Travel Time for access time, morning peak (evening and off-peak show similar trends).

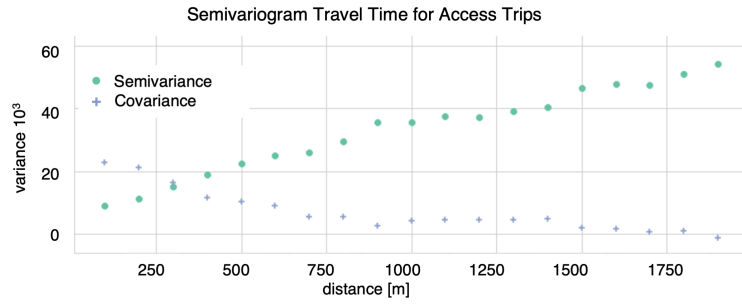


Figure 11: Spatial correlation of travel time observations

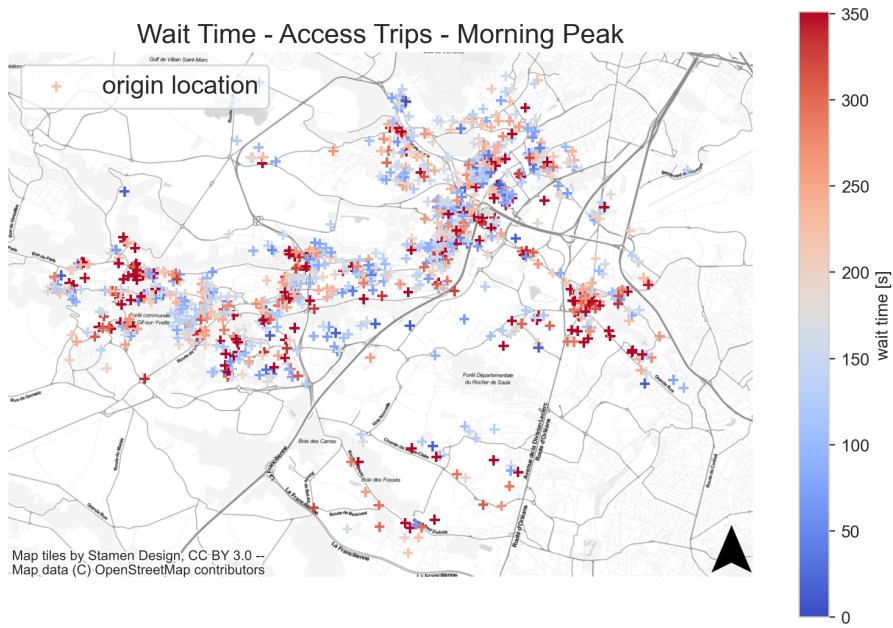


Figure 12: Spatial Trend Wait Time - No clear pattern can be identified, indicating low spatial autocorrelation

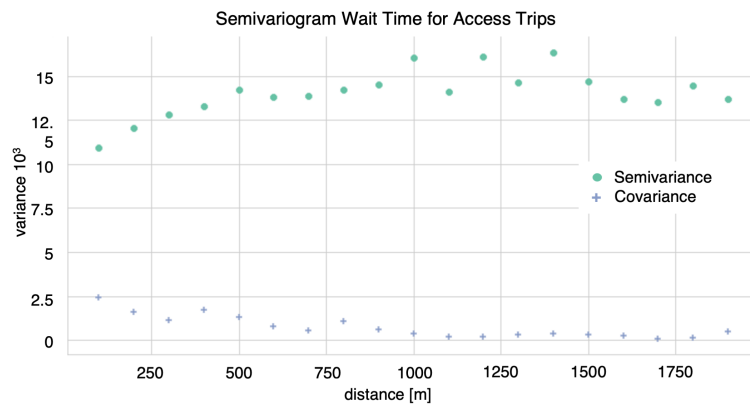


Figure 13: Spatial correlation of wait time observations

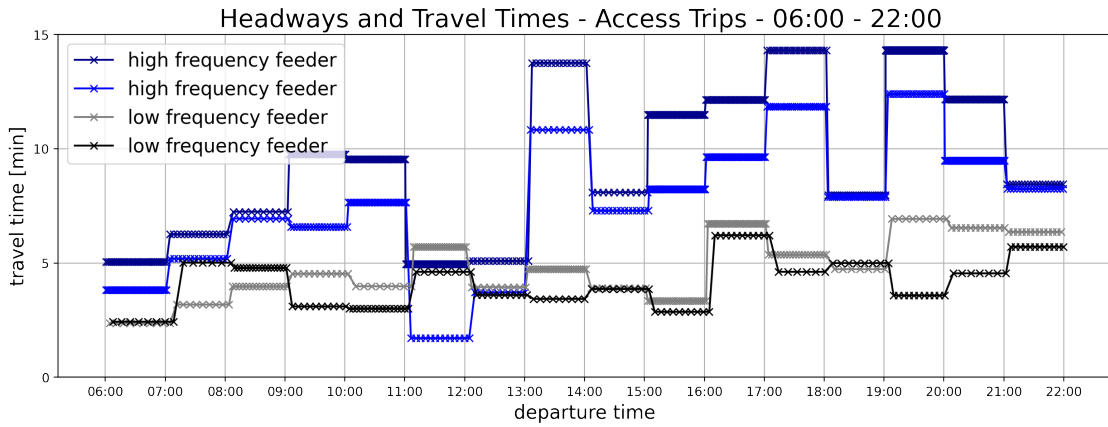


Figure 14: Headway and Travel Times of some examples of virtual DRT trips. Each departure time of a virtual DRT trip is indicated by a cross. The respective travel time is indicated by the y-axis value.

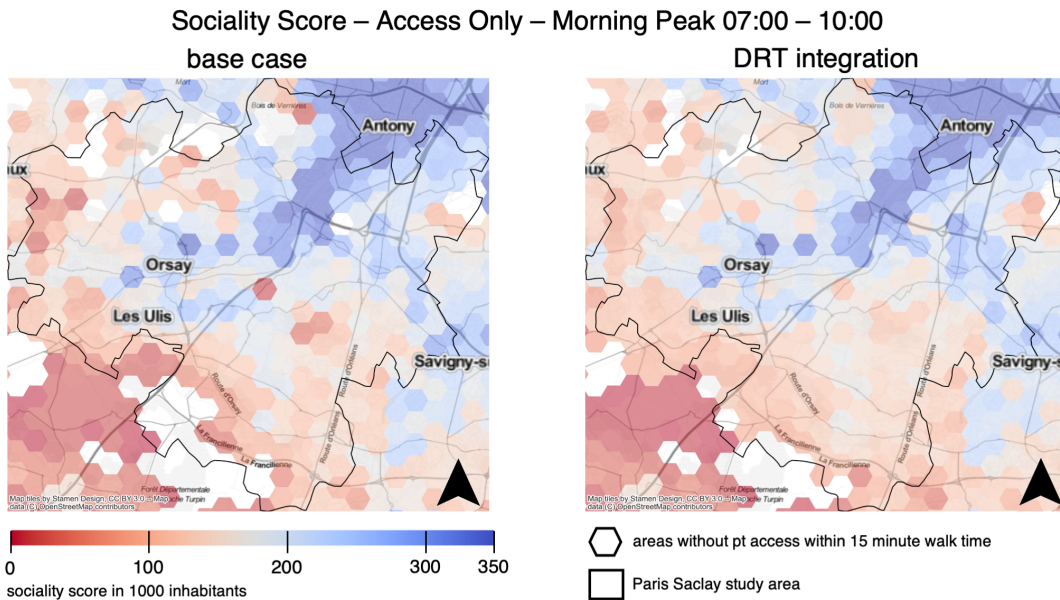


Figure 15: Sociality Score - Access Only - Morning Peak 07:00 - 10:00

Improvement of Accessibility Brought by DRT

Figure 14 shows headway and travel times of the virtual DRT trips added to the PT graph. We can then compute accessibility on this graph. Note that accessibility varies with the time of day (1). However, in the following figures we show averages over the time periods mentioned.

First, we study a system with DRT access services only (no egress). Figure 15 shows that the catchment area is expanded, especially in the south: hexagons with no access to PT within 15 minutes walk, can now use PT. Figure 16 shows more clearly the improvement in accessibility brought by improved access to PT thanks to DRT. As only access SMS feeder is added in Paris Saclay, the areas outside Saclay do not show any changes, except slight improvement in some locations, for instant south of Versailles, possibly due to the possibility for travelers starting from there to make changes in Saclay, which are enhanced by DRT.

Accessibility improvements are even greater in peak hours (Figures 17 and 18, as DRT compensates for the low frequency of conventional PT.

Figure 19 shows the improvement in accessibility when both access and egress trips are added, averaged over the entire day. Improvement is much greater than the access-DRT only case. Moreover, improvement is also visible also outside Saclay: users from everywhere can now reach opportunities in Saclay faster, thanks to DRT egress connections.

Sociality Score Improvement – Access Only – Morning Peak 07:00 – 10:00

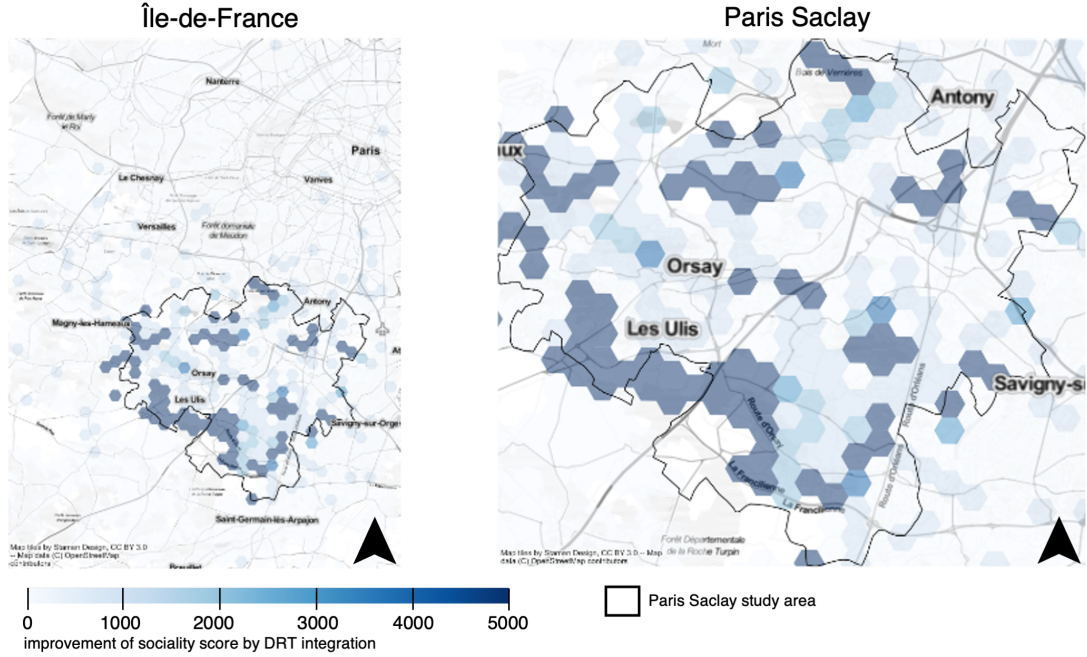


Figure 16: Sociality Score Improvement - Access Only - Morning Peak 07:00 - 10:00

Sociality Score – Access Only – Off Peak 10:00 – 16:00

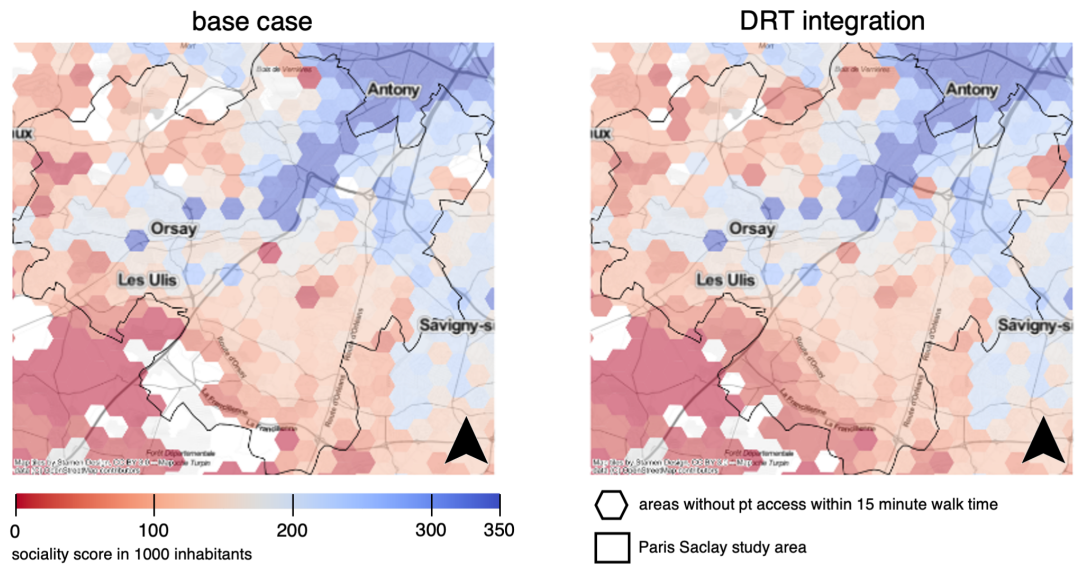


Figure 17: Sociality Score - Access Only - Off Peak 10:00 - 16:00

Sociality Score Improvement – Access Only – Off Peak 10:00 – 16:00

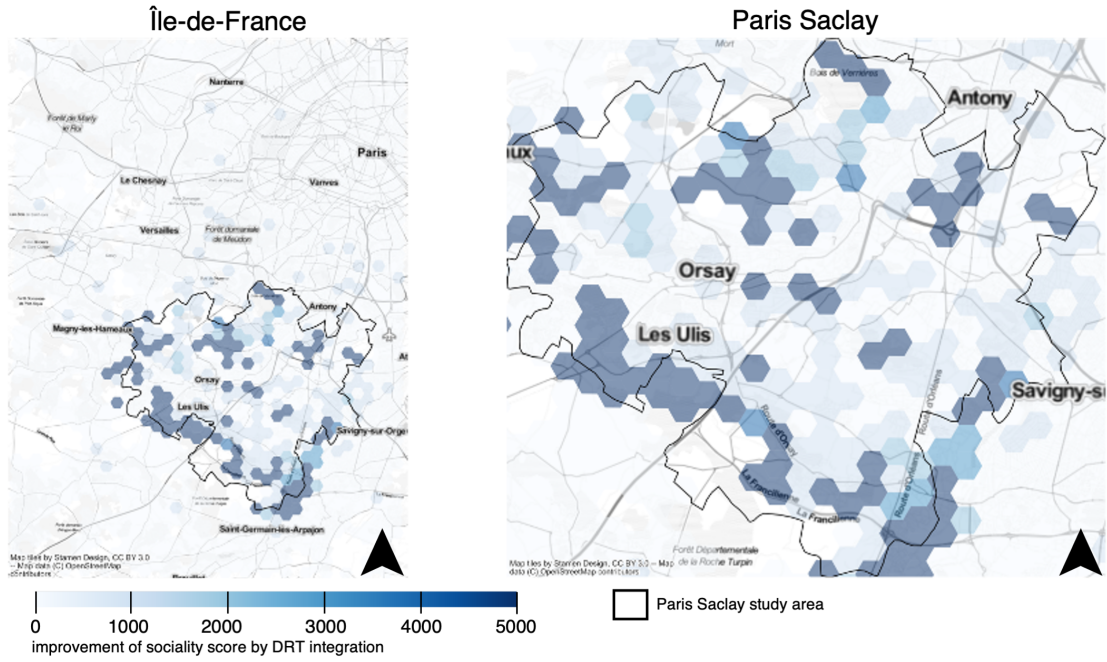


Figure 18: Sociality Score Improvement - Access Only - Off Peak 10:00 - 16:00

Sociality Score Improvement – Access & Egress – Full Day 05:00 – 23:00

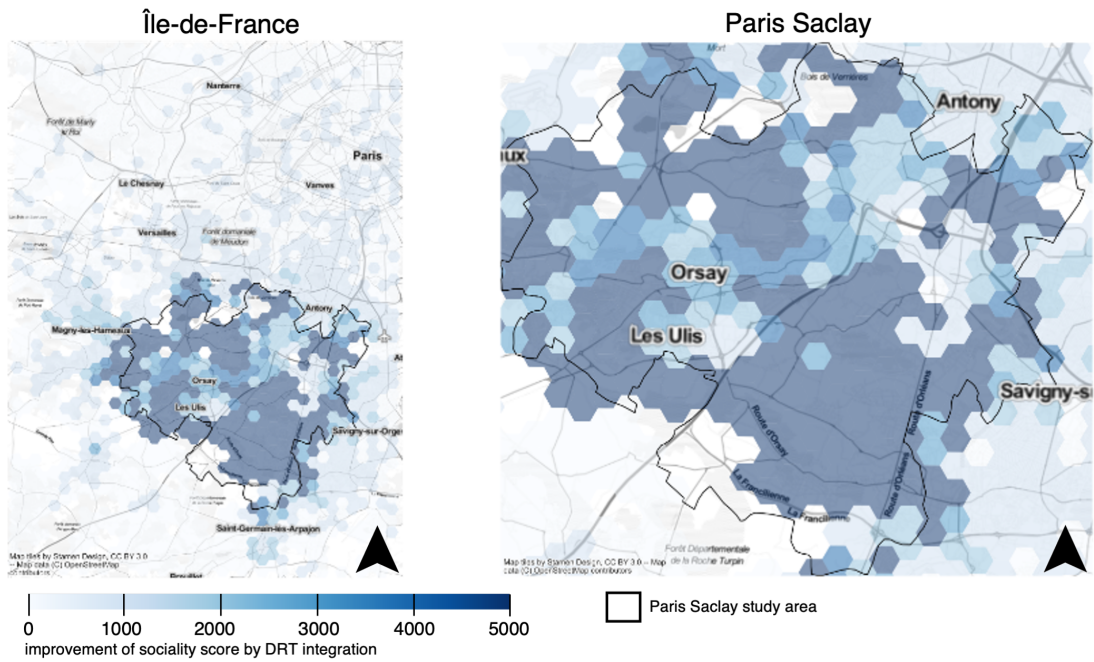


Figure 19: Sociality Score Improvement - Access & Egress - Full Day 05:00 - 23:00

5 CONCLUSIONS

We proposed a method to compute the impact of SMS on accessibility, based on empirical observations of SMS trips. Our method can support transport agencies and authorities in future deployment of SMS. In our future work, we will empirically validate the results by running simulations where we replaced simulated SMS with our estimated virtual trips. Finally, we will apply our method to car- or bike-sharing feeder and, possibly, on observations from real deployments.

ACKNOWLEDGEMENTS

This work has been supported by The French ANR research project MuTAS (ANR-21-CE22-0025-01) and by BayFrance.

REFERENCES

- AA.VV. (2018). Advances in Sensitivity Analysis of Uncertainty to Changes in Sampling Density When Modeling Spatially Correlated Attributes. In B. S. D. Sagar, Q. Cheng, & F. Agterberg (Eds.), *Handbook of mathematical geosciences* (chap. 19). Springer Open.
- Badeanlou, A., Araldo, A., & Diana, M. (2022). Assessing transportation accessibility equity via open data. *hEART*.
- Biazzo, I., Monechi, B., & Loreto, V. (2019). General scores for accessibility and inequality measures in urban areas. *Royal Society open science*.
- Calabrò, G. (2023). Adaptive Transit Design: Optimizing Fixed and Demand Responsive Multi-Modal Transport via Continuous Approximation. In *Transportation Research Part A*.
- Cascetta, E. (2009). *Transportation systems analysis: models and applications*. Springer Science & Business Media.
- Chandra, S., Bari, M. E., Devarasetty, P. C., & Vadali, S. (2013). Accessibility evaluations of feeder transit services. *Transportation Research Part A*.
- Chilès, J.-P., & Desassis, N. (2018). Fifty years of kriging. In B. Daya Sagar, Q. Cheng, & F. Agterberg (Eds.), *Handbook of Mathematical Geosciences*. Springer International Publishing.
- Chouaki, T., Hörl, S., & Puchinger, J. (2023). Towards reproducible simulations of the grand paris express and on-demand feeder services. In *Trb*.
- Chouaki, T., & Puchinger, J. (2021). Agent based simulation for the design of a mobility service in the paris-saclay area. *Transportation Research Procedia*.
- Craig, T., & Shippy, W. (2020). Gtfs flex—what is it and how is it used? *National Center for Applied Transit Technology*.
- Diepolder, S. (2023). *Accessibility of dynamic transport*. Retrieved 01.03.2023, from <https://github.com/severindiepolder/AccessibilityOfDynamicTransport>
- Erhardt, G. D., Roy, S., Cooper, D., Sana, B., Chen, M., & Castiglione, J. (2019). Do transportation network companies decrease or increase congestion? *Science Advances*(5).
- Fortin, P., Morency, C., & Trépanier, M. (2016). Innovative gtfs data application for transit network analysis using a graph-oriented method. *Journal of Public Transportation*.
- GTFS Reference Document*. (2023). Google. Retrieved 01.03.2023, from <https://developers.google.com/transit/gtfs/reference?hl=en>
- Handcock, M. S., & Wallis, J. R. (1994). An approach to statistical spatial-temporal modeling of meteorological fields. *Journal of the American Statistical Association*(426).
- Henao, A., & Marshall, W. E. (2019). The impact of ride-hailing on vehicle miles traveled. *Transportation*(6).

- Kriging interpolation*. (2023). Columbia University. Retrieved from <https://www.publichealth.columbia.edu/research/population-health-methods/kriging-interpolation>
- Le Hasif, C., Araldo, A., Dumbrava, S., & Watel, D. (2022). A graph-database approach to assess the impact of demand-responsive services on public transit accessibility. In *15th acm sigspatial international workshop on computational transportation science*.
- Miller, E. J. (2020). Measuring accessibility: Methods and issues. International Transport Forum.
- Moliński, S. (2022). Pyinterpolate: Spatial interpolation in python for point measurements and aggregated datasets. *Journal of Open Source Software*.
- Nahmias-Biran, B.-h., Oke, J. B., Kumar, N., Lima Azevedo, C., & Ben-Akiva, M. (2021). Evaluating the impacts of shared automated mobility on-demand services: An activity-based accessibility approach. *Transportation*.
- Saeidizand, P., Fransen, K., & Boussauw, K. (2022). Revisiting car dependency: A worldwide analysis of car travel in global metropolitan areas. *Cities*.
- Yakowitz, S. J., & Szidarovsky, F. (1985). A comparison of kriging with nonparametric regression methods. *Journal of Multivariate Analysis*.
- Zhou, M., Le, D.-T., Nguyen-Phuoc, D. Q., Zegras, P. C., & Ferreira Jr, J. (2021). Simulating impacts of automated mobility-on-demand on accessibility and residential relocation. *Cities*.

On allowing endogenous minimum consumption bounds in the Multiple Discrete Continuous Choice Model: An application to expenditure patterns

Andrea Pellegrini¹, John Rose¹

¹ Institute of Transport and Logistics Studies,
Business School, University of Sydney

SHORT SUMMARY

This study develops a novel econometric model that allows for the endogenous identification of minimum bounds on consumption. This is done by combining a censored Tobit model with a Multiple Discrete Continuous Extreme Value model. Whilst the former is employed to identify the minimum consumption of a good/service based upon the socio-demographic characteristics of the consumer, the latter is used to assess multiple discrete continuous consumption patterns. The proposed modelling framework is applied to investigate individuals' expenditure behaviour with the attention being placed on the following expenditure categories: Transport, Shopping, Child Care, Entertainment, Household Bills, and Rent/Mortgage.

Keywords: Multiple Discrete Continuous Decisions, Minimum Consumption Patterns, Expenditure Behavior, Censored Tobit Model

1. INTRODUCTION

Numerous consumption choices result in economic-agents selecting multiple goods or services at the same time. Along with the selection of the type of good to purchase, consumers also usually decide how much of the selected product to consume, with the latter representing a continuous quantity dimension within the underlying decision-making process. These choice situations have thus far been widely analysed in the economics literature via the use of multiple discrete continuous (MDC) demand models. Unlike traditional single discrete choice models, MDC methods allow for assessing interior (goods are assumed to be imperfect substitutes) and corner solutions (goods are treated as perfect substitutes), whilst also accounting for potential satiation effects. The latter effects typically arise with increasing consumption of a good or a service. Since its advent, the Multiple Discrete Continuous Extreme Value (MDCEV) model (Bhat, 2005; 2008) has become the benchmark framework within the transportation literature for analysing individuals' decisions being discrete and continuous in nature. Further, the MDCEV model has been extended in a variety of directions. These are, for example, a) heterogeneity in parameters (Bhat et al., 2016; Shobani et al., 2013), b) flexible utility profiles (Palma and Hess, 2022; Pellegrini et

al., 2021; Bhat, 2018, Bhat et al., 2015), and c) multiple constraints (Mondal and Bhat, 2021; Castro et al., 2012).

A downside of many MDC models is that they fail to account for the impact of minimum consumption patterns for goods/services on the analysis of consumer behaviour. For example, it is reasonable to believe that leisure activities (going to a museum) cannot be performed without devoting a certain minimum amount of time for it, or travelling to holiday destinations cannot be done without individuals spending a certain minimum amount of monetary expenditure. The first attempt to incorporate lower bounds on consumption quantities traces back to Van Nostrand et al. (2012) who recast the utility function formulated in Bhat (2008) to capture the existence of a minimum amount of time allocated to location destination vacation choices (see, Astroza et al., 2017, for an application involving time-use data). Specifically, the authors observed the minimum time that each individual allotted for every destination under investigation and imposed this value as the lower bound when estimating the log-likelihood function. The empirical evidence from the study suggested the implementation of lower bounds on consumption patterns provided more realistic predictions of the time use allocation decisions, whilst also resulting in a better goodness of fit relative to the traditional MDCEV model. Recently, Saxena et al. (2021) proposed an extension of the MDCEV model in which the analyst is able to integrate both lower and upper bounds on consumption choices into a trackable and flexible framework. Despite such improvements, lower bounds on consumption are imposed disregarding the socio-demographic and economics characteristics of the consumers. It is quite unrealistic to assume that households with children versus childless households allot the same minimum time value for entertainment activities. Likewise, house owners versus tenants are unlikely to engage with similar minimum household related expenses. As such, it is likely necessary to exploit individual consumer differences to better understand consumer preference behaviour.

Given the above considerations, the aim of this paper is to develop a model of MDC demand wherein lower bounds on consumptions are specified based upon the demographic information of different decision-makers. To do this, we propose using a censored Tobit regression model to identify minimum consumption patterns for the discrete alternatives, which subsequently serve as lower bounds within a MDCEV model. The developed framework is adopted for investigating consumer expenditure behaviour of 858 Australian residents of New South Wales (NWS), Australia. The expenditure categories involved in this study comprise of nine different categories, consisting of transport, shopping, childcare, entertainment, household bills, and rent/mortgage payments, savings, miscellaneous costs, and other expenditure items.

METHODOLOGY

The framework that we formulate for this study consists of the joint estimation of two empirical methods, namely the censored Tobit model and the MCEV model. Specifically, a censored Tobit model is first employed to the identify the minimum amount of money that each respondent i spent on the expenditure category j , such that

$$y_{ij}^* = \beta_0 + \beta K + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma_j^2) \quad (1)$$

where K is a $N \times T$ matrix that describes the socio-demographic and economic characteristics of the consumers, y_{ij}^* is a latent variable, β is a vector $T \times 1$ of parameters to be estimated, and ϵ_{ij} are independently and identically normally distributed error terms. The amount of money spent by the respondent i on the alternative j is specified such that

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{Lj} < y_{ij}^* < y_{Uj} \\ y_{Lj} & \text{if } y_{ij}^* < y_{Lj} \\ y_{Uj} & \text{if } y_{ij}^* > y_{Uj} \end{cases} \quad (2)$$

In the above equation, the Tobit model is assumed to be bounded between (y_{Lj}, y_{Uj}) , with y_{Lj} being the below censored point whereas y_{Uj} representing the above censored point. By computing the conditional expectation, $y_{ij}^0 = E[y_{ij} | y_{ij}^* \text{ if } y_{Lj} < y_{ij}^* < y_{Uj}, X]$, we are able to endogenously determine the minimum expenditure that respondents made for each alternative j . The estimated minimum expenditure is then integrated into the utility function of the MDCEV model,

$$U_i(\mathbf{Y}) = y_{i1}\psi_1 + \sum_{j=2}^J \delta_{ij}(y_{ij}) \cdot (3)$$

In Equation (3), y_{ij} are inside goods (i.e., the expenditure categories under scrutiny) whilst y_{i1} is the linear outside good whose role is to acknowledge the possibility of consumers spending money on expenditure categories other than those analysed in the analysis. The appeal of a linear formulation for the outside good resides in the fact that the corresponding first derivate equals one and hence drops out in the calculation of the KKT conditions for optimality. Further, $\delta_{ij}(y_{ij})$ can be written such as

$$\begin{aligned} \delta_{ij}(y_{ij}) &= \psi_{ij}y_{ij} \quad \text{if } y_{ij} \leq y_{ij}^0 \\ \delta_{ij}(y_{ij}) &= \psi_{ij}y_{ij}^0 + \gamma_{ij}\psi_{ij}\ln\left(\frac{y_{ij}-y_{ij}^0}{\gamma_{ij}} + 1\right) \quad \text{if } y_{ij} > y_{ij}^0 \end{aligned} \quad (4)$$

where γ_{ij} govern satiation patterns, ψ_{ij} is the baseline marginal utility at the point of zero expenditure, and y_{ij}^0 is the minimum required expenditure of an alternative j (if it is selected). Both γ_{ij} and ψ_{ij} can be further parametrized to incorporate demographic and economic variables attached to the consumer i as below

$$\begin{aligned} \gamma_{ij} &= \exp(\delta r_{ij}) \\ \psi_{ij} &= \exp(\alpha r_{ij} + \tau_j) \end{aligned} \quad (5)$$

where τ_j are independently and identically error terms with a type 1 extreme value distribution

The underlying assumption is that the economic-agent i is assumed to maximize the utility function expressed in Equation (3), $U_i(\mathbf{Y})$, subject to a monetary constraint $y_{i1} + \sum_{j=2}^J y_{ij} = M_i$, where M_i represents the available monetary budget. The optimal expenditure allocations $y_{ij}^*(j = 1, \dots, J)$ can be obtained by forming the Lagrangian function and applying the KKT conditions for optimality, such that

$$\mathcal{L} = U(Y) - \lambda(y_{i1} + \sum_{j=2}^J y_{ij} - M), \quad (6)$$

where λ is the Lagrangian multiplier associated with the specified monetary budget constraint. The KKT conditions for the optimal expenditure allocations for the individual i are given by:

$$\begin{aligned} u'_i(y_{ij}^*) - \lambda &= 0, \text{ if } y_{ij}^* > 0, j = 1, \dots, J \\ u'_i(y_{ij}^*) - \lambda &< 0, \text{ if } y_{ij}^* = 0, j = 1, \dots, J \end{aligned} \quad (7)$$

where $\lambda = \psi_1$, with $\psi_1 = \exp(\tau_1)$

The KKT conditions formalized in Equation (7) can be re-written as:

$$\begin{aligned} V_{ij} + \tau_j &= \psi_1 \text{ if } y_{ij}^* > 0, j = 1, \dots, J \\ V_{ij} + \tau_j &< \psi_1 \text{ if } y_{ij}^* = 0, j = 1, \dots, J \end{aligned} \quad (8)$$

where $V_{ij} = \psi_{ij}$ if $y_{ij} \leq y_{ij}^0$ whilst $V_{ij} = \psi_{ij} \left(\frac{y_{ij} - y_{ij}^0}{\gamma_{ij}} + 1 \right)^{-1}$ if $y_{ij} > y_{ij}^0$

The maximum likelihood estimation method is next adopted for estimating the parameters of the two methodological approaches involved in the optimization process.

2. DATA

Respondents 18 years and older drawn from New South Wales Australia were recruited using the online panel QOR Surveys ([www. https://www.qorsurveys.com.au/](https://www.qorsurveys.com.au/)) between 27th March and 20th April 2022 and asked to complete an online survey associated with environmental issues related to the Murray Darling Basin area. As part of the survey, respondents were asked to provide a summary of their average monthly household expenditure across nine different expenditure categories. The nine expenditure categories consisted of expenditure on transport, shopping, child care, entertainment, household bills, and rent/mortgage payments, savings, miscellaneous costs, and other expenditure items. After providing information regarding actual expenditure, respondents were next asked what would be the absolute minimum amount of expenditure that could be allocated to each category over a given month. As such, the survey captured both actual expenditure as well as the minimum perceived expenditure amount across each expense category.

A total of 2,056 respondents completed the survey. After extensive data cleaning involving removing data from respondents whose total survey response time was considered to be too fast to meaningfully complete the survey, who undertook straight lining behaviour in answering attitudinal questions, and who provided inconsistent responses to survey questions such as suggesting minimum expenditure amounts were necessary that were greater than actual amounts reported, the final sample consists of data collected from 858 respondents. The socio-demographic characteristics of this final sample match the recent 2021 Census values, with the sole exception of being slightly more skewed towards males (54.6% relative to 49.3%).

3. RESULTS

In addition to the proposed model, we also estimated a MDCEV model in which minimum bounds on consumption are imposed without accounting for the differences among consumers (bounds

are exogenous to the model). The goodness of fit measures reveal that the framework developed herein outperforms the model with exogenous bounds on expenditure patterns. Focusing on the empirical findings, we found that large households (with at least one adult) are more likely to spend on utility bills than single-person households. The presence of kids results in households spending more on childcare followed by entertainment and shopping. Full-time workers are inclined to allocate a larger portion of the monetary budget to transportation and utility bills categories. Finally, we conducted a simulation study to compare the forecasting performance of the two models estimated in this study. The evidence suggests that the proposed modelling framework provides more accurate predictions relative to the model where bounds on consumptions are exogenously specified, suggesting that respondents do indeed have different minimum consumption patterns.

4. CONCLUSIONS

In this study, we restrict our attention to the formulation of a novel econometric model that permits the identification of minimum bounds on consumption based upon the sociodemographic and economics characteristics of the decision-makers. Unlike traditional applications where minimum bounds are exogenous to the model, we exploit the features of a censored Tobit model to endogenously determine the minimum amount of money that can be potentially spent by respondents. The estimated minimum consumption value is then used in the MCEV model which is employed to assess individuals' expenditure behaviour.

REFERENCES

1. A., Mondal and C.R. Bhat. A New Closed Form Multiple Discrete-Continuous Extreme Value (MDCEV) Choice Model with Multiple Linear Constraints. *Transportation Research Part B*, (2021) Vol. 147, pp. 42-66.
2. A., Pellegrini, Pinjari, A.R., and R. Maggi. A multiple discrete continuous model of time use that accommodates non-additively separable utility functions along with time and monetary budget constraints. *Transportation Research Part A: Policy and Practice* 144, (2021) 37-53.
3. A., Sobhani, N., Eluru, and A., Faghih-Imani. A latent segmentation based multiple discrete continuous extreme value model. *Transportation Research Part B: Methodological*. 58, (2013) 154-169.
4. C. Van Nostrand, V. Sivaraman, and A.R. Pinjari. Analysis of long-distance vacation travel demand in the United States: A multiple discrete-continuous choice framework. *Transportation*, 40 (1) (2013), pp. 151-171.
5. C.R. Bhat, S. Astroza, and A.C. Bhat. On Allowing a General Form for Unobserved Heterogeneity in the Multiple Discrete-Continuous Probit Model: Formulation and Application to Tourism Travel. *Transportation Research Part B: Methodological*, 86 (2016), pp. 223-249.
6. C.R. Bhat. A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions. *Transportation Research Part B: Methodological*., 39 (8) (2005), pp. 679-707.
7. C.R. Bhat. A new flexible multiple discrete-continuous extreme value (MDCEV) choice model. *Transportation Research Part B: Methodological*., 110 (2018), pp. 261-279.
8. D. Palma, and S. Hess/ Some Adaptations of Multiple Discrete-Continuous Extreme Value (MDCEV) Models for a Computationally Tractable Treatment of

- Complementarity and Substitution Effects, and Reduced Influence of Budget Effects. *Transportation Research Part B: Methodological*, (2022).
9. M. Castro, R. Paleti, and C.R. Bhat. A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. *Transportation Research Part B: Methodological.*, 46 (1) (2012), pp. 253-272.
 10. S Saxena, AR Pinjari, A Roy, and R Paleti. Multiple discrete-continuous choice models with bounds on consumptions. *Transportation Research Part A: Policy and Practice* 149, 237-265.
 11. S., Astroza, A.R. Pinjari, C.R. Bhat, and S.R. Jara-Diaz. A Microeconomic Theory-based Latent Class Multiple Discrete-Continuous Choice Model of Time Use and Goods Consumption. *Transportation Research Record*, (2017) Vol. 2664, pp. 31-41.

MATSim-based assessment of fast charging infrastructure needs for a full-electric passenger car fleet on long-distance trips in Sweden

Hamoun Pourroshanfekr Arabani*¹, Mattias Ingelström²,
Mats Alaküla³, and Francisco J. Márquez-Fernández⁴,

¹ Ph.D. student, Division of Industrial Electrical Engineering and Automation (IEA), Lund University, Sweden

² Ph.D. student, IEA, Lund University, Sweden

³ Professor, IEA, Lund University, Sweden

⁴ Associate Senior Lecturer, IEA, Lund University, Sweden

SHORT SUMMARY

This paper assesses the fast-charging infrastructure requirements to satisfy all long-distance trips by fully electric passenger cars. The main goals of this study are developing an accurate model for electric vehicle (EV) owner trips and their charging behavior, identifying candidate charging station locations and the number of chargers per station for fast charging infrastructure based on the developed model. The transition to EVs is gaining momentum, but the success of this shift relies heavily on the availability and accessibility of charging infrastructure. Several aspects of the fast-charging infrastructure planning problem are investigated based on the developed multi-agent model of EVs' usage using MATSim. The main contribution of this study is the introduction of a novel methodology to identify candidate locations for fast-charging infrastructure needs based on the missing energy event (MEE) in the MATSim EVcontrib and its application to assess the fast-charging infrastructure needs for passenger cars in Sweden.

Keywords: Electric vehicles, Fast charging infrastructure, Long-distance trips, MATSim.

1. INTRODUCTION

There is no doubt that transportation is a major contributor to greenhouse gas emissions, but electric vehicles (EVs) offer solutions for a more environmentally friendly way of transport. The popularity of EVs is growing due to their potentially environmental-friendly energy sources' usage that reduces dependence on fossil fuels. Several factors are preventing EVs from becoming widely used, including the limited range, long recharging time, and lack of charging infrastructure.

The EVs range refers to how far they can travel on a single charge. Many EV owners can charge their EVs at home enough to complete their daily trips. It is estimated that 90% of European Union trips do not exceed 80 km, whereas the typical range of an EV is higher than 200 km (Metias et al., 2022).

Charging times for EVs vary depending on the battery size and charging power used. For a typical EV with a 60kWh battery, it takes just under 8 hours with a 7kW charger to charge it from empty to full, which is suitable for overnight charging, or about 30 minutes with a fast charger (> 150 kW) while on the route.

The availability and accessibility of fast-charging infrastructure are paramount for the widespread adoption of EVs, especially for long-distance travel. Since the cost of such infrastructure is high, it is important to carefully choose their location and size to maximize the number of EVs they serve. There are several studies on the matter.

(Liu et al., 2021) proposes a strategy for placing EV charging stations along German motorways, considering cost and driver satisfaction.

(Baltazar, Vallet, & Garcia 2022) provides a multi-perspective analysis of the potential of EVs for long-distance mobility, highlighting the interest in assessing environmental impacts, user behavior, and EV diversity and proposing general suggestions for EV deployment, including the need for real-world traffic data and consideration of the diversity of EVs and their states of charge (SOC).

(Ge and MacKenzie, 2022) investigates the factors that determine EV users' charging behavior on long-distance trips, using data from a stated choice experiment. The study found that EV drivers' decisions to charge are mostly influenced by their battery SOC and the ability to reach the next station without deviating from their original travel plan, in addition to other secondary factors such as charging cost, time, detour time to reach a station, and amenities at the station.

The main goals of this study are developing an accurate model for EV owner trips and their charging behavior, identifying candidate charging station locations, and the number of chargers per station for fast-charging infrastructure based on the developed model. To achieve that, a novel methodology to identify candidate locations for fast-charging infrastructure is proposed. The developed approach is based on using missing energy event (MEE) in the Multi-Agent Transport Simulation (MATSim) EVcontrib (Horni, Nagel, & Axhausen, 2016) to assess the fast-charging infrastructure needs for passenger cars in Sweden.

2. METHODOLOGY

In this paper, the MATSim tool is used for modeling electric passenger cars' owner trips and their charging behavior based on the previous Sweden case study (Bischoff et al., 2019; Márquez-Fernández et al., 2019; Márquez-Fernández et al., 2021). Using MATSim, every EV movement on the transportation network, its energy consumption while driving, and each charging activity can be modeled and tracked. These results can later be aggregated for each vehicle type/fleet as well as for a certain type of charging infrastructure or for a specific geographical area.

Synthetic population

The Sweden study case is based on the SAMPERS (Sveder, 2002) aggregated travel demand model modified to account only for passenger car flows on roads. To be able to use it in MATSim, the flows in the SAMPERS model have been disaggregated into individual agents using the original origin and destination (OD) information combined with Corine Land Cover Data to determine specific OD locations for each agent (Bischoff et al. 2019). Only trips that are estimated to be longer than 150 km are considered, shorter trips are assumed not to need fast charging.

In cases where the travel distance exceeds 400 km and a round-trip on the same day is improbable, the model results are translated directly, without any adjustments, and for travel distances ranging between 150 km to 400 km, the model included a probability of scheduling a same day return trip for the agent, which decreases as the distance increases.

Regarding departure time selection the following cases are considered: if the distance of the trip is more than 1000 km then the departure time is selected randomly between 8 am and 10 am. For shorter trips, private trips start between 5 am and 12 pm in 70% of cases, between 12 pm and 4 pm in 20% of cases, and only 10% of cases after 4 pm. Business trips start between 5 am and 10 am in 80% of the cases and the rest of the cases start after 10 am.

Charging activity

Apart from the general plans, each agent is also assigned specific charging activities before every iteration in MATSim.

For each long-distance trip, the shortest route and the corresponding energy consumption profile are established. This, together with the capacity and initial SOC of the EV battery allows for determining the locations for the charging activities along the route. Charging activities are included in the plan according to: in 80% of cases, a charging activity is considered when the SOC reaches a value between 20 - 30% (randomly selected); for the rest 20% a charging activity is considered when the SOC reaches a value between 30 - 50% (randomly selected).

Charging activities are integrated into MATSim's activity-based modeling, but they do not provide any positive score (utility function of each agent in the MATSim model). The estimated time required for charging the vehicle defines the duration of each charging activity and in this model, it depends on the desired SOC after charging the EV battery capacity, and the rated power of the charging station.

At this point, the EVs' movements and their charging activities are established in the model, but the location of charging infrastructures and their specifications must be defined. For each charging station the total number of chargers, their type, and rated power must be defined. In the next part, an approach for finding candidate fast-charging infrastructure is introduced.

Candidate Fast Charging Infrastructure

The flowchart of the proposed strategy for finding candidate locations for fast-charging infrastructure is shown in Figure 1. First, it is assumed that there are no charging stations on the road and that all EVs start their long-distance trips with an initial SOC (SOC_{init}) distributed as follows: in 50% of cases $SOC_{init} = [90\% \ 100\%]$, in 30% of the cases $SOC_{init} = [70\% \ 90\%]$ and in 20% of the cases $SOC_{init} = [50\% \ 70\%]$ (randomly assigned except for the first step of proposed approach at Figure 1 where the lowest value is used).

Then the locations where the energy level of the EV battery runs down to zero are identified using the MEE. Afterward, the MEEs are aggregated in an area within a 30km radius. These areas are ranked based on the frequency of MEEs and the first 100 top ranks of them are selected as the selected candidate locations for charging stations. In the next step, the developed MATSim model is run with these 100 charging stations in place, and their impact is assessed based on the % of successful trips (the long-distance trips do not contain any MEEs). If this percentage is lower than 90%, more fast-charging station is added and the simulation reruns again until the required performance is achieved.

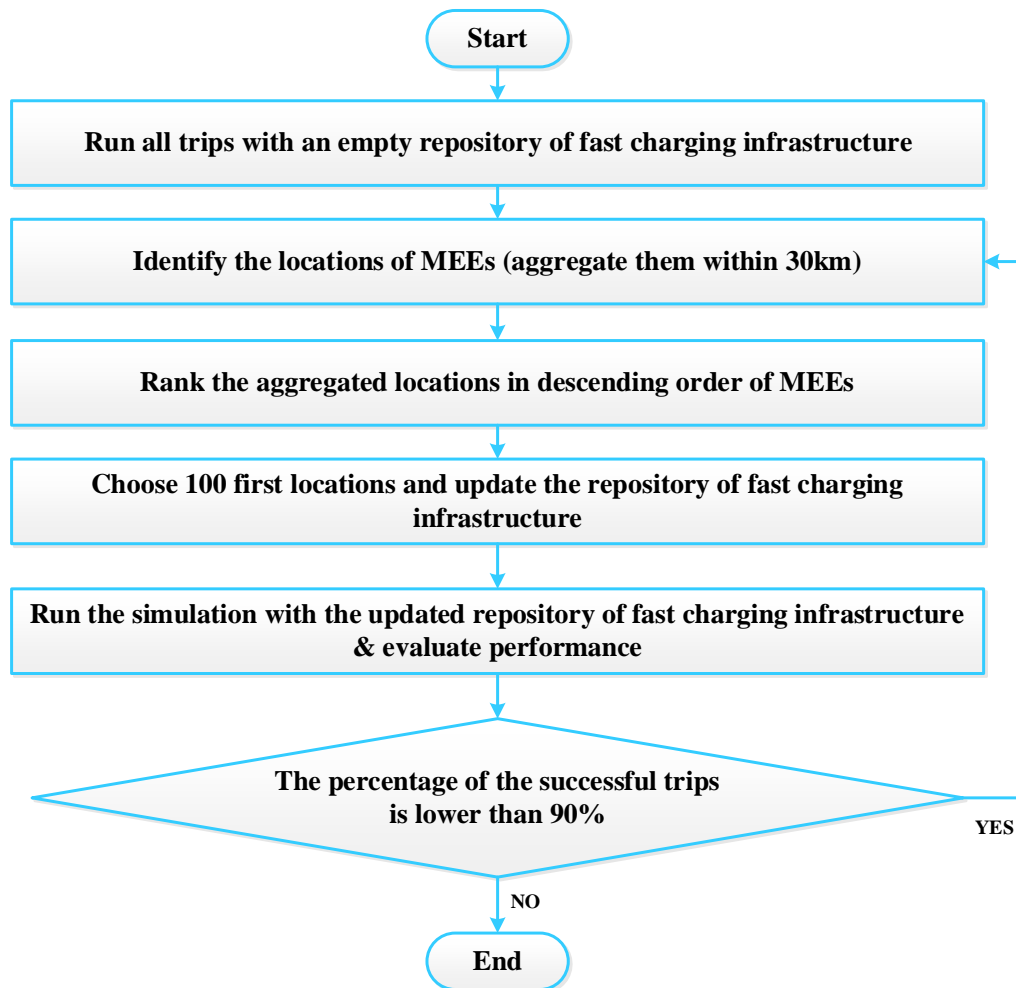


Figure 1: Flowchart of the proposed approach for identifying fast charging infrastructure locations.

Figure 2. shows the distribution of aggregated MEEs in a hexagonal network within a 30 km radius. This distribution results from running the developed MATSim model without any fast-charging infrastructure. As expected, a high number of MEEs occur on the route between Gothenburg (the second-largest city in Sweden) and Stockholm (the capital and largest city of Sweden). There is also a high number of MEEs near the border between Sweden and Denmark and around Malmö (the third-largest city in Sweden).

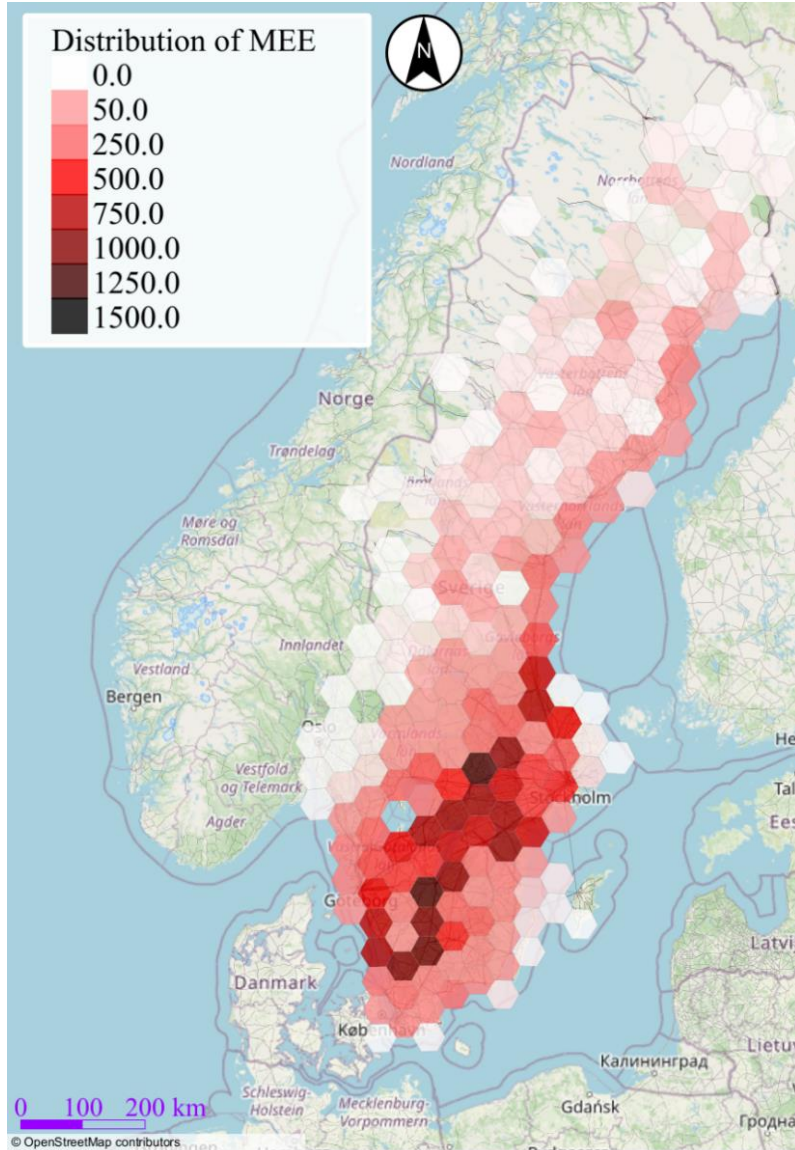


Figure 2: The Distribution of MEEs in the hexagonal network with a 30 km radius.

3. RESULTS AND DISCUSSION

In the simulation setup, the first part that must be defined is the mix of EVs fleet and their related specification which are mentioned in Table 1. The initial energy level of EV batteries (SOC_{init}) and the starting time of the trips are already mentioned in previous sections.

Table 1: Assignment of EVs

Vehicle Type	Battery Capacity kWh	Fleet Share %
Small	60	15
Medium	80	50
SUV	100	35

However, to make the simulation more manageable, a standard practice of utilizing a 10% sample of the population (the total number of trips in the MATSim model) is employed.

In the next step, the location of charging infrastructures must be introduced as input to the model. In Figure 3, all candidate locations are achieved based on the first step of the proposed approach (no-charging station scenario of developed MATSim model) and the 100 selected locations are shown based on the distribution of aggregated of MEEs in Fig 2. The total number of candidate locations is equal to 350. For the first iteration, charging stations are placed at the 100 locations with the highest number of MEEs. Each of these stations is equipped with an unlimited number of chargers, so every vehicle reaching the stations will be allowed to charge.

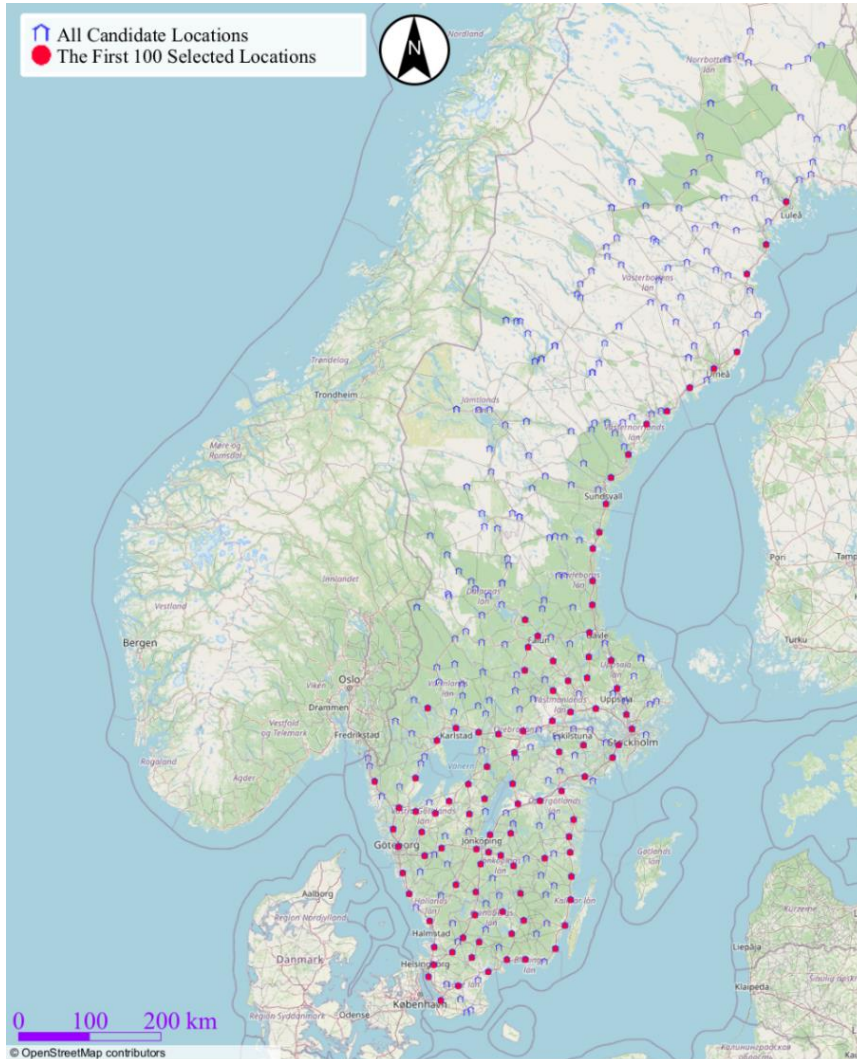


Figure 3: The candidate location and the 100 top-rank selected locations for fast-charging infrastructure based on the distribution of MEEs.

For subsequent iterations, more fast-charging infrastructure locations are added to the charging infrastructure repository of the MATSim model, by a step of 100 new fast charging stations at each iteration. These 100 new fast-charging stations are selected following the same procedure, based on the distribution of MEEs resulting from the previous iteration.

In Table 2, the percentage of successful trips before and after adding candidate fast-charging infrastructures at each iteration of the proposed methodology is shown.

Table 2: The percentage of successful trips before and after adding fast charging infrastructure with an unlimited number of chargers.

Number of fast-charging stations	Percentage of successful trips %
0	9.43
100	85.56
200	88.90
300	89.04
400	89.07

If all candidate locations (350 locations in Figure 3) for fast-charging infrastructure in the first step of the proposed approach are considered, then the percentage of successful trips is equal to 89.28. The reason for the higher percentage of the successful trip with all candidate locations based on the no-charging station scenario compared to the result of 400 fast charging stations based on the proposed method in Figure 1 is that in the no-charging scenario, the distribution of initial SOC is set to lower bound of each case, then the candidate locations are found, but in the proposed approach the distribution of initial SOC is set based on the mentioned range for each case and the new candidate locations are found based on the higher value of initial SOC, therefore since the charging activities are initiated at the higher value of 20% EVs battery capacity the newly found candidate location cannot cover the trips that are started with the lower value of initial SOC and also the improvement in the percentage of successful trips decreases while the number of the charging station is increased.

Another important reason for this result is the consideration of the unlimited number of chargers at each fast-charging station.

Based on the results, it is observed that with 200 fast-charging infrastructure locations almost 90% of long-distance trips by passenger electric cars can be satisfied if there is an unlimited number of chargers in each location.

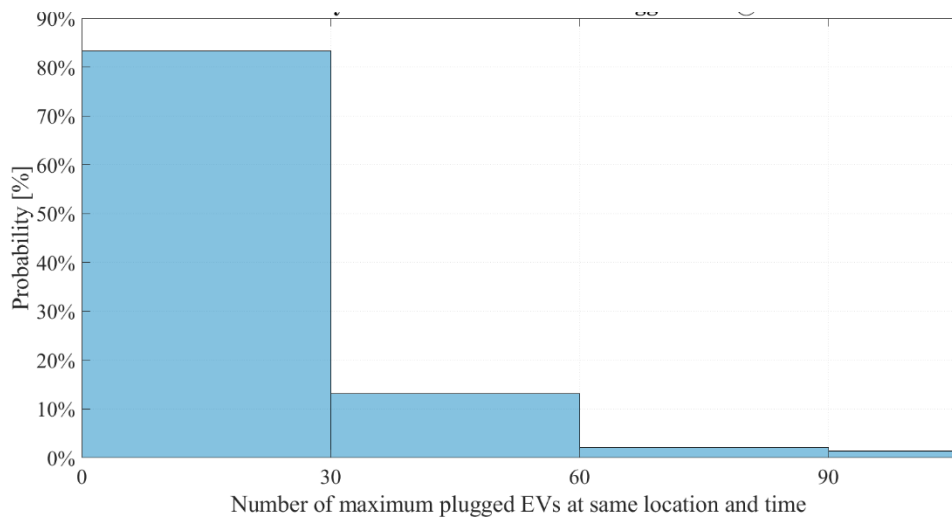


Figure 5: The Distribution of the maximum number of simultaneously charging EVs at a fast-charging station.

As can be seen in Figure 5, in more than 80% of cases the maximum number of EVs that are charging at the same fast charging station location simultaneously is not more than 30 in the scenario with 400 charging stations with unlimited chargers at each location.

Therefore, the limitation of the number of chargers is added to the proposed methodology with the consideration of a maximum of 30 chargers at each fast-charging station in the real world. Since 10% of all long-distance trips above 150km are considered in the MATSim model, each fast-charging infrastructure location is equipped with 3 chargers.

Table 3 displays the success rate percentages of trips before and after the inclusion of candidate fast-charging infrastructures that consist of three chargers at each charging station in every iteration of the proposed methodology.

Table 3: The percentage of successful trips before and after fast charging infrastructure with 3 chargers on each location

Number of fast-charging infrastructure	Percentage of successful trips %
0	9.43
100	46.93
200	53.71
300	59.29
400	61.41

If all candidate locations (350 locations in Figure 3) for the fast-charging infrastructure in the no-charging station scenario are considered with 3 chargers on each location, then the percentage of successful trips is equal to 54.40.

Based on the results of Table 3, after considering 3 chargers at each location the percentage of successful trips based on the proposed methodology (adding new charging stations based on the iterative method Figure 1) is higher compared to considering all found locations in the no charging case is higher.

4. CONCLUSIONS

This paper presents a novel methodology based on the missing energy event concept of MATim EVcontrib to assess the fast-charging infrastructure needs for fully electric passenger cars on long-distance trips in Sweden. The proposed methodology accurately models EV owner trips and charging behavior, identified candidate charging station locations, and assessed them and the number of chargers per station required for fast-charging infrastructure. The results highlight the importance of investing in fast-charging infrastructure to promote the widespread adoption of EVs and pave the way for a sustainable transportation system.

Future research will improve the accuracy of our methodology by refining the synthetic population of the MATSim model for long-distance trips by electric passenger cars. This can involve developing a more detailed representation of EV owner demographics, travel patterns, and charging behavior, and incorporating data on the spatial distribution of charging infrastructure. Such research will provide a more comprehensive assessment of the fast-charging infrastructure requirements for EVs and inform the planning and deployment of charging infrastructure.

REFERENCES

Baltazar, J., Vallet, F., & Garcia, J. 2022. A model for long-distance mobility with battery electric vehicles: a multi-perspective analysis. *Procedia CIRP*, Vol. 109, No. 5, pp. 334-339.

Bischoff J., Márquez-Fernández F.J., Domingues-Olavarría G., Maciejewski M., Nagel K. 2019. Impacts of vehicle fleet electrification in Sweden—a simulation-based assessment of long-distance

trips. In 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), 5-7 June, Cracow, Poland, pp. 1-7.

Ge, Y. and MacKenzie, D., 2022. Charging behavior modeling of battery electric vehicle drivers on long-distance trips. *Transportation Research Part D: Transport and Environment*, Vol. 113, No. 12, pp. 103490.

Liu, J., Peper, J., Lin, G., Zhou, Y., Awasthi, S., Li, Y. and Rehtanz, C., 2021. A planning strategy considering multiple factors for electric vehicle charging stations along German motorways. *International Journal of Electrical Power & Energy Systems*, Vol. 124, No. 1, pp. 106379.

Márquez-Fernández F. J., Bischoff J., Domingues-Olavarría G., Alaküla, M. 2019. Using multi-agent transport simulations to assess the impact of EV charging infrastructure deployment. In 2019 IEEE transportation electrification conference and expo (ITEC), 19-21 June, Detroit, USA, pp. 1-6.

Márquez-Fernández, F.J., Bischoff, J., Domingues-Olavarría, G. and Alaküla, M., 2021. Assessment of future EV charging infrastructure scenarios for long-distance transport in Sweden. *IEEE Transactions on Transportation Electrification*, Vol. 8, No. 1, pp. 615-626.

Metais, M.O., Jouini, O., Perez, Y., Berrada, J. and Suomalainen, E., 2022. Too much or not enough? Planning electric vehicle charging infrastructure: A review of modeling options. *Renewable and Sustainable Energy Reviews*, Vol. 153, No. 1, pp. 111719.

Sveder, G.J., 2002. The new Swedish national model—SAMPERS: System and validation. *National transport models: Recent developments and prospects*, pp.93-100. Springer

Applying a latent class cluster analysis to identify consumer segments of electric vehicle charging styles

Elham Hajhashemi ^{*1}, Patricia Sauri Lavieri ², Neema Nassir ³

¹ PhD student, Department of Infrastructure, The University of Melbourne, Australia

² Lecturer, Department of Infrastructure, The University of Melbourne, Australia

³ Senior Lecturer, Department of Infrastructure, The University of Melbourne, Australia

SHORT SUMMARY

Electric vehicle market growth makes understanding user charging behaviour essential for policy design and EV adoption facilitation. In this study, we examined the heterogeneity in charging preferences of 994 respondents across Australia, using a latent class cluster model that considers indicators of charging behaviour as outcomes of interest. We used sociodemographic characteristics, travel needs, and EV adoption status as covariates to predict class membership. Our findings identify five segments of consumers with distinct charging style preferences: cost-sensitive planners, cost-sensitive on-demanders, predictability seekers, flexibility seekers, and indifferent late adopters. We provide targeted policies for each segment based on their charging style and profile, aimed at facilitating EV adoption and meeting their charging needs. Our results suggest that two broad categories of action are necessary to facilitate EV adoption and meet charging needs of upcoming EV users: improving EV-related knowledge and providing economical home charging options.

Keywords: Charging style, electrification and decarbonization of transport, latent class cluster analysis

1. INTRODUCTION

Transitioning to Electric Vehicles (EVs) from Internal Combustion Engine Vehicles (ICEVs) requires significant behavioural changes and increased cognitive effort from consumers, as charging decisions are multidimensional, involving scheduling, location and charger type choice, and highly variable prices. While current activity-based demand models tend to assume that drivers deliberate about charging before or after every trip based on the battery state of charge and charger availability, this assumption may not be behaviourally realistic for many EV users. That is, users may reduce the cognitive load of charging decisions by using heuristics, relying on daily routine cues and habit. In this sense, segmenting individuals based on their behavioural patterns can be a more effective way of modelling charging behaviour.

The notion of “style”, as in “lifestyle” (Talvitie, 1997), “mobility style” (Lanzendorf, 2002), “modality style” (Vij et al., 2013), has been adopted by researchers to represent behavioural patterns together with their underlying motivations and attitudes towards different aspects of life, travel, and/or modal preferences. Analogously, the term “charging style” can be used to represent charging behaviour patterns (including heuristics and cues that individuals may use) associated with underlying personal subjective orientations. Yet, only a couple of studies have explored this idea.

Franke et al. (2013) proposed that EV users adopt a charging style as a preferred coping strategy to interact with the limited battery resources of their vehicles. They developed the concept of user battery interaction style (UBIS) to measure differences in coping strategies related to charging. An individual with high UBIS makes charging decisions based on the vehicle's state of charge, while people with low UBIS use other cues to make a charging decision, such as routine or opportunity. Subsequently, Daina et al. (2015) used this concept of UBIS to predict charging demand. Although UBIS can capture the coping strategies of EV users, it does not define charging style as a representation of a general pattern of charging. That is, it requires additional factors to generate a prediction of charging choices.

Considering that a comprehensive construct representing charging styles can benefit integrated energy and transport demand models, the current study utilises empirical data from a survey with EV owners and potential owners together with a latent class cluster analysis (LCCA) approach to (1) identify classes of electric vehicle charging styles, (2) define user profiles for each style, and (3) provide tailored policy recommendations to facilitate charging and potentially increase EV adoption among consumers with different styles.

2. METHODOLOGY

The conceptual framework of this study, as shown in Figure 1, aims to classify current and prospective EV users into different classes based on their charging style. The dimensions of charging style were identified based on literature review and include three main categories: charging attributes, coping strategy with battery resources, and risk attitude.

Charging attributes include three dimensions representing user preferences regarding charging: (1) time regularity, (2) location, (3) trade-offs between cost and charging speed, and cost and perceived convenience. To measure individuals' coping styles related to charging, a scale was developed based on UBIS, measuring the trade-off between battery level and opportunity, and battery level and routine. Risk attitude was not directly considered, but it was inferred through the trade-off between planning for charging and deciding on the go. The framework also takes into account socio-demographic characteristics, EV adoption cohort, and travel needs to understand the charging profile of people in different classes.

2.1 Latent class cluster analysis

LCCA is used to reveal the charging style classes using poLCA package (Linzer and Lewis, 2011). LCCA groups individuals into distinct charging style classes based on observable charging preferences and individual characteristics. The model has a measurement component, which links the underlying latent categorical variable to its indicators, and a structural component, which defines the relationships between explanatory covariates (active covariates) to determine class membership. The model simultaneously estimates these sub-models using a probabilistic approach. Inactive covariates provide additional profiling of the identified classes.

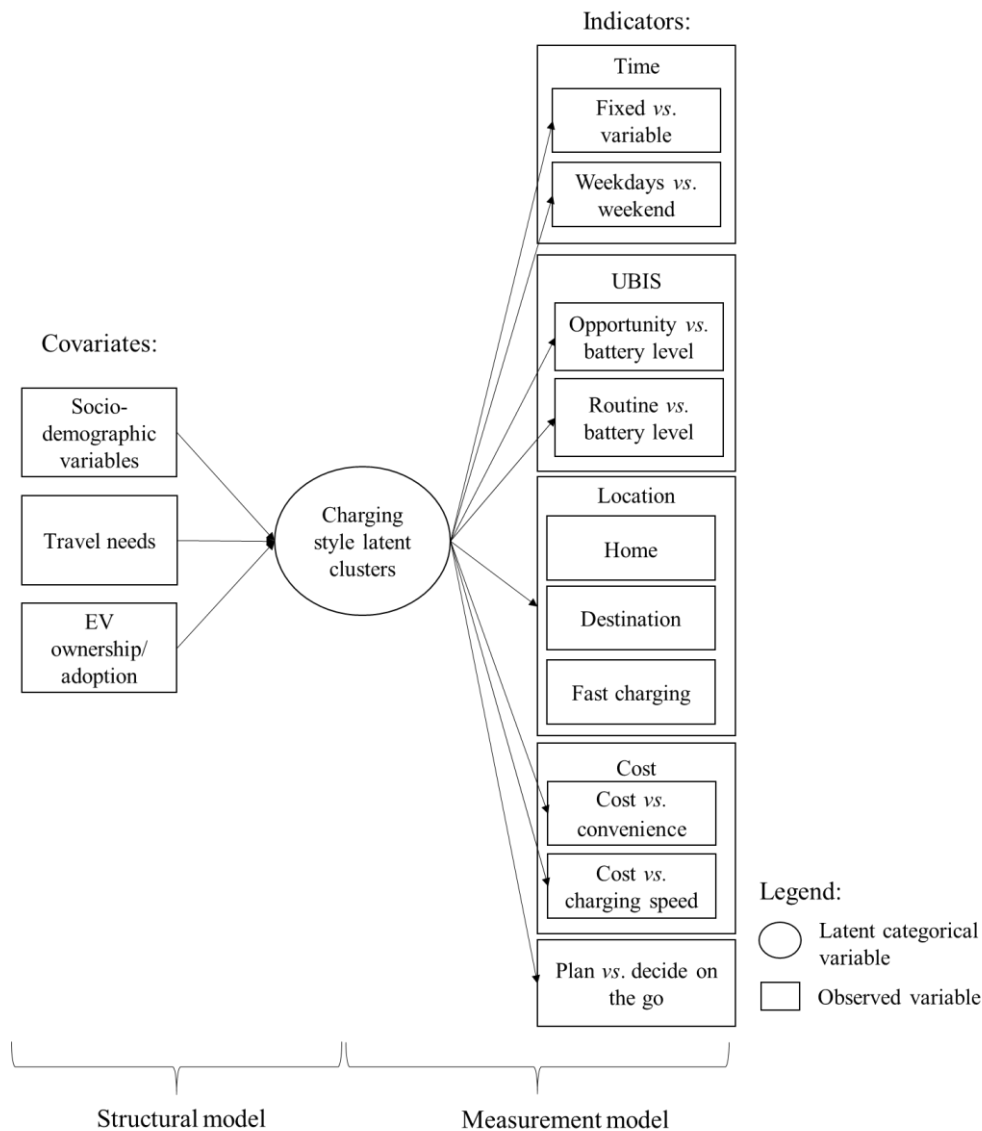


Figure 1: Conceptual Framework

2.2 Data collection and sample

We obtained our data through an online survey conducted between July and August of 2021 as part of the EV Integration (2020-2022) project in Australia (University of Melbourne, 2022). The sampling strategy aimed for around 10% responses from EV drivers and a sample of ICEV drivers representative of the Australian driver population. The final sample size comprised 994 observations, including 97 EV drivers. The survey collected information on socio-demographic characteristics, travel needs, EV ownership and intentions to purchase, and charging preferences. We used three cohorts of EV adoption: EV owners, early majorities (likely to own an EV within five years), and late majorities (likely to own an EV within ten years or have no plans to own one). Detailed information about the data can be found in the project's report (Lavieri and Oliveira, 2021).

The last column of Table 2 shows descriptive statistics of the sample. While the sample is not representative of the overall driving-age population due to the oversampling of EV owners, the sub-sample of ICEV drivers is. Relative to the driving-age population, EV owners in the sample are more likely to be men, in 35-54 age category, have tertiary education, be employed full-time, and have high income. Furthermore, they drive twice the national average of annual distances driven per person.

3. RESULTS AND DISCUSSION

Models with 2 to 6 classes were compared using AIC and BIC, and the five-class model was chosen as the best fit. Tables 1 and 2 present the behaviour and the profile of each class while Table 3 presents the class membership model results.

3.1 *Charging styles and profiles*

Class 1: 26.7% of the total sample are cost-sensitive planners who prefer home charging and prioritize cost over convenience or speed. This group mostly charges based on routine and opportunity. They usually charge on weekdays and plan their charging in advance. This class has the highest proportion of mid-low-income households, the largest share of households with solar panels, and high vehicle ownership.

Class 2: Comprising 27.5% of the total sample, cost-sensitive "on-demanders" prefer home charging, prioritize cost over convenience, and charge based on battery level. This class has a high proportion of homeowners, households with off-street parking, and late majorities (in terms of EV adoption).

Class 3: Named as predictability seekers, they make up 18.9% of the total sample and prioritize convenience over cost. They tend to charge their vehicles based on routine and opportunity. They prefer home charging. They have the highest average weekly distance travelled and plan their charging in advance to meet their travel needs. This class has the highest proportion of high-income households, women, and those in early majority cohort.

Class 4: Flexibility seekers comprise 18.5% of the total sample and prioritize charging convenience and speed over cost. They charge based on battery level. This group prefers fast charging and destination charging the most among the five classes. This class has the highest proportion of men, young individuals, high-income earners, highly educated individuals, and EV owners.

Class 5: known as Indifferent Late Adopters, they make up 8.5% of the sample. They have no defined charging preferences yet. This group has the highest proportion of individuals who are unemployed or not in the workforce, low-income households without solar panels and/or off-street parking, or individuals living in rented properties. They are mostly among the late adopter cohort.

Table 1: Summary Statistics of Indicators

	Class 1	Class 2	Class 3	Class 4	Class 5	Sample total
Class name	<i>Cost-sensitive planners</i>	<i>Cost-sensitive "on-demanders"</i>	<i>Predictability seekers</i>	<i>Flexibility seekers</i>	<i>Indifferent late adopters</i>	
Class share (%)	26.7	27.5	18.9	18.5	8.5	100
Class size (n)	271	274	185	179	85	994
Opportunity vs. battery level:						
<i>Opportunity</i>	80.2%	17.9%	43.1%	30.1%	2.2%	40.2%
<i>Equal</i>	13.9%	46.0%	32.6%	17.0%	94.6%	33.6%
<i>Battery level</i>	5.9%	36.1%	24.4%	52.9%	3.3%	26.2%
Routine vs. battery level:						
<i>Routine</i>	96.1%	32.0%	57.9%	18.6%	0.0%	48.9%
<i>Equal</i>	2.8%	29.4%	17.0%	10.1%	95.6%	21.9%
<i>Battery level</i>	1.1%	38.6%	25.1%	71.4%	4.4%	29.2%
Charging location preference:						
<i>Home</i>	70.7%	79.0%	74.0%	51.1%	57.8%	68.9%
<i>Destination charging</i>	10.9%	8.2%	9.1%	17.9%	16.2%	11.6%
<i>Fast charging</i>	18.3%	12.8%	16.9%	31.0%	26.0%	19.5%
Cost vs. convenience:						
<i>Cheapest</i>	90.4%	83.6%	2.4%	8.0%	0.0%	49.1%
<i>Equal</i>	1.7%	10.0%	54.4%	7.4%	89.0%	22.3%
<i>Most convenience</i>	7.9%	6.4%	43.2%	84.6%	11.0%	28.6%
Cost vs. charging speed:						
<i>Cheapest</i>	95.3%	97.5%	19.7%	11.6%	0.5%	58.1%
<i>Equal</i>	2.9%	2.5%	61.9%	12.4%	92.4%	23.2%
<i>Fastest</i>	1.8%	0.0%	18.4%	76.1%	7.1%	18.6%
Day:						
<i>Weekdays</i>	84.2%	26.7%	63.4%	36.1%	2.2%	48.7%
<i>Equal</i>	9.5%	50.7%	28.6%	20.6%	94.4%	33.6%
<i>Weekends</i>	6.3%	22.7%	8.0%	43.3%	3.4%	17.7%
Time of day:						
<i>Same</i>	91.1%	38.2%	58.7%	37.1%	0.0%	52.8%
<i>Equal</i>	6.6%	41.8%	25.0%	22.0%	99.3%	30.4%
<i>Different</i>	2.3%	20.1%	16.3%	40.9%	0.7%	16.8%
Plan vs. decide on the go:						
<i>Plan</i>	85.8%	61.5%	51.3%	36.8%	0.0%	56.3%
<i>Equal</i>	10.0%	34.3%	35.2%	15.8%	98.7%	30.0%
<i>Decide on the go</i>	4.2%	4.2%	13.4%	47.4%	1.3%	13.7%

Table 2: Summary Statistics of Covariates

Covariates	Class 1	Class 2	Class 3	Class 4	Class 5	Sample total
EV ownership and adoption status:						
<i>Early majority</i>	39.9%	35.5%	40.6%	31.8%	33.6%	36.8%
<i>EV owners</i>	4.2%	1.9%	12.9%	29.3%	3.5%	9.8%
<i>Late majority</i>	55.9%	62.6%	46.6%	38.9%	62.9%	53.4%
Gender:						
<i>Female</i>	52.9%	43.4%	55.5%	32.1%	49.7%	46.7%
<i>Male</i>	47.1%	56.6%	44.5%	67.9%	50.3%	53.3%
Age:						
<i>18 to 34</i>	29.7%	20.0%	23.1%	36.3%	33.7%	27.4%
<i>35 to 54</i>	33.8%	35.7%	38.5%	42.1%	27.5%	36.2%
<i>55 and older</i>	36.5%	44.3%	38.4%	21.6%	38.8%	36.4%
Income:						
<i>\$100,000 or more</i>	41.0%	43.9%	52.5%	73.4%	46.3%	50.4%
<i>\$35,000 to \$99,999</i>	42.2%	41.6%	31.4%	14.1%	36.1%	34.3%
<i>Less than \$34,999</i>	16.8%	14.5%	16.1%	12.5%	17.6%	15.3%
Education: <Inactive>						
<i>Bachelor and higher</i>	35.8%	40.6%	37.9%	59.1%	38.5%	42.1%
<i>Below bachelor</i>	64.2%	59.4%	62.1%	40.6%	61.5%	57.9%
Employment status: <Inactive>						
<i>Full-time</i>	40.2%	36.0%	44.0%	65.4%	43.3%	44.7%
<i>Not in workforce or unemployed</i>	35.3%	41.0%	34.8%	25.9%	42.3%	35.6%
<i>Part time</i>	24.5%	23.1%	21.3%	8.7%	14.4%	19.7%
Family composition:						
<i>Have children</i>	43.4%	32.3%	27.3%	52.8%	30.15	37.9%
<i>No children</i>	56.6%	67.7%	72.7%	47.2%	69.9%	62.1%
Average number of cars in household	1.81	1.78	1.77	1.54	1.43	1.71
Living situation: <Inactive>						
<i>Own</i>	67.4%	74.6%	74.1%	72.8%	65.0%	71.4%
<i>Rent</i>	32.6%	25.4%	25.9%	27.2%	35.0%	28.6%
Building type: <Inactive>						
<i>Flat or apartment</i>	10.3%	7.2%	9.8%	15.2%	10.4%	10.3%
<i>Separate house</i>	69.7%	76.5%	66.1%	59.4%	74.7%	69.4%
<i>Townhouse</i>	9.8%	7.8%	12.8%	17.2%	8.4%	11.1%
<i>Other</i>	10.2%	8.5%	11.3%	8.2%	6.5%	9.2%
Having off-street parking:						
<i>No</i>	10.2%	7.1%	15.9%	17.3%	23.9%	12.9%
<i>Yes</i>	89.8%	92.9%	84.1%	82.7%	76.1%	87.1%
Having solar panels:						
<i>No</i>	63.1%	67.7%	64.8%	73%	75%	67.5%
<i>yes</i>	36.9%	32.3%	35.2%	27%	25%	32.5%
Average typical weekly distance travelled (km)	229.0	154.3	246.8	237.2	174.4	208.8
Average time window (hour)	25.0	28.3	25.5	23.8	29.0	26.1
Solar Panel condition: <Inactive>						
<i>Do not have</i>	40.6%	49.2%	47.4%	56.6%	69.2%	49.6%
<i>Already have</i>	36.9%	32.3%	35.2%	27.0%	25.0%	32.5%
<i>Will adopt if buy an EV</i>	22.5%	18.5%	17.4%	16.4%	5.8%	17.9%

3.2 Class membership model

The cost-sensitive planners serve as the reference category, and coefficients are interpreted accordingly. In summary, belonging to:

- Class 2 is more likely for individuals with high income and no children, but less likely for those who drive long distances.

- Class 3 is more likely for EV owners, those aged 35 or more, and those with income between \$35,000 to \$99,999 and no children. Less likely for those with off-street parking.
- Class 4 is more likely for EV owners, males, and those with income between \$35,000 to \$99,999 or \$100,000+, but less likely for those with more cars or off-street parking and solar panels.
- Class 5 is more likely for those with income between \$35,000 to \$99,999 and less likely for those with more cars or off-street parking.

Table 3: Class Membership Model

Covariates	Class 2 vs. class 1	Class 3 vs. Class 1	Class 4 vs. Class 1	Class 5 vs. class 1
	Coef. (t-stat)	Coef. (t-stat)	Coef. (t-stat)	Coef. (t-stat)
Intercept	-0.690(-1.042)	-0.621(-0.903)	1.805(2.644)	1.343(1.815)
EV adoption status (reference: early majority)				
<i>EV owners</i>	-	1.174(2.320)	2.269(4.753)	-
<i>Late majority</i>	-	-	-	-
Gender (reference: female)				
<i>Male</i>	-	-	0.700(3.044)	-
Age (reference: 18 to 34)				
<i>35 to 54</i>	-	0.787(2.161)	-0.631(-1.800)	-
<i>55 and older</i>	-	0.819(2.343)	-	-
Income (reference: Less than \$34,999)				
<i>\$35,000 to \$99,999</i>	-	0.626(2.132)	1.497(4.114)	0.623(1.960)
<i>\$100,000 or more</i>	0.658(1.912)	-	0.784(1.685)	-
Family composition (reference: have children)				
<i>Do not have children</i>	0.532(2.134)	1.076(3.657)	-	-
Average number of cars in household	-	-	-0.407(-2.472)	-0.659(-2.592)
Having off-street parking (reference: no)				
<i>Yes</i>	-	-0.663(-1.835)	-0.706(-1.916)	-0.949(-2.317)
Having solar panel (reference: no)				
<i>Yes</i>	-	-	-0.763(-2.572)	-
Typical weekly distance (km)	-0.002(-3.265)	-	-0.001(-1.998)	-

3.3 Policy recommendation

Our study found that EV users may exhibit diverse charging behaviours, and therefore we discuss tailored policies that could be targeted at each charging style segment. We conducted a literature review of pertinent policies and identified three main categories of recommendations that could be customised to each class: 1) financial and regulatory support for residential charging, 2) financial and regulatory support for solar charging, and 3) educational campaigns.

Class 1:

1. Offer financial incentives to landlords for installing level 2 chargers or low-interest loans to tenants for home charger installation. Streamline permitting processes for installation in rental properties.
2. Offer tax credits or incentives for solar panel installation and promote shared solar programs. Offer bundled incentives for EV and solar purchases to encourage the adoption of both technologies.
3. Provide information about the financial incentives available for EV purchase and home and solar charging, along with the long-term cost savings associated with solar charging.

Class 2:

1. Offer cash rebates or discounts for EV adoption and home charger installation.
2. Offer tax credits and incentives for solar panel purchase and installation.
3. Offer educational campaigns that raise awareness about financial incentives for EV purchases, home charging, and solar charging.

Class 3:

1. Allocate curbside or public charging options for those without access to off-street parking.
2. Offer EV and solar panel bundles, emphasizing their convenience and environmental benefits. Encourage community solar programs, solar panel installation, and flexible work policies.
3. Educate consumers on the increasing driving range of EVs and how home charging can adequately meet their travel needs. Highlight the expanding network of public chargers as a backup. Additionally, emphasize the long-term benefits associated with solar charging.

Class 4:

1. Facilitate approval for home charging installation and offer curbside or public charging for those without access to off-street parking.
2. Programs that promote the adoption of solar batteries and encourage weekend charging, as well as support working from home to overcome the limited flexibility barrier for solar charging in this group.
3. Highlight the increasing driving range of EVs, educate consumers on the adequacy of home charging for their travel needs, and emphasize the growing network of public chargers as a backup.

Class 5:

Educational campaigns to increase the knowledge about EVs is a prerequisite for this class, which does not even consider EV adoption in a distant future. This can be achieved through EV festivals, test drives, and incentives, while also making them aware of available support for charging their EVs, including financial and regulatory assistance.

4. CONCLUSIONS

In this study, we performed a LCCA to identify five distinct charging style classes among current and prospective EV users in Australia. Indifferent late adopters have no preferred charging style, while flexibility seekers prioritize speed, convenience, and battery level. Predictability seekers have a fixed charging time and location preference. The largest classes, cost-sensitive planners and on-demanders, prioritize cost savings and home charging. Based on the likely behavioural pattern of each class, as EV adoption continues to grow, home charging will become increasingly crucial, and offering an affordable home charging option is key to meeting the needs of many upcoming EV adopters.

REFERENCES

- DAINA, N., POLAK, J. W. & SIVAKUMAR, A. J. T. R. R. 2015. Patent and latent predictors of electric vehicle charging behavior. *2502*, 116-123.
- FRANKE, T., KREMS, J. F. J. T. R. P. F. T. P. & BEHAVIOUR 2013. Understanding charging behaviour of electric vehicle users. *21*, 75-89.
- LANZENDORF, M. 2002. Mobility Styles and Travel Behavior: Application of a Lifestyle Approach to Leisure Travel. *Transportation Research Record*, 1807, 163-173.
- LAVIERI, P. S. & OLIVEIRA, G. J. M. 2021. Electric Vehicle Charging Consumer Survey: Insights Report. Technical Report, Faculty of Engineering and Information Technology, The University of Melbourne.
- LINZER, D. A. & LEWIS, J. B. 2011. poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, 42, 1 - 29.
- TALVITIE, A. 1997. Things planners believe in, and things they deny. *Transportation*, 24, 1-31.
- UNIVERSITY OF MELBOURNE. 2022. *EV Integration* [Online]. Available: <https://electrical.eng.unimelb.edu.au/power-energy/projects/ev-integration> [Accessed].
- VIJ, A., CARREL, A. & WALKER, J. L. 2013. Incorporating the influence of latent modal preferences on travel mode choice behavior. *Transportation Research Part A: Policy and Practice*, 54, 164-178.

Modelling the impact of activity duration on utility-based scheduling decisions: a comparative analysis

Janody Pougala¹, Tim Hillel², and Michel Bierlaire¹

¹Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, Switzerland

²Behaviour and Infrastructure Group, University College London, United Kingdom

SHORT SUMMARY

There exist two major categories of activity-based models: on one side, utility-based or econometric models are founded on the principles of random utility maximisation, and use discrete choice modelling techniques to solve activity-based problems. On the other hand, rule based approaches refute the assumption that decision-makers are perfect optimisers and their activity-travel behaviour is the product of context-dependent rules. Recently hybrid models (e.g. OASIS) combine both approaches, to keep the flexibility and theoretical robustness of utility-based models, with the addition of spatio-temporal constraints which increase the behavioural realism and simplify the estimations. However, hybrid models suffer from two main issues: specifying a utility function that accurately reflects the decision-making process, and estimating parameters in a highly complex space due to the constraints. In this paper, we answer these questions within the context of the OASIS framework (Pougala et al., 2022). We first estimate the parameters of two state-of-the-art utility functions (Charypar & Nagel, 2005; Feil, 2010), using data from the Swiss Mobility and Transport Microcensus (Office fédéral de la statistique and Office fédéral du développement Territorial, 2017) and compare them with OASIS' default linear-in-parameters utility function of the model.

Keywords: activity-based modelling, discrete choice modelling, parameter estimation

1 INTRODUCTION

Activity-based models (ABMs) stem from the fundamental assumption that travel demand is derived from the need to perform activities, rooted in a spatiotemporal context, and influenced by personal and environmental factors. By focusing on individuals and explicitly considering these interactions, ABMs aim to be more behaviourally realistic than traditional trip-based models, and to provide more flexible and targeted insights on individual mobility. Successful applications of activity-based models have demonstrated the added value of shifting the focus to individuals and their activities. Two main approaches can be cited: utility-based and rule-based models. The scheduling process is a result of random utility maximisation for the former, and the satisfaction of a set of spatio-temporal rules for the latter. While historically both approaches were considered contradictory, there has been increased research in hybrid models which combine elements from either theory. One limitation of hybrid approaches can be the significant complexity of the models, which require both a robust utility specification able to reflect the dynamics and interactions involved in the decision-making process, and informative constraints that can efficiently reduce the solution space without impacting the completeness of the results. We present the case of the OASIS framework, an integrated framework to simulate daily activity schedules by considering all choice dimensions (activity participation, timing decisions, mode and location choice) simultaneously. The model was designed to accommodate any utility specification and constraints depending on the context, but current implementations have relied on many simplifying assumptions both for the utility specification and formulation of constraints, mainly to compensate for the lack of available data. As a result, the utility function is too simple to properly capture all facets of activity-travel behaviour, which limits possible extensions of the model (including multiday or multiperson scheduling).

The formulation of behaviourally realistic utility functions has been the focus of many works in utility-based models for activity-travel behaviour. In particular, authors have been interested on

the impact of timing (start time and duration) on the utility of performing an activity. One common assumption, derived from studies on departure time (Small, 1982), is that individuals have preferences regarding the timings of their activities, and penalise deviations from such preferences (e.g. Pougala et al., 2022; Charypar & Nagel, 2005; Allahviranloo & Axhausen, 2018). These deviations can be asymmetrical (e.g. being late is more penalised than being early). Other authors consider these preferences more implicitly: for example, Joh et al. (2005) formalise an S-shaped utility for duration, which captures the behavioural assumption that a frustration (e.g. being involved in an activity for a duration too short) and satiation (e.g. the activity duration is too long) effects exist. This formulation always considers an increasing utility function, but at different rates depending on whether the threshold of satiation has been reached. This is not the case for formulations that consider explicit scheduling preferences, where the utility (of activity duration) can decrease. Joh et al.’s S-shaped function was adopted by multiple authors, such as Feil (2010) and Ettema & Timmermans (2003). An important gap in the literature is the calibration of utility parameters to observed data, either due to lack of data or due to the complexity of the model.

In this paper, we investigate different utility specifications for activity participation and timing, to be used within the OASIS framework. The specifications presented here are: the OASIS default utility function and two utility functions (Charypar & Nagel, 2005; Feil, 2010) used as scoring functions in the state-of-the-art agent-based microsimulator MATSim. We use the estimation component of the OASIS framework to estimate the parameters of all three utility specifications, using the Mobility and Transport Microcensus (MTMC), a nationwide travel survey of Switzerland (Office fédéral de la statistique and Office fédéral du développement Territorial, 2017). The purpose of this investigation is two-fold:

1. first, to test our methodology for parameter estimation on state-of-the-art utility functions,
2. to empirically investigate the differences of these utility specifications, and their implications in terms of behaviour.

This paper is organized as follows. In Section 2, we present the OASIS framework, and in particular the methodology to estimate the parameters of the model based on Metropolis-Hastings sampling. We apply this procedure on the three utility functions described, using a sample of the MTMC data. Finally, we discuss key differences between the specifications, and propose some insights for future investigations.

2 METHODOLOGY

OASIS framework

In this section, we briefly introduce the OASIS framework (Pougala et al., 2022). The framework (Fig. 1) is composed of two main elements:

1. A simulation model that outputs distributions of feasible schedules for given individuals, based on a mixed-integer optimisation model where the objective function is the utility of the schedule.
2. An estimation component to calibrate the parameters of the said utility function.

Parameter estimation

In the OASIS framework, the estimation of parameters is two-fold:

1. For a given individual n , generate a set of K feasible alternatives $\tilde{C}_n = \{S_0, \dots, S_{k-1}, S_k\}$ which includes the chosen schedule $S_k = S^*$. The non-chosen schedules $S_i \forall i \neq k$ are sampled using the Metropolis-Hastings algorithm (Algorithm 1), where the target distribution is proportional to the utility function of the problem. The parameters of the target distribution are estimated on a random choice set:
 - (a) Begin with an initial schedule S_0
 - (b) At each iteration i , propose a candidate state S_{new} by modifying the current schedule in one dimension (activity participation, timing, or travel).

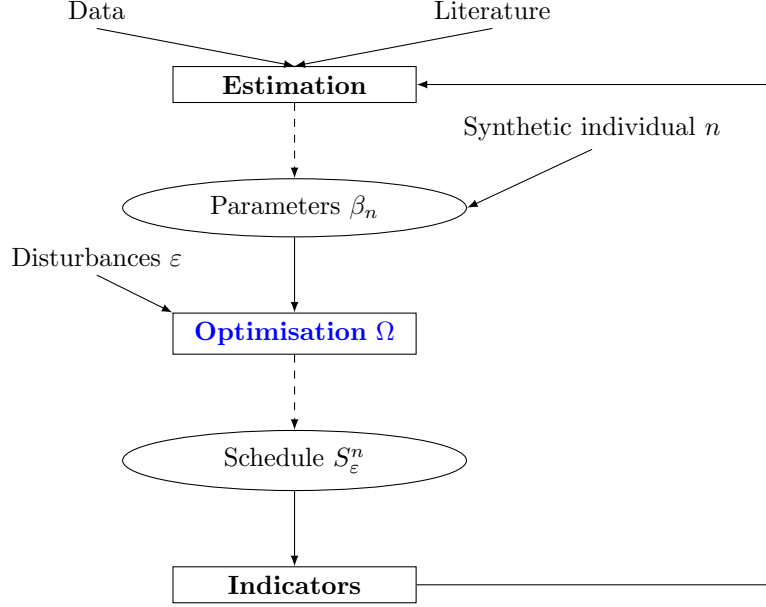


Figure 1: OASIS framework

- (c) Accept or reject S_{new} based on the predefined acceptance probability.
2. Considering the choice probability in Eq.1, the maximum likelihood estimators of the parameters $\hat{\beta}$ are derived from the corresponding likelihood function.

$$P_{in} = P_n(i|\tilde{C}_n) = \frac{e^{\mu V_{in} + \ln P_n(\tilde{C}_n|i)}}{\sum_{j \in \tilde{C}_n} e^{\mu V_{jn} + \ln P_n(\tilde{C}_n|j)}} \quad (1)$$

Given that the parameters are estimated on a sample of alternatives \tilde{C}_n , the choice probability is corrected with a term $\ln P_n(\tilde{C}_n|j)$, in order to obtain unbiased estimators (Ben-Akiva & Lerman, 1985). This term corresponds the sampling probability of the choice set \tilde{C}_n , and is directly obtained from the Metropolis-Hastings routine.

Algorithm 1 Choice set generation for the ABM with Metropolis-Hastings

$t \leftarrow 0$, initialise state with random schedule $X_t \leftarrow S_0$
 Initialise utility function with random parameters \tilde{U}_S
for $t = 1, 2, \dots$ **do**
 Choose operator ω with probability P_ω
 $X^*, q(X_t, X^*) \leftarrow \mathbf{ApplyChange}(\omega, X_t)$
 function $\mathbf{APPLYCHANGE}(\omega, \text{state } X)$
 return new state X' , transition probability $q(X, X')$
 end function
 Compute target weight $p(X^*) = U_S(\tilde{X}^*)$
 Compute acceptance probability $\alpha(X_t, X^*) = \min\left(\frac{p(X^*)q(X_t|X^*)}{p(X_t)q(X^*|X_t)}\right)$
 With probability $\alpha(X_t, X^*)$, $X_{t+1} \leftarrow X^*$, else $X_{t+1} \leftarrow X_t$
end for

Simulation

The scheduling process for a given individual n , a set of activities to be scheduled A_n , and sets of possible modes M_n and $L_{a,n}$ is summarised in 2. The set of estimated parameters $\beta_{a,n}$ is also provided as input.

The objective function of the maximisation problem is the utility function of the schedule. This function is expressed as a sum of a deterministic element and an error term (Eq.2). We assume the distribution of the error term to be known.

$$U_S = V_S + \varepsilon_S \quad (2)$$

The decision variables (activity participation, start time, duration, sequence,...) are chosen such as to maximise the utility of the schedule subject to constrained. For a draw of the error term ε_S^r , the maximisation problem becomes deterministic, and yields one optimal schedule S^r which is a draw from the distribution of schedules for n .

The constraints of the optimisation problem can be classified in two categories:

1. Mathematical constraints: they ensure that the resulting schedule is valid. For example, the time budget T cannot be exceeded, activities cannot overlap,...
2. Context-dependent constraints: they are additional rules that influence the activity-travel behaviour. For example, some activities must take precedence over others (e.g. picking up children from school must happen after they were picked up),...

Algorithm 2 Simulation of activity schedules

Initialise $n, \beta_n, A_n, M_n, L_n$

for $r = 1, 2, \dots, R$ **do**

Draw ε_S^r from distribution of error terms.

Draw schedule S_n^r by solving $\Omega = \max U_S(X_n, \beta_n, \varepsilon_S^r)$ s.t. constraints

end for

Utility specification

For an individual n , each activity a provides a utility $U_{a,n}$, composed of the following elements:

1. A *participation* term, which is constant with respect to time.
2. A utility with respect to *activity start time*.
3. A utility with respect to *activity duration*.
4. Utility terms with respect to *travel*, considering the influence of travel time and cost to the activity.

In this paper, we test different specifications of the utility function, more specifically, the utility terms associated with start time and duration.

Default OASIS utility function

In the default OASIS utility function (Eq.3), the influences of start time x_a and duration τ_a are considered by penalising deviations from preferred start time x_a^* (earlier or later starts are penalised, see Eq. 4) and duration τ_a^* (shorter or longer durations are penalised, see Eq. 5). For each activity, we therefore estimate four parameters corresponding to the penalties: $\{\theta_{\text{early}}, \theta_{\text{late}}, \theta_{\text{short}}, \theta_{\text{long}}\}$, as well as activity-specific constants. The assumption is that individuals have asymmetrical penalties for positive and negative deviations from their preferences¹, and penalise differently each activity (implying different ranges of flexibility).

Further assumptions are taken for the *home* activity:

1. The start time is constrained: the day must start at home. There is therefore no utility associated with start time.
2. The duration of the *home* activity is not associated with a preference, but results from the other scheduled activities. There is therefore no utility associated with duration.

¹The preferred start time and duration can be a fixed point or a continuous range.

$$U_S = U + \sum_{a=0}^{A-1} (U_a^{\text{participation}} + U_a^{\text{start time}} + U_a^{\text{duration}} + \sum_{b=0}^{A-1} U_{a,b}^{\text{travel}}) \quad (3)$$

$$U_a^{\text{start time}} = \theta_a^{\text{early}} \max(0, x_a^* - x_a) + \theta_a^{\text{late}} \max(0, x_a - x_a^*) + \varepsilon_{\text{start time}} \quad (4)$$

$$U_a^{\text{duration}} = \theta_a^{\text{short}} \max(0, \tau_a^* - \tau_a) + \theta_a^{\text{long}} \max(0, \tau_a - \tau_a^*) + \varepsilon_{\text{duration}} \quad (5)$$

MATSIM scoring function

The utility function used in MATSIM was formalised by Charypar & Nagel (2005). The utility of activity duration has a logarithm form (Eq. 7), which implies a decreasing marginal utility. In addition, a too short duration is penalised. For start time (Eq. 8), schedule deviations such as being late or early are penalised.

The parameters are: a parameter common to all activities β_{act} , a typical duration τ_a^* (considered known), a scaling factor A and a priority term ρ . $\beta^{\text{short}}, \beta^{\text{early}}, \beta^{\text{late}}$ penalise schedule deviations (δ).

$$U_S = U \sum_{a=0}^{A-1} (U_a^{\text{duration}} + U_a^{\text{start time}} + U_a^{\text{travel}}) \quad (6)$$

$$U_a^{\text{duration}} = \max \left[0, \beta_{\text{act}} \tau_a^* \ln \left(\frac{\tau_a}{\tau_a^* \exp(-A/(\rho \tau_a^*))} \right) \right] + \beta_a^{\text{short}} \delta_a^{\text{short}} \quad (7)$$

$$U_a^{\text{start time}} = \beta_a^{\text{early}} \delta_a^{\text{early}} + \beta_a^{\text{late}} \delta_a^{\text{late}} \quad (8)$$

PlanomatX utility function

We test the utility specification proposed by Feil (2010), which is a modification of the MATSIM utility function (Section 2). The utility function considers the impact of activity duration with an asymmetric S-shaped curve with an inflection point, as formalised by Joh et al. (2005) (Eq. 10). The parameters of the S-shape are: the inflection point α_a , the slope β_a , and the relative vertical position of the inflection point γ_a . When $\gamma_a = 1$, α_a can be considered as the duration where the utility reaches its maximum. They do not consider start time in their utility function.

$$U_S = \sum_{a=0}^{A-1} (U_a^{\text{act}} + U_a^{\text{travel}}) \quad (9)$$

$$U_a^{\text{act}} = U_a^{\text{min}} + \frac{U_a^{\text{max}} - U_a^{\text{min}}}{(1 + \gamma_a \exp \beta_a [\alpha_a - \tau_a])^{1/\gamma_a}} \quad (10)$$

3 RESULTS AND DISCUSSION

Case study

We use the Mobility and Transport Microcensus (MTMC), a Swiss nationwide survey gathering insights on the mobility behaviours of local residents (Office fédéral de la statistique and Office fédéral du développement Territorial, 2017). Respondents provide their socio-economic characteristics (e.g. age, gender, income) and those of the other members of their household. Information on their daily mobility habits and detailed records of their trips during a reference period (1 day) are also available. The 2015 edition of the MTMC contains 57'090 individuals, and 43'630 trip diaries. In order to illustrate a real-life application of the simulator, we focus on the sample of full-time students residing in Lausanne (236 individuals).

Following the methodology described in Section 2, we start by generating the choice sets of daily schedules for each individual in the sample. Each choice set is composed of 10 alternatives, including the chosen (recorded) schedule.

The models are estimated with PandasBiogeme (Bierlaire, 2020). The estimation process is done using 70% of observations in the sample data, where one observation is the daily schedule of one individual.

Estimation results

Tables 1, 2, and 3 present the parameter estimates for the OASIS utility function, the MATSIM scoring function, and the PlanomatX function, respectively. We have chosen to display only significant parameters at a 5% level. For the estimation of the MATSIM function, we have considered the same assumptions as described by Charypar & Nagel (2005) for the values of the scaling parameter ($A = -200$) and the priorities for each activity ($\rho_a = 1$ for $a \in \{\text{home, education, work}\}$ and $\rho_a = 3$ otherwise).

Similarly, we have assumed for the estimation of the PlanomatX function that $U_a^{\min} = 0$ and $\gamma_a = 1 \forall a$, as described by Feil (2010).

Finally, given that in the OASIS context, *home* is the reference alternative and therefore associated with a null utility, we do not have estimated any parameter for this activity. The magnitudes and signs of the other coefficients should therefore be considered relative to the home baseline.

OASIS

For *education*, being early seems to be slightly more penalised than being late, although the penalties are almost symmetrical. For duration, cutting the activity short is associated with a negative penalty, whereas a long duration is not regarded negatively (the associated parameter is not significantly different from 0). Surprisingly, for *work* the penalty for being late is not statistically significant, while being early or a short duration are significantly penalised.

For *leisure*, only being late is penalised. This can be explained by the fact that, while leisure is usually considered as a discretionary activity, it is likely constrained by the participation of other individuals or feasible times (e.g. opening hours of facilities). The penalty for a short duration for both leisure and shopping is not significant, which implies that these activities are more flexible than education or work for scheduling trade-offs.

Parameter	Param. estimate	Rob. std err	Rob. <i>t</i> -stat	Rob. <i>p</i> -value
Education: ASC	7.62	1.26	6.04	1.55e-09
Education: early	-1.15	0.282	-4.07	4.62e-09
Education: late	-0.89	0.214	-4.15	3.28e-05
Education: short	-0.452	0.23	-1.97	0.0493
Leisure: ASC	5.02	0.679	7.38	1.57e-13
Leisure: late	-0.747	0.135	-5.54	3.1e-08
Leisure: long	-0.137	0.0497	-2.75	0.00593
Shopping: ASC	5.04	0.807	6.25	4.07e-10
Shopping: early	-0.652	0.144	-4.53	5.88e-06
Shopping: late	-0.534	0.0944	-5.65	1.57e-08
Shopping: long	-0.17	0.06	-2.84	0.00456
Work: ASC	4.34	1.53	2.84	0.00448
Work: early	-0.71	0.223	-3.19	0.00145
Work: short	-1.37	0.481	-2.86	0.00423
Summary statistics				
	$L(0) = -454.1869$			
	$L(\hat{\beta}) = -152.0466$			
	$\hat{\rho}^2 = 0.621$			

Table 1: Estimation results for OASIS utility function. Only statistically significant parameters were included.

MATSIM

A similar behaviour is implied by the estimates of these parameters. For *education* both start time deviations are penalised (being early slightly more than being late) in comparable magnitudes. Being early at a leisure activity is not associated with a statistically significant penalty, as opposed to being late. For *work*, we have once again an insignificant parameter for being late.

Parameter	Param. estimate	Rob. std err	Rob. <i>t</i> -stat	Rob. <i>p</i> -value
β_{act}	0.0514	0.00974	5.27	1.34e-07
Education: early	-1.6	0.449	-3.57	0.00036
Education: late	-1.01	0.291	-3.48	0.00051
Leisure: late	-0.467	0.122	-3.84	0.00012
Shopping: early	-0.476	0.119	-4.01	6.04e-05
Shopping: late	-0.293	0.0842	-3.48	0.00049
Work: early	-2.75	0.712	-3.87	0.000111
Work: short	-1.59	0.493	-3.22	0.00126

Summary statistics

$L(0) = -593.8925$

$L(\hat{\beta}) = -248.568$

$\bar{\rho}^2 = 0.56$

Table 2: Estimation results for MATSIM utility function. Only statistically significant parameters were included.

PlanomatX

On the other hand, the parameters estimates for the S-shaped utility function are more difficult to interpret, especially the values of the inflection point α , which is the duration when the utility function reaches its maximum. For *education* and *work*, this parameter is around 2 hours, which means that beyond this duration, the utility increases at a decreasing rate (satiation effect). The fact that longer durations are usually scheduled for these activities suggests that the time allocation for education and work is more constraint-driven than utility-driven. For *shopping*, we observe the opposite. The inflection point is at a very high duration as compared to the typical values in the dataset. However, the negative slope suggests a decreasing utility.

Comparison of utilities

Figures 2-5 illustrate the utilities as functions of activity duration for education, work, leisure and shopping.

Given the similarity of their specifications, the OASIS and MATSIM utility functions have comparable trends for the *education* and *work* activities. The OASIS utility seems to converge towards the PlanomatX utility for very long durations.

For *leisure* and *shopping*, the results are more heterogeneous. For the MATSIM specification, the impact of duration seems negligible as the utility varies very little with duration. For these two activities the OASIS function and PlanomatX are in concordance: about the inflection point for leisure, and the overall decreasing trend for shopping.

These results show that a linear-in-parameter specification is able to capture overall utility trends. An in-depth investigation of simulation results is now required to understand the impacts of the utility specifications on results (accuracy and interpretability), and on model performance.

Parameter	Param. estimate	Rob. std err	Rob. t -stat	Rob. p -value
Education: U^{\max}	4.79	0.443	10.8	0.00
Education: α	1.57	0.202	7.75	9.1e-15
Education: β	7.56	4.84	1.56	0.119
Leisure: U^{\max}	4.47	0.379	4.50	9.1e-15
Leisure: α	0.668	0.213	3.13	0.00172
Leisure: β	2.53	0.686	3.69	0.000225
Shopping: U^{\max}	2.12	0.333	6.36	2.04e-10
Shopping: α	3.66	0.975	3.75	0.000175
Shopping: β	-4.85	2.3	-2.1	0.0353
Work: U^{\max}	3.31	0.637	5.19	2.08e-07
Work: α	2.07	0.0459	45.	0.00
Work: β	11.5	0.792	14.5	0.00

Summary statistics

$$L(0) = -454.1869$$

$$L(\hat{\beta}) = -187.871$$

$$\hat{\rho}^2 = 0.56$$

Table 3: Estimation results for PlanomatX utility function, considering $U_a^{\min} = 0$, and $\gamma_a = 1$.

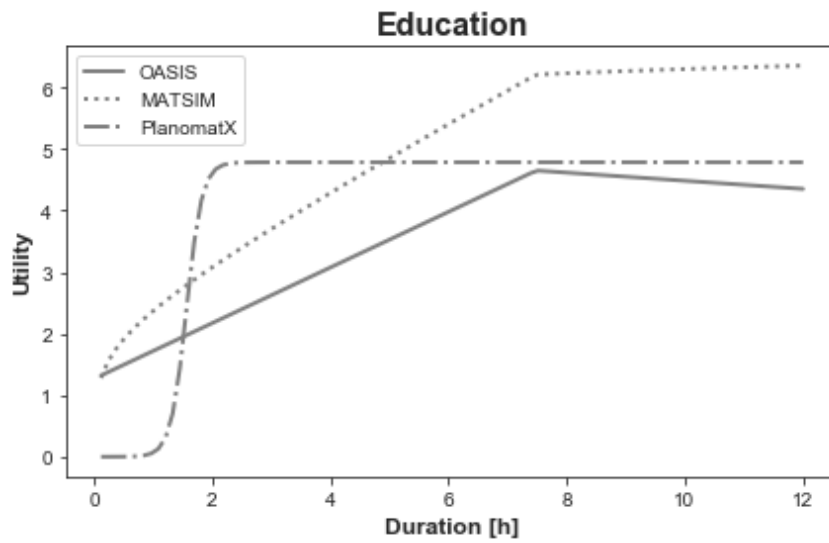


Figure 2: Utility of activity duration for education

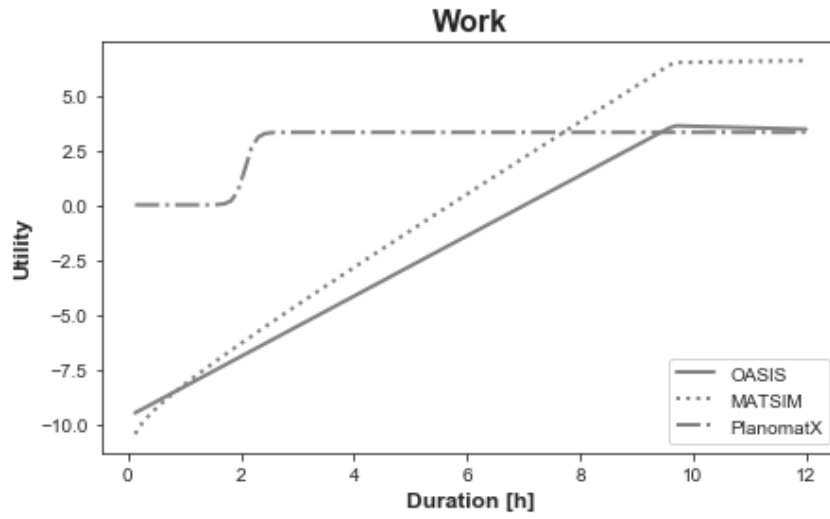


Figure 3: Utility of activity duration for work

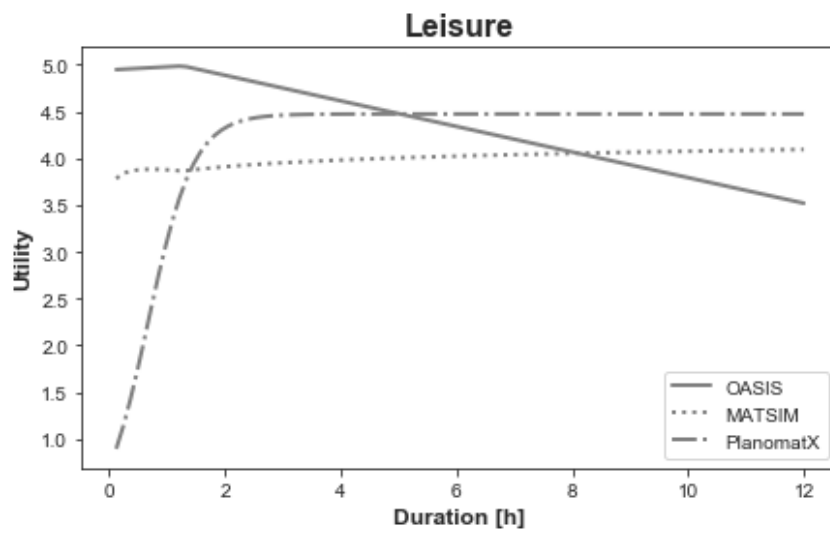


Figure 4: Utility of activity duration for leisure

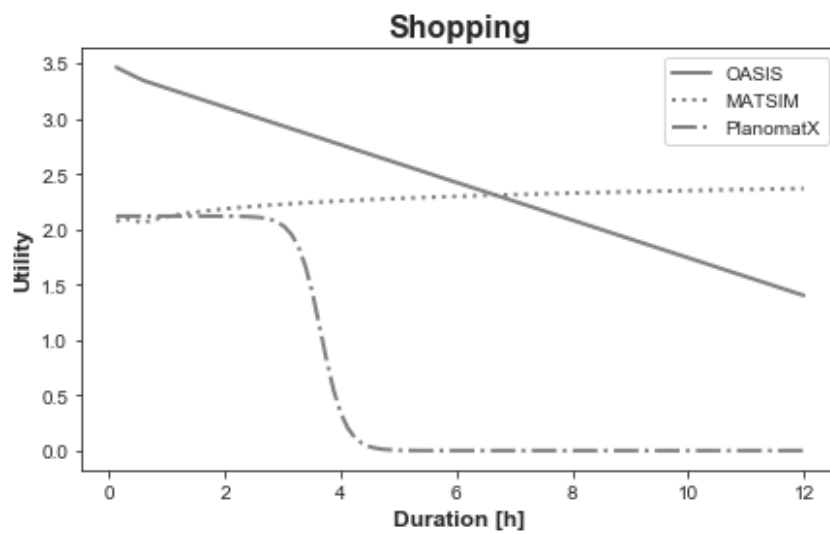


Figure 5: Utility of activity duration for shopping

4 CONCLUSIONS

In this work, we have tested three different utility specifications to explain activity-travel behaviour, with a particular emphasis on the sensitivity to activity duration. The contributions of this work are a demonstration of the estimation methodology of the OASIS framework applied to different utility functions. In addition, we have compared some behavioural insights provided by the three models, and started identifying focus areas for further investigations. Therefore, future work will include:

1. A comparative analysis of simulation outputs (schedule distributions) using the three utility functions. This analysis will be based on dedicated indicators of performance.
2. A sensitivity analysis to the model inputs, in particular to the size of the choice set, and its composition (e.g. diversity of the alternatives).
3. A relaxation of some assumptions on the values of the parameters. For instance, estimating the parameters of the PlanomatX function without the assumption of $\gamma_a = 1$.
4. An investigation of the influence of other variables and their effect on the utility of activities. For example, travel time and cost or socio-demographic characteristics.

REFERENCES

- Allahviranloo, M., & Axhausen, K. (2018). An optimization model to measure utility of joint and solo activities. *Transportation Research Part B: Methodological*. doi: 10.1016/j.trb.2017.12.004
- Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand* (Vol. 9). Cambridge, MA, USA: MIT press.
- Bierlaire, M. (2020). A short introduction to pandasbiogeme. *A short introduction to PandasBiogeme*.
- Charypar, D., & Nagel, K. (2005). Generating complete all-day activity plans with genetic algorithms. *Transportation*, 32(4), 369–397. Retrieved from <https://doi.org/10.1007/s11116-004-8287-y> doi: 10.1007/s11116-004-8287-y
- Ettema, D., & Timmermans, H. (2003). Modeling Departure Time Choice in the Context of Activity Scheduling Behavior. In *Transportation research record* (pp. 39–46). National Research Council. doi: 10.3141/1831-05
- Feil, M. (2010). *Choosing the Daily Schedule: Expanding Activity-Based Travel Demand Modeling* (Doctoral dissertation, ETH Zürich). doi: 10.3929/ethz-a-006200573
- Joh, C. H., Arentze, T. A., & Timmermans, H. J. (2005). A utility-based analysis of activity time allocation decisions underlying segmented daily activity-travel patterns. *Environment and Planning A*. doi: 10.1068/a36214
- Office fédéral de la statistique and Office fédéral du développement Territorial. (2017). *Comportement de la population en matière de transports. Résultats du microrecensement mobilité et transports 2015* (Tech. Rep.). Neuchâtel, Berne. Retrieved from www.are.admin.ch/mtmc
- Pougala, J., Hillel, T., & Bierlaire, M. (2022, jun). Capturing trade-offs between daily scheduling choices. *Journal of Choice Modelling*, 43, 100354. doi: 10.1016/J.JOCM.2022.100354
- Small, K. A. (1982). The scheduling of consumer activities: work trips. *American Economic Review*. doi: 10.2307/1831545

Mixed Integer Formulation with Linear Constraints for Integrated Service Operations and Traveler Choices in Multimodal Mobility Systems

Haoye Chen^{*1}, Jan Kronqvist², Wilco Burghout³, Erik Jenelius⁴, and Zhenliang Ma⁵

¹Doctoral Student, Department of Civil and Architectural Engineering, KTH Royal Institute of Technology, Sweden

²Assistant Professor, Department of Mathematics, KTH Royal Institute of Technology, Sweden

³Director, Centre for Traffic Research, KTH Royal Institute of Technology, Sweden

⁴Associate professor, Department of Civil and Architectural Engineering, KTH Royal Institute of Technology, Sweden

⁵Assistant Professor, Department of Civil and Architectural Engineering, KTH Royal Institute of Technology, Sweden

SHORT SUMMARY

Multimodal mobility systems provide seamless travel by integrating different types of transportation modes. Most existing studies model service operations and travelers' choices independently or limited in multimodal travel options. We propose a choice-based optimization model for optimal operations of multimodal mobility systems with embedded travelers' choices using a multinomial logit (MNL) model. We derive a mixed-integer linear formulation for the problem by linearizing transformed MNL constraints with bounded errors. The preliminary experimental test for a small mobility on demand and public transport network shows the model provides a good solution quality.

Keywords: Integrated service operations and user choices, Linearization of discrete choice constraints, Multimodal mobility systems.

1 INTRODUCTION

Multimodal mobility systems integrate different modes of transportation, such as walking, cycling, driving, public transportation, and ride-sharing services, into a seamless and efficient network. Recent advances in autonomous vehicles have the potential to increase coordination among traffic modes, especially for Mobility-on-Demand (MoD) services (e.g., taxis, Lyft, Uber, and DiDi) which provide point-to-point services and can connect travelers to public transportation.

In the multimodal mobility area, from the supply side, service providers decide on operations like vehicle routing. For the demand side, travelers choose the modes (e.g., Subway, ride-sharing services, or a combination of them) and path according to the features of available options like travel time and price. However, most studies model service providers' operations and travelers' choices independently or limited in multimodal travel options. For example, Wollenstein-Betech et al. (2022) proposed an integrated Autonomous Mobility-on-Demand (AMoD) system with public transportation. They optimize the routing and rebalancing of the AMoD fleet from the system-optimum perspective, while the travelers' choices of modes are exogenous to the operation optimization model. Liu et al. (2019) developed a multimodal transportation system integrating a choice model in which travelers can choose either public transport or MoD services but not multimodal travel options. Pi et al. (2019) integrated a choice model in a multimodal dynamic traffic assignment model by repeatedly updating travelers' pre-defined multimodal mode choices and assignment results until convergence.

Conceptually, an effective operation modeling in a multimodal mobility system should jointly consider service operations (supply) and travelers' choice preferences (demand). Mathematically, this

can be modeled as a choice-based optimization problem. Choice-based optimization models are mathematical models that optimize decision-making based on choices made by individuals. These models are commonly used in marketing, economics, and other fields where individuals make choices based on various factors such as price, quality, and convenience. Many studies found that the choice-based optimization model is highly relevant in the real world since it can provide more accurate predictions of decision-making behavior Roemer et al. (2023).

The paper proposes a choice-based optimization model for cooperative travels in the multimodal mobility system, which aims to minimize the total system travel time by deciding part of service operations while satisfying travelers' choice preferences characterized by the multinomial logit model (MNL). Due to the non-linearity and non-convexity of the MNL model embedded in the problem, limited research has thus far been devoted to solving it. Pacheco Paneque et al. (2021) proposed a mixed-integer linear formulation based on simulation for the related discrete choice models. Later, Pacheco Paneque et al. (2022) adopted scenario decomposition and scenario grouping based on their aforementioned paper into a novel Lagrangian decomposition method to solve a choice-based optimization problem.

Different from the simulation and the Lagrangian decomposition-based methods, we propose and explore a new mixed-integer formulation for the choice-based optimization model for the studied problem. The main contributions of this paper are two-fold:

- Propose a choice-based optimization model for cooperative travels in multimodal mobility systems. It optimizes system travel times by deciding part of service operations while satisfying travelers' choice preferences.
- Propose a novel mixed-integer formulation to effectively solve the choice-based optimization problem by linearizing transformed MNL constraints with bounded errors.

Note that we validate our model on a simple network with MoD and public transport services and compared it with sampling and simulation-based approach in linearizing MNL constraints. More experimental tests on the real-world network will be conducted and also compared with state-of-art models and general nonlinear optimization solvers, such as numerical optimization and meta-heuristics.

2 METHODOLOGY

Model description

We define a multi-modal transportation network G containing multiple layers. One layer represents MoD services, and each other layer can represent a specific mode, such as subway, buses, shared bikes, or walking. For simplicity, we only discuss one MoD layer and one public transportation layer (e.g., subway), as shown in Fig 1. However, our formulation can be easily expanded to one MoD layer and multiple other layers.

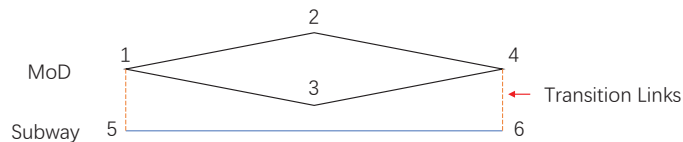


Figure 1: A simple example of a multi-modal transportation system

Denote $G = \{G_m \cup G_p\}$ that MoD layer $G_m = (V_m, E_m)$ has vertices V_m and edges E_m . The same for the public transportation layer $G_p = (V_p, E_p)$. There are transition links T connecting two different layers to represent possible transfers of modes. Denote $G = (V, E)$ that $V = \{V_m \cup V_p\}$ is the set of all points in the network and $E = \{E_m \cup E_p \cup T\}$ is the set of all edges in the network.

To model demand over different OD pairs, denote by $w = (w_s, w_t)$ an OD pair starting from vertex w_s to vertex w_t and $d_w \geq 0$ as the travel demand rate in this OD pair. Define W as the set of OD pairs. We denote by x_e^w the travel flow by OD pair w on edge e , $x_e = \sum_{w \in W} x_e^w$ the total flow on edge e , and x_e^b the rebalancing flow of MoD services. $\mathbf{X} = \{x_e^w | e \in E, w \in W\}$ and

$\mathbf{X}^b = \{x_e^b | e \in E\}$ are non-negative continuous decision variables.

For OD pair w , denote R^w as the set of possible routes and θ^{wr} as the percentage of travelers taking route r . To model relationships between edges and routes, define δ_{er}^w as equal to 1 if edge e belongs to route r of OD pair w , 0 otherwise. Along with each edge $e \in E$, we also have associated non-negative travel time $t_e \geq 0$ and non-negative cost $p_e \geq 0$. We can set travel time and prices for edges to express different travel time patterns and price policies. The travel time t^r and cost p^r can then be expressed simply by the equations

$$t^r = \sum_{e \in E} \delta_{er}^w t_e \quad \forall w \in W, \forall r \in R^w, \quad (1)$$

$$p^r = \sum_{e \in E} \delta_{er}^w p_e \quad \forall w \in W, \forall r \in R^w. \quad (2)$$

We assume the utility function of route r for the OD pair w as follows:

$$\mu^{wr} = -\beta_1 t^r - \beta_2 p^r, \quad (3)$$

where β_1 and β_2 are marginal costs for time and price respectively.

To model our problem, denoted \mathbf{P}_0 , we define $E^+(i) \in E$ as the set of edges starting from vertex i , and $E^-(i) \in E$ as the set of edges ending with vertex i . Then, \mathbf{P}_0 can be expressed as follows:

$$\min_{\mathbf{X}, \mathbf{X}^b} \sum_{e \in E} t_e(x_e) x_e \quad (4)$$

$$\text{s.t.} \quad \sum_{e \in E^-(i)} x_e^w + 1_{i=w_s} d_w = \sum_{e \in E^+(i)} x_e^w + 1_{i=w_t} d_w \quad \forall w \in W, i \in V, \quad (5)$$

$$\sum_{e \in E_m^-(i)} (x_e^b + x_e) = \sum_{e \in E_m^+(i)} (x_e^b + x_e) \quad \forall i \in V_m, \quad (6)$$

$$x_e^w = d_w \sum_{r \in R^w} \delta_{er}^w \theta^{wr} \quad \forall w \in W, \forall e \in E, \quad (7)$$

$$\theta^{wr} = \frac{\exp(\mu^{wr})}{\sum_{r' \in R^w} \exp(\mu^{wr'})} \quad \forall r \in R^w, \forall w \in W, \quad (8)$$

$$x_e^w \geq 0 \quad \forall w \in W, \forall e \in E, \quad (9)$$

$$x_e^b \geq 0 \quad \forall e \in E. \quad (10)$$

Objective (4) minimizes total travel time in the multi-modal system. Constraint (5) complies with flow conservation and demand. Constraint (6) regulates the rebalancing flow of MoD service. Constraint (7) describes the relationship between flow and route. Constraint (8) ensures the percentages of routes' choices fulfill an MNL model. Constraints (9) and (10) define non-negative ranges for the decision variables.

Linearizing transformed MNL constraints

The main computational challenge of the model is due to the nonlinear parts, especially the nonlinear constraint (8). Based on three reasonable assumptions, we are able to separately deal with distinct cases for constraint (8) to avoid some computational challenges.

For OD pair w , we select an arbitrary route r_0 as the base route. Then, constraint (8) can be represented by the constraints

$$\frac{\theta^{wr}}{\theta^{wr_0}} = \frac{\exp(\mu^{wr})}{\exp(\mu^{wr_0})}, \forall r \in R^w / r_0, \forall w \in W, \quad (11)$$

$$\sum_{r \in R^w} \theta_r = 1, \forall w \in W. \quad (12)$$

Take the natural logarithm on both sides of constraint (11):

$$\ln \theta^{wr} - \ln \theta^{wr_0} = \mu^{wr} - \mu^{wr_0}, \forall r \in R^w / r_0, \forall w \in W \quad (13)$$

Then, We give an equivalent formulation \mathbf{P}'_0 for our original problem: (4), s.t.{(5), (6), (7), (9), (10), (12), (13)}

In \mathbf{P}'_0 , constraint (13) is still non-linear and could be handled by a piece-wise linear approximation. However, for $\varphi = \ln\theta$, it is difficult to obtain accurate piece-wise linearization when θ is close to 0 since $\varphi' = 1/\theta \rightarrow +\infty$, when $\theta \rightarrow 0$. Therefore, we apply slight transformations of the original problem to avoid that θ takes a value close to 0 when doing piece-wise linearization.

Constraint (8) describes the MNL model. We revise the model by adding three assumptions for an OD pair w :

1. θ^{wr} can only take values of $0 \cup [\epsilon, 1]$ where ϵ is a small threshold, such as 0.1%, that fulfills the accuracy requirements of the application.
2. if $\theta^{wr_1} \geq \epsilon$ and $\theta^{wr_2} \geq \epsilon, r_1, r_2 \in R^w$, then:

$$\frac{\theta^{wr_1}}{\theta^{wr_2}} = \frac{\exp(\mu^{wr_1})}{\exp(\mu^{wr_2})}. \quad (14)$$

3. if $\theta^{wr} = 0$, then $\forall \theta^{wr'} \geq \epsilon, r' \in R^w, r' \neq r$:

$$\theta^{wr'} \frac{\exp(\mu^{wr})}{\exp(\mu^{wr'})} < \epsilon. \quad (15)$$

We argue that these assumptions are reasonable. In practice, the original MNL model assigns extremely small probabilities to options/routes no matter how inferior they are according to constraint (8). It is common to round the probabilities for these unattractive options to 0 as long as they are lower than a threshold similar to assumption 1. Assumption 2 ensures that all non-zero probabilities fulfill the MNL relationship. Assumption 3 ensures that options with 0 probability are unattractive options. Thus, the assumptions allow us to exclude parts of the search space that are not interesting for the application but may create numerical challenges.

We now calculate the error boundary when these three assumptions are used to replace constraint (8). For simplicity, we restrict the discussion to one OD pair and assume there are n options/routes. Denote by $\hat{\theta}_i$ the probability of route i computed based on the assumptions and by θ_i the probability computed by the original constraint. μ_i is the utility function of route i . We define sets N_0 and N_1 that route $i \in N_0$ if $\hat{\theta}_i = 0$, $i \in N_1$ if $\hat{\theta}_i \geq \epsilon$. N is the union of N_0 and N_1 .

For route $i \in N_0$:

$$\begin{aligned} \Delta_i &= |\theta_i - \hat{\theta}_i| = \theta_i \\ &= \frac{\exp(\mu_i)}{\sum_{j \in N} \exp(\mu_j)} \leq \frac{\exp(\mu_i)}{\sum_{j \in N_1} \exp(\mu_j)} \\ \text{Since: } \exp(\mu_j) &> \frac{\hat{\theta}_j \exp(\mu_i)}{\epsilon}, \forall j \in N_1 \\ &< \frac{\epsilon}{\sum_{j \in N_1} \hat{\theta}_j} = \epsilon \end{aligned} \quad (16)$$

For route $i \in N_1$:

$$\begin{aligned}
\Delta_i &= |\theta_i - \hat{\theta}_i| \\
&= \frac{\exp(\mu_i)}{\sum_{j \in N_1} \exp(\mu_j)} - \frac{\exp(\mu_i)}{\sum_{j \in N} \exp(\mu_j)} \\
&= \frac{\exp(\mu_i) \sum_{j \in N_0} \exp(\mu_j)}{\sum_{j \in N_1} \exp(\mu_j) \sum_{j \in N} \exp(\mu_j)} \\
&= \hat{\theta}_i \frac{\sum_{j \in N_0} \exp(\mu_j)}{\sum_{j \in N} \exp(\mu_j)} \leq \frac{\sum_{j \in N_0} \exp(\mu_j)}{\sum_{j \in N} \exp(\mu_j)} \\
&= \frac{1}{1 + \frac{\sum_{j \in N_1} \exp(\mu_j)}{\sum_{j \in N_0} \exp(\mu_j)}} \leq \frac{1}{1 + \frac{\sum_{j \in N_1} \exp(\mu_j)}{|N_0| \max_{j \in N_0} \exp(\mu_j)}} \\
\text{Since: } \exp(\mu_j) &> \frac{\hat{\theta}_j \max_{k \in N_0} \exp(\mu_k)}{\epsilon}, \forall j \in N_1 \\
&< \frac{|N_0| \epsilon}{|N_0| \epsilon + \sum_{j \in N_1} \hat{\theta}_j} = \frac{|N_0| \epsilon}{|N_0| \epsilon + 1} \\
&\leq \frac{|N| \epsilon}{|N| \epsilon + 1}
\end{aligned} \tag{17}$$

Therefore, with a threshold of an acceptable error bound $\max_i |\Delta_i|$, we can define our new problem P_1 with the replacement of constraint (8) by the three assumptions.

To encode the logic implied by the assumptions into our optimization problem, we introduce binary variables b^{wr} . These binary variables work as indicators with $b^{wr} = 1$ if $\theta^{wr} \geq \epsilon$, and 0 otherwise. We assume the absolute value of the utility function (3) has an upper bound $|U|_{max}$. Then, define continuous variables $\varphi^{wr} \in [\ln \epsilon - |U|_{max}, 0]$, $w \in W, r \in R^w, \tilde{\varphi}^{wr} \in [\ln \epsilon, 0], w \in W, r \in R^w$, and $\tilde{\theta}^{wr} \in [\epsilon, 1], w \in W, r \in R^w$ for auxiliary. Define a small positive value τ to deal with the strict inequality in assumption 3 and avoid numerical issues in (20).

P_1 :

$$\begin{aligned}
\min_{\mathbf{x}, \mathbf{x}^b} & \sum_{e \in E} t_e(x_e) x_e \\
\text{s.t. } & (5), (6), (7), (9), (10), (12), \\
& \theta^{wr} = 0 \cup [\epsilon, 1] \quad \forall w \in W, r \in R^w, \tag{18} \\
& b^{wr} \geq \theta/2 \quad \forall w \in W, r \in R^w, \tag{19} \\
& b^{wr} \leq \theta/\epsilon + \tau \quad \forall w \in W, r \in R^w, \tag{20} \\
& \varphi^{wr} - \varphi^{wr_0} = \mu^{wr} - \mu^{wr_0} \quad \forall w \in W, \forall r \in R^w/r_0, \tag{21} \\
& \tilde{\varphi}^{wr} = \ln \tilde{\theta}^{wr} \quad \forall w \in W, \forall r \in R^w, \tag{22} \\
& b^{wr} = 1 \rightarrow \varphi^{wr} = \tilde{\varphi}^{wr} \quad \forall w \in W, \forall r \in R^w, \tag{23} \\
& b^{wr} = 1 \rightarrow \theta^{wr} = \tilde{\theta}^{wr} \quad \forall w \in W, \forall r \in R^w, \tag{24} \\
& b^{wr_1} = 0 \text{ and } b^{wr_2} = 1 \rightarrow \\
& \quad \mu^{wr_1} - \mu^{wr_2} \leq \ln \epsilon - \varphi^{wr_2} - \tau \quad \forall w \in W, \forall r_1, r_2 \in R^w, r_1 \neq r_2, \tag{25} \\
& b^{wr} \in \{0, 1\} \quad \forall w \in W, \forall r \in R^w, \tag{26} \\
& \ln \epsilon - |U|_{max} \leq \varphi^{wr} \leq 0 \quad \forall w \in W, \forall r \in R^w, \tag{27} \\
& \ln \epsilon \leq \tilde{\varphi}^{wr} \leq 0 \quad \forall w \in W, \forall r \in R^w. \tag{28}
\end{aligned}$$

Constraints (18), (19), and (20) restrict the ranges of θ^{wr} . They indicate that θ^{wr} is non-zero by $b^{wr} = 1$ and zero by $b^{wr} = 0$. Constraints (21), (22), (23), and (24) ensure assumption 2 holds when probabilities of two routes are greater than ϵ and make piece-wise linearization of natural log function starts from $\ln \epsilon$. Constraints (25) ensure assumption 3 holds. Constraints (26), (27), and (28) define ranges of variables.

In a sophisticated mixed-integer linear programming solver, such as Gurobi, it is possible to include constraint (18) by semi-continuous variables, and constraint (22) through so-called general constraints. The formulation P_1 , thus, allows us to utilize powerful mixed-integer linear programming software.

3 RESULTS AND DISCUSSION

In this section, we present experiments based on an artificial network to show the accuracy of the results. As shown in Fig (2), the case has two layers of MoD services and Subway connected by some transition links. In the MoD layer, points represent districts in a virtual city while edges are abstracted roads. In Subway layer, there are two lines. Each station is connected with a point in the MoD layer by a transition link.

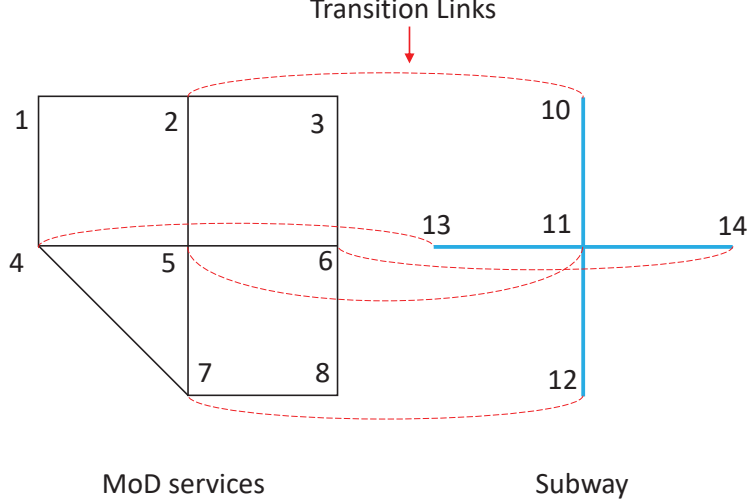


Figure 2: The case illustration

For fare, we set a distance-based fare for MoD services that $p_e = 10\$, \forall e \in E_m$ and an entrance-based fare for the subway that $p_e = 4\$, \forall e = (i, j), i \in V_m, j \in V_p$. As for the travel time, we assume a constant travel time of 5 min for transition links and 15 min for all edges in the public transportation layer. We use Bureau of Public Roads function (29) for edges in the MoD layer, given by

$$t_e(x_e) = t_e^0 \left(1 + \alpha (x_e/m_e)^\beta \right), \quad (29)$$

and we use the typical values $\alpha = 0.15$ and $\beta = 4$. $t_0 = 10\sqrt{2}$ min is set in function (29) for edge (4, 7) and $t_0 = 10$ for the rest of edges. We also assume all edges have the same capacity $m = 20k$ pcu/h. We set a fixed time of 15 min for edges in Subway layer and 5 min for transition links. To solve the model, we apply piece-wise linearization to (29). Gurobi can handle the quadratic terms in the objective (4) by "translating them into bilinear form and applying spatial branching" according to its website (<https://www.gurobi.com/documentation/10.0/refman/nonconvex.html>).

The parameters in the utility function (3) are normally estimated by real data. However, we directly define $\beta_1 = 1/\text{min}$ and $\beta_2 = 1/\$$ for two reasons: (1) Our contribution is in the algorithm. Such settings are enough for illustration and basic exploration. (2) For utility function's parameter estimation, most literature on the multi-modal transportation system normally do modal split first to decide demands for each mode. However, our problem jointly considers mode and route choices. It is difficult to find a perfectly suitable estimation in current research.

In the given case settings, we solve \mathbf{P}_1 with a threshold ϵ of 0.01 by an i9-12900H CPU and Gurobi 10.0.0 in 22.42s. Table 1 displays the solution results of our model. There are 4 OD pairs and corresponding flows as shown by $w : d_w$. $|R^w|$ represents the number of available routes of OD pair w . $r \in R^w$ shows the detailed information of active routes which have non-zero choice probabilities and how many inactive routes. $\hat{\theta}^{wr}$ is the choice probability obtained by \mathbf{P}_1 and θ^{wr} is the one computed by the original MNL model based on the utility values in the solution. Δ^{wr} is the difference between two computations of choice probabilities to measure the solution quality.

Table 1 shows the solution results of the proposed formulation. The differences Δ^{wr} between the linearized MNL and the original MNL models are quite small, which illustrates that the proposed

Table 1: Results of the proposed formulation

OD: flow $w : d_w$	Routes $ R^w $	Route $r \in R^w$	P_1 $\hat{\theta}^{wr}$	MNL θ^{wr}	Error Δ^{wr}
(1, 8): 15k	22	[1, 4, 7, 8]	100.00%	100.00%	0.00%
		21 routes remain...	0.00%	$\leq 0.00\%$	$\leq 0.00\%$
(8, 1): 30k	22	[8, 6, 3, 2, 1]	4.05%	4.02%	0.03%
		[8, 6, 5, 2, 1]	6.67%	6.65%	0.03%
		[8, 7, 4, 1]	89.27%	89.10%	0.18%
		19 routes remain...	0.00%	$\leq 0.21\%$	$\leq 0.21\%$
(3, 7): 23k	21	[3, 6, 5, 7]	43.71%	43.25%	0.46%
		[3, 6, 14, 11, 12, 7]	7.67%	7.58%	0.09%
		[3, 2, 5, 7]	41.68%	41.24%	0.44%
		[3, 2, 10, 11, 12, 7]	6.94%	6.86%	0.09%
		17 routes remain...	0.00%	$\leq 0.99\%$	$\leq 0.99\%$
(7, 3): 25k	21	[7, 8, 6, 3]	19.15%	19.14%	0.01%
		[7, 5, 6, 3]	29.83%	29.82%	0.01%
		[7, 5, 2, 3]	42.18%	42.17%	0.00%
		[7, 12, 11, 14, 6, 3]	3.25%	3.22%	0.03%
		[7, 12, 11, 10, 2, 3]	5.60%	5.59%	0.01%
		16 routes remain...	0.00%	$\leq 0.03\%$	$\leq 0.03\%$

method gives a good approximation.

We also tried the sampling and simulation method inspired by Pacheco Paneque et al. (2021) to linearize the choice constraint (Table 2). We set 100 draws for each OD pair and solved the same case in 3039.57s. Here, $\hat{\theta}^{wr}$ is the probability obtained by the simulation-based method and θ^{wr} is the one calculated by the MNL model based on the utility values in the solution. We still use the differences between the two probabilities to measure the solution quality.

Table 2: Results of the simulation-based formulation

OD: flow $w : d_w$	Routes $ R^w $	Route $r \in R^w$	Simulation $\hat{\theta}^{wr}$	MNL θ^{wr}	Error Δ^{wr}
(1, 8): 15k	22	[1, 4, 7, 8]	100.00%	100.00%	0.00%
(8, 1): 30k	22	[8, 6, 3, 2, 1]	4.00%	5.61%	1.61%
		[8, 6, 5, 2, 1]	6.00%	8.98%	2.98%
		[8, 7, 4, 1]	90.00%	85.08%	4.92%
(3, 7): 23k	21	[3, 6, 5, 7]	46.00%	37.67%	8.33%
		[3, 6, 14, 11, 12, 7]	7.00%	7.73%	0.73%
		[3, 2, 5, 7]	40.00%	45.64%	5.64%
		[3, 2, 10, 11, 12, 7]	7.00%	7.96%	0.96%
(7, 3): 25k	21	[7, 8, 6, 3]	18.00%	19.30%	1.30%
		[7, 5, 6, 3]	31.00%	27.02%	3.98%
		[7, 5, 2, 3]	42.00%	44.93%	2.93%
		[7, 12, 11, 14, 6, 3]	4.00%	3.03%	0.97%
		[7, 12, 11, 10, 2, 3]	5.00%	5.66%	0.66%

As shown in Table 2, the simulation-based formulation produces the same routes of non-zero probabilities as the ones by our proposed formulation. However, the error of the simulation-based formulation can be up to 8.33% which is much greater than the error shown in Table 1. The comparison suggests our proposed formulation has good solution quality.

4 CONCLUSIONS

This paper proposes a choice-based optimization model for integrated service operations and traveler choices in multimodal mobility systems. We derive a mixed-integer formulation by linearizing the MNL-based discrete choice constraints with bounded errors. Preliminary experiments show that the proposed formulation provides a good solution quality. Future work will derive the computation complexity and test the methodology on large-size problems, as well as compare it with state-of-art solution methods.

ACKNOWLEDGEMENTS

The project is funded by the TRENOP strategic research funding from the Swedish Government and KTH Digital Futures.

REFERENCES

- Liu, Y., Bansal, P., Daziano, R., & Samaranayake, S. (2019). A framework to integrate mode choice in the design of mobility-on-demand systems. *Transportation Research Part C: Emerging Technologies*, 105, 648–665. Retrieved 2023-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S0968090X18313718> doi: 10.1016/j.trc.2018.09.022
- Pacheco Paneque, M., Gendron, B., Sharif Azadeh, S., & Bierlaire, M. (2022). A lagrangian decomposition scheme for choice-based optimization. *Computers & Operations Research*, 148, 105985. Retrieved 2023-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S0305054822002234> doi: 10.1016/j.cor.2022.105985
- Pacheco Paneque, M., Bierlaire, M., Gendron, B., & Sharif Azadeh, S. (2021). Integrating advanced discrete choice models in mixed integer linear optimization. *Transportation Research Part B: Methodological*, 146, 26–49. Retrieved 2023-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S0191261521000266> doi: 10.1016/j.trb.2021.02.003
- Pi, X., Ma, W., & Qian, Z. S. (2019). A general formulation for multi-modal dynamic traffic assignment considering multi-class vehicles, public transit and parking. *Transportation Research Part C: Emerging Technologies*, 104, 369–389. Retrieved 2023-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S0968090X19302074> doi: 10.1016/j.trc.2019.05.011
- Roemer, N., Müller, S., & Voigt, G. (2023). A choice-based optimization approach for contracting in supply chains. *European Journal of Operational Research*, 305(1), 271–286. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0377221722004544> doi: <https://doi.org/10.1016/j.ejor.2022.05.052>
- Wollenstein-Betech, S., Salazar, M., Houshmand, A., Pavone, M., Paschalidis, I. C., & Cassandras, C. G. (2022). Routing and rebalancing intermodal autonomous mobility-on-demand systems in mixed traffic. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 12263–12275. Retrieved 2023-02-15, from <https://ieeexplore.ieee.org/document/9541261/> doi: 10.1109/TITS.2021.3112106

The valuation of arrival and departure delays in the UK passenger rail using satisfaction survey data

Kacper Rossa*¹, Andrew S J Smith², Richard P Batley³, Phillip Hudson⁴

¹ PhD Student in Rail Economics, Institute for Transport Studies, University of Leeds, United Kingdom

² Professor of Transport Performance and Economics, Institute for Transport Studies, University of Leeds, United Kingdom

³ Professor of Transport Demand and Valuation, Institute for Transport Studies, University of Leeds, United Kingdom

⁴ Arup, London, United Kingdom

SHORT SUMMARY

Stated preference surveys are typically used to derive reliability multipliers, defined as a trade-off between a minute of lateness and scheduled journey time. In this study, travel satisfaction data from the National Rail Passenger Survey in the United Kingdom is used to estimate the impact of scheduled journey length and delays on passenger satisfaction. An ordered logit model with OD fixed effects is estimated and reliability multipliers are subsequently derived. The estimated values are slightly larger than previously suggested, ranging from 4 to 9 for arrival delay and 2 to 6 for departure delay. The study offers some degree of novelty in terms of the type of data used in the estimation process. As a result, some caution is needed in using and interpreting the estimated multipliers. On the other hand, this study highlights the potential that satisfaction surveys may have in transport economics.

Keywords: lateness valuation, passenger satisfaction, rail economics, reliability multiplier, transport economics and policy

1. INTRODUCTION,

Transport researchers are interested in the impact that journey lengths, fares and delays have on passengers. Ticket sales data is often used to estimate the effect that generalized journey time (*GJT*) components have on rail demand. Following Wheat and Wardman (2017), the demand function is specified as:

$$V = \mu GJT^\lambda F^\gamma GVA^\delta \quad (1)$$

where F is the fare, GVA income, λ , γ , δ are respective elasticities and μ represents all the other factors impacting the demand. Generalised journey time is a composite index specified as:

$$GJT = T + \alpha H + \beta I \quad (2)$$

where T is the station-to-station journey time, H is the service headway and I is the number of interchanges with α and β being the respective penalty multipliers converting interchanges and headway into equivalent journey time.

Extending the demand specification presented in the equation 1, Batley, Dargay and Wardman (2011) used the following relationship between demand and average lateness at the destination prescribed by the Passenger Demand Forecasting Handbook (ATOC, 2005) in the UK:

$$Y = \left[1 + \frac{w(\bar{L}_{new}^+ - \bar{L}_{base}^+)}{GJT_{base}} \right]^\lambda \quad (3)$$

where Y is the proportionate change in demand, \bar{L}_{new}^+ and \bar{L}_{base}^+ represent average lateness at the destination in the base and new scenarios, GJT_{base} generalized journey time in the base scenario, λ is the elasticity of demand to GJT and w is the reliability (lateness) multiplier.

The lateness multiplier w defines the conversion rate of 1 minute of lateness to the equivalent of journey time. It is estimated as the ratio of the utility of lateness to the utility of scheduled journey time. Wardman and Batley (2014) provides a review of estimates of reliability multipliers with most of the initial values being around 2.5 to 3. Similar studies conducted throughout the years generally supported that figure but suggested values of up to 6.5 for airport journeys with Preston et al. (2009) estimating the different reliability multipliers by journey purpose and length as showed in Table 1.

Table 1: Reliability multipliers (Preston et al., 2009)

Journey Purpose	Short	Long
Business	2.70	1.80
Commute	3.22	2.10
Leisure	5.30	1.88

Stated preference (SP) surveys are most often used in studies where reliability multipliers are estimated (e.g. Bates et al., 2001; Preston et al., 2009; Batley and Ibáñez, 2012). In such cases, passengers are presented with alternative hypothetical travel options and make a choice regarding their preferred scenario. The differences in the options presented to the respondent are the ticket prices, scheduled journey lengths and performance. While the SP data can be subjected to biases, such as systematic bias (divergence between hypothetical and actual choices), justification bias (rationalizing actual choices) or strategic bias (influencing policy) (for review see Wardman, 1988), it has become a standard approach as it is often the only possible source of such data (Bates et al., 2001). An alternative to SP data is revealed preference (RP) data where passengers' actual travel choices are investigated. While economists typically prefer data on actual choices, RP data is more difficult to obtain and based on the assumptions of perfect information about the travel alternatives (Wardman, 1988; Bates et al., 2001; Preston et al., 2009).

An alternative to SP and RP surveys can be sought in satisfaction surveys where passengers score their satisfaction with an actual travel experience *ex post*. There is an abundance of literature looking at the impact of different journey aspects on passenger satisfaction (for reviews see de Oña and de Oña, 2014; Rong et al., 2022). In the rail context, Monsuur et al. (2021) used the National Rail Passenger Survey to estimate the impact of delays on passenger satisfaction, suggesting that passengers are very unlikely to remain satisfied with journeys delayed by over 30 minutes. Satisfaction data, typically from longitudinal household panels, have been widely used

in economic valuation in labour (Layard, Mayraz and Nickell, 2008), health (Ferrer-i-Carbonell and van Praag, 2002) and environmental economics (Frey, Luechinger and Stutzer, 2009). However, similar approaches have not been as widely used in transport economics, possibly resulting from a lack of transport surveys with such detailed information or from household surveys lacking enough transport-related information. The most important exception is a study by Dickerson, Hole and Munford (2014) looking at the relationship between life satisfaction and commuting.

This piece of work draws on earlier work using SP surveys to estimate reliability multipliers (e.g. Bates et al., 2001; Preston et al., 2009; Batley and Ibáñez, 2012). At the same time, the methodology used in this study is similar to the large body of literature using data from surveys on life satisfaction (e.g. Layard, Mayraz and Nickell, 2008; Dickerson, Hole and Munford, 2014). The major difference is the use of a survey on journey, not life, satisfaction and its cross-sectional nature. National Rail Passenger Survey is used where passengers' satisfaction with a journey is reported on a 5-point Likert scale. The passenger responses are matched to operational data to study the impact of the scheduled journey time and delays on passenger satisfaction. An ordered logit model of passenger satisfaction is used to estimate the utilities of both scheduled journey length and delay (at departure and arrival) that are subsequently used in the estimation of a reliability multiplier.

2. METHODOLOGY

National Rail Passenger Survey (NRPS) dataset

275,000 responses from 10 waves (between 2015 and 2020) of NRPS in the UK were obtained directly from Transport Focus (for more details see Transport Focus, 2020). The data has been used in multiple studies, i.e. Monsuur et al. (2021) looking at the impact of delays on satisfaction, Stead, Smith and Ojeda-Cabral (2019) comparing satisfaction with open access and franchised operators or Lyons, Jain and Weir (2016) looking at passengers' use of in-vehicle time. The responses were subsequently matched with operational data using the Historic Service Performance platform to compute scheduled journey length and corresponding delay lengths for each of the passengers. Following initial analysis, some responses were discarded based on:

- 1) Frequency of travel
Out of the 46% of passengers responding to the question regarding the frequency of travel on a given route, 73% admitted to travelling at least every 2 months. It is assumed that only the frequent travellers are affected by potential changes in scheduled journey length on a given route.
- 2) Journey purpose
Airport journeys, commuter journeys longer than 60 minutes and passengers using special ticket types were removed from the dataset.
- 3) Recorded delay length and delay perception
Responses where a passenger reported late arrival, but no delay was matched using the operational data (5.7%), were discarded as well the delays of more than 30 minutes to remove outliers and possibly erroneous responses.
- 4) Number of responses for a given origin-destination pair

OD pairs with more than 10 and 25 responses were selected as 792 OD pairs were identified with more than 10 responses over 26,026 responses and 270 pairs with more than 25 responses over 17,695 responses.

The passengers scored their overall satisfaction with their journey on a 5-point Likert scale, from ‘very satisfied’ to ‘very dissatisfied’ as showed in Figure 1. Similarly, passengers scored their satisfaction with the train, station, value for money and service frequency.

4 Your overall opinion of your journey today

Q16 Taking into account Glasgow Central station where you boarded the train and the actual train travelled on after being given this questionnaire, how satisfied were you with your journey today?

Very satisfied Fairly satisfied Neither satisfied nor dissatisfied Fairly dissatisfied Very dissatisfied Don't know/no opinion

Figure 1. Overall satisfaction questions (Transport Focus, 2020)

Deriving the reliability multiplier

As the dependent variable (satisfaction) can take one of the five outcome categories, which are in sequential order, an ordered logit model is used for estimating the latent continuous variable y^* . In this case, the probability of choosing a satisfaction category i is estimated for a given number of k categories, thus:

$$P(Y = i) = P(k_{i-1} < y^* \leq k_i) \quad (4)$$

where journey satisfaction is modelled as follows:

$$P(Y = i) = P(k_{i-1} < \beta_0 + \beta_1 SJT + \beta_2 L_D + \beta_3 L_A + \sum_{n=1}^n \beta_n Sat_n \leq k_i) \quad (5)$$

where SJT is scheduled journey time, L_A and L_D - length of delay at arrival (destination) and departure (origin) and Sat_n is a dummy variable representing passengers' satisfaction with train or station (models 1-4) and also value for money or frequency (model 4). It takes value of 1 if the passenger is satisfied with a given journey aspect or 0 otherwise.

In models 2-4, OD pair fixed-effects are included in the models by introducing a dummy variable representing each of the OD pairs represented in the sample. This allows treatment of the dataset, which is cross-sectional in nature, as a pseudo-panel of frequent rail travellers to estimate the impacts of both changes in journey times and delays on passenger satisfaction.

The ordered logit model is conceptually most suitable for modelling ordinal data (Dickerson, Hole and Munford, 2014; for review see Boes and Winkelmann, 2006), but its major disbenefit is the difficulty in interpreting the coefficients. However, as noted in Dickerson, Hole and Munford (2014), the ratios of the coefficients in the ordered model can be used to evaluate the trade-offs between variables. In this case, reliability multipliers are estimated as a ratio of utility of departure and arrival delay β_2 and β_3 to the utility of scheduled journey length β_1 . The multipliers are calculated separately for the two types of delays, at departure (w_D) and arrival (w_A) following Batley and Ibáñez (2012) for the selected three journey types. In line with the literature (i.e. Bates et al.,

2001; Preston et al., 2009; Batley and Ibáñez, 2012), the reliability multiplier represents the value of delayed time respective to the scheduled time.

$$w_D = \frac{\beta_2}{\beta_1} \text{ and } w_A = \frac{\beta_3}{\beta_1} \quad (6)$$

3. RESULTS AND DISCUSSION

The models of passenger satisfaction are estimated using an ordered logit model with estimated coefficients presented in Table 2.

Table 2: Model estimates

	1	t-stat	2	t-stat	3	t-stat	4	t-stat
Constant								
Business	0	.	0	.	0	.	0	.
Commute	-0.0778	-0.93	-0.225	-1.85	-0.186	-1.25	-0.117	-0.69
Leisure	-0.0124	-0.14	-0.0553	-0.48	-0.0223	-0.16	-0.0793	-0.49
Station_Sat								
Business	1.345***	20.32	1.350***	17.26	1.325***	15.10	1.137***	12.27
Commute	1.111***	27.59	1.183***	20.87	1.195***	16.40	0.980***	12.94
Leisure	1.401***	27.85	1.465***	21.53	1.413***	17.07	1.172***	13.56
Train_Sat								
Business	3.059***	45.77	3.066***	38.20	3.127***	34.03	2.866***	28.90
Commute	3.049***	72.16	2.998***	51.96	3.042***	41.61	2.770***	35.98
Leisure	3.438***	62.16	3.418***	46.02	3.401***	36.61	2.999***	30.14
Freq_Sat								
Business							0.803***	8.15
Commute							0.888***	12.25
Leisure							0.919***	9.88
VfM_Sat								
Business							1.049***	15.92
Commute							1.120***	15.00
Leisure							1.123***	19.61
L_A (β_3)								
Business	-0.0505***	-8.82	-0.0567***	-8.89	-0.0521***	-7.68	-0.0537***	-7.53
Commute	-0.101***	-18.36	-0.1000***	-14.12	-0.114***	-13.05	-0.109***	-12.03
Leisure	-0.0593***	-12.30	-0.0583***	-9.74	-0.0570***	-8.45	-0.0576***	-8.20
L_D (β_2)								
Business	-0.0690***	-7.77	-0.0683***	-6.46	-0.0729***	-5.99	-0.0758***	-6.01
Commute	-0.0522***	-7.71	-0.0472***	-5.20	-0.0296**	-2.60	-0.0349**	-2.97
Leisure	-0.0404***	-6.30	-0.0421***	-5.02	-0.0354***	-3.54	-0.0402***	-3.84
SJT (β_1)								
Business	-0.0009*	-1.99	-0.0120***	-5.43	-0.0121***	-5.00	-0.0134***	-5.27
Commute	-0.0058***	-6.76	-0.0134***	-5.26	-0.0142***	-4.74	-0.0123***	-3.93
Leisure	0.0002	0.53	-0.0105***	-4.95	-0.0102***	-4.34	-0.0125***	-5.06
N	40363		25457		17316		16632	
LL	-36770.8		-22388.3		-15181.8		-13920.9	
r2	0.234		0.246		0.231		0.267	
Fixed effects	X		V		V		V	
VfM and Freq	X		X		X		V	
Minimum N	1		10		25		25	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Model 1 is based on estimating the ordered logit without OD fixed effects. In this model, the delays at arrival and departure both have a statistically significant negative impact on satisfaction while the impact of journey length is less clear. However, it is not expected that passengers simply travelling longer are less satisfied with their journeys, as journey lengths naturally increase with distance. However, it is worth noting the significant and negative coefficient on scheduled journey time for commuters that may be potentially explained by their general dissatisfaction with longer travel for work (irrespective of distance). Nevertheless, it can be expected that respondents who travel on the same OD pair may be sensitive to changes in scheduled journey times and it is further assumed that these impacts are similar for travellers on the same OD. With the introduction of OD fixed effects in models 2-4, the coefficient on scheduled journey length becomes significant and negative for all journey purposes.

Using the estimated coefficients, reliability multipliers for arrival and departure delay are calculated for the models with OD fixed effects as showed in Table 3. The estimated reliability multipliers at arrival are around 4.0-4.7 for business travellers, 7.4-8.9 for commuters and 4.6-5.6 for leisure travellers. The respective departure reliability multipliers are 5.6-6.0 for business travellers, 2.1-3.5 for commuters and 3.2-4.0 for leisure travellers. The reliability multiplier is larger at departure for business travellers, slightly larger at arrival for leisure travellers and much larger for commuters at arrival. This would suggest that 1 minute of delay is valued as around 4 minutes at arrival and 6 minutes at departure for business travellers, 8 minutes at arrival and 3 at departure for commuters, and 5 minutes at arrival and 3 at departure for leisure travellers.

Table 3: Reliability multipliers

Journey Purpose	w_A			w_D		
	(2)	(3)	(4)	(2)	(3)	(4)
Business	4.74	4.31	3.99	5.72	6.02	5.64
Commute	7.43	8.02	8.86	3.51	2.08	2.83
Leisure	5.52	5.61	4.61	3.99	3.49	3.21

4. CONCLUSIONS

This study adds a degree of novelty in using passenger satisfaction data instead of the typically used SP survey data to estimate reliability multipliers. This study combined the previous work using life satisfaction surveys (e.g. Layard, Mayraz and Nickell, 2008) with work using passenger satisfaction surveys to study the impact of delays on passengers (e.g. Monsuur et al., 2021) and studies using SP surveys to estimate reliability multipliers (e.g. Batley and Ibáñez, 2012). Passenger satisfaction data was used to estimate an ordered logit model with origin-destination pair fixed-effects to estimate the utilities of delay and scheduled journey length. Subsequently their ratios were calculated to derive reliability multipliers, a conversion rate between lateness and scheduled journey length.

The estimated reliability multipliers are slightly larger than the ones typically estimated in the SP studies and some caution is needed while applying these values. To the best of our knowledge, it is the first study attempting to use journey satisfaction data in such an application. Therefore, it is believed that the methodology outlined in this study can be applied to similar satisfaction datasets for comparison of results. The study does, however, highlight the potential of using such data in transport economics. One of the key recommendation of this study is to consider including

more questions relating to income or fares in journey satisfaction questionnaires that could allow the estimation of more sophisticated metrics, including the value of time.

ACKNOWLEDGEMENTS

This research received funding under the iCASE doctoral studentship from EPSRC and Arup. The authors are also grateful for the assistance of Transport Focus and their permission to use the NRPS data.

REFERENCES

ATOC. 2005. *Passenger Demand Forecasting Handbook*. London: Association of Train Operating Companies.

Bates, J., Polak, J., Jones, P., Cook, A. 2001. The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 37, No. 2-3, pp. 191-229.

Batley, R., Dargay, J., Wardman, M. 2011. The impact of lateness and reliability on passenger rail demand. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 47, No. 1, pp. 61-72.

Batley, R., Ibáñez, N. 2012 Randomness in preference orderings, outcomes and attribute tastes: An application to journey time risk. *Journal of Choice Modelling*, Vol. 5, No. 3, pp. 157–175.

Boes, S., Winkelmann, R. 2006. Ordered response models. *Allgemeines Statistisches Arch*, Vol. 90, pp. 167–181.

De Oña, J. and de Oña, R. 2014 Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches. *Transportation Science*, Vol. 49, No. 3, pp. 605-622.

Dickerson, A., Hole, A. R., Munford, L. A. 2014. The relationship between well-being and commuting revisited: Does the choice of methodology matter? *Regional Science and Urban Economics*, Vol. 49, pp. 321-329.

Ferrer-i-Carbonell, A., van Praag, B. M. S. 2002. The subjective costs of health losses due to chronic diseases. An alternative model for monetary appraisal. *Health Economics*, Vol. 11, pp. 709-722.

Frey, B.S., Luechinger, S., Stutzer, A. 2009. The life satisfaction approach to environmental valuation. *IZA Discussion Paper* 4478.

Layard, R., Mayraz, G., Nickell, S. 2008. The marginal utility of income. *Journal of Public Economics*, Vol. 92, No. 8-9, pp. 1846-1857.

Lyons, G., Jain, J., Weir, I. 2016. Changing times – A decade of empirical insight into the experience of rail passengers in Great Britain. *Journal of Transport Geography*, Vol. 57, pp. 94-104.

Monsuur, F., Enoch, M., Quddus, M., Meek, S. 2021. Modelling the impact of rail delays on passenger satisfaction. *Transportation Research Part A: Policy and Practice*, Vol. 152, pp. 19-35.

Preston, J., Wall, G., Batley, R., Ibáñez, J. N., Shires, J. 2009. Impact of Delays on Passenger Train Services: Evidence from Great Britain. *Transportation Research Record*, Vol. 2117, No. 1, pp. 14-23.

Rong, R., Liu, L., Jia, N., Ma, S. 2022. Impact analysis of actual traveling performance on bus passenger's perception and satisfaction. *Transportation Research Part A: Policy and Practice*, Vol. 160, pp. 80-100.

Stead, A. D., Wheat, P., Smith, A. S., Ojeda-Cabral, M. 2019. Competition for and in the passenger rail market: Comparing open access versus franchised train operators' costs and reliability in Britain. *Journal of Rail Transport Planning & Management*, Vol. 12.

Transport Focus. 2020. *National Rail Passenger Survey. Spring 2020. Wave 42.*

Wardman, M. 1988. A Comparison of Revealed Preference and Stated Preference Models of Travel Behaviour. *Journal of Transport Economics and Policy*, Vol. 22, No. 1, pp. 71-91.

Wardman, M., Batley, R. 2014 Travel time reliability: a review of late time valuations, elasticities and demand impacts in the passenger rail market in Great Britain. *Transportation*, Vol. 41, No. 5, pp. 1041-1069.

Wheat, P., Wardman, M. 2017. Effects of timetable related service quality on rail demand. *Transportation Research Part A: Policy and Practice*, Vol. 95, pp. 96-108.

Understanding the Capacity of Airport Runways

Kailin Chen^{*1}, Daniel J. Graham¹, Richard J. Anderson¹, Anupriya¹, and Prateek Bansal²

¹Transport Strategy Centre, Department of Civil and Environmental Engineering, Imperial College London, United Kingdom

²Department of Civil and Environmental Engineering, National University of Singapore, Singapore

SHORT SUMMARY

Understanding the capacity of runway system under different operational conditions is of critical importance to airport operators and planners. The availability of granular data on day-to-day runway operations facilitates the development of models that allow a precise comprehension of runway capacity. However, the exercise is empirically challenging due to statistical biases that emerge via the complex interactions between air traffic control and runway capacity. This paper develops a novel causal statistical framework based on a confounding-adjusted Stochastic Frontier Analysis (SFA) to deliver estimates of runway capacity and its parameters that are robust to such biases. The model captures the key factors and interactions affecting runway capacity in a computationally intensive manner. The performance of the model is demonstrated via benchmarking of the estimated capacities of three major airports around the world.

Keywords: Airport operations; Runway capacity; Empirical estimation; Confounding; Causal statistical modelling; Stochastic frontier analysis.

1 INTRODUCTION

Runway capacity, the maximum number of aircraft movements that a runway system in an airport can operate in a given time period, is a primary input to air traffic management and planning (De Neufville et al., 2013). Knowledge of runway capacity supports decisions in planning and operations, including (1) distributing the daily runway demand over the available runway capacity (that is, available slots) (Gilbo, 1993; Cheung et al., 2021), (2) modeling airport delay or even delay propagation within the airport network (Pyrgiotis et al., 2013), and (3) appraising investments in runway capacity expansion (Hansen, 2004), among others. By nature, runway capacity is highly dynamic because it is determined by various time-varying operational factors such as weather conditions, runway configuration, and fleet mix (Ashford et al., 2011). However, even after many endeavors to estimate runway capacity, the literature lacks methods that can robustly quantify its dynamic nature, while being less resource-intensive in terms of time, labour, data, and other monetary requirements (such as expenses for a software license). This study attempts to address this gap by developing a model to precisely estimate runway capacity and the parameters of its associated factors.

Previous models of assessing runway capacity can be grouped into four main categories: (1) table lookup and spreadsheet (FAA, 1983; TRB, 2012), (2) analytical (Blumstein, 1959; Cheung et al., 2017; Mascio et al., 2020), (3) simulation (Bubalo & Daduna, 2011; Kuzminski, 2013; Barrer et al., 2005), and (4) empirical (Gilbo, 1993; Hansen, 2004; Kim & Hansen, 2010; O’Flynn, 2016; Kim et al., 2015). The first two categories of models carry several assumptions on runway operations; such as the absence of airspace constraints and the control of air traffic controllers; that seldom hold true in practical airport operation conditions. Simulation models offer the flexibility to understand the dynamic nature of runway capacity under different operational scenarios. However, the construction of such varying scenarios requires highly granular data on time-varying factors such as air traffic control regulations, which are often difficult to obtain. Moreover, the high cost in time, money and human resources (well-trained programmers) make simulation models remain frequent use only in big and hairy projects (such as detailed airfield design) rather than assessing runway capacity solely. Empirical models provide the ability to understand runway capacity in

the least data-hungry manner. Such models derive estimates of capacity from data on historical throughput that implicitly represent varied operational conditions. Nonetheless, we highlight that estimating capacity from the throughput data is not straightforward as there are external unobserved (unquantifiable) factors such as air traffic control regulations, that are correlated with both capacity and its observed determinants (such as the mix of aircraft served). We note that state-of-the-art empirical approaches in the literature, such as censored regression, fail to adjust for such unobserved sources of confounding, which limits their ability to provide a statistically robust and reproducible characterisation of runway capacity.

To fill these research gaps, we propose an empirical method to estimate runway capacity by employing a causal statistical approach, confounding-adjusted SFA, first developed by Karakaplan & Kutlu (2017); Karakaplan (2022) and further applied in Karakaplan & Kutlu (2019); Ojo & Baiye-gunhi (2020); Xu et al. (2022). The proposed model uses historical throughput as the dependent variable and associated operational condition factors as covariates. Demand and delay are introduced in the inefficiency to determine the deviation from throughput to capacity. The adopted SFA delivers estimates of runway capacity by constructing a confounding-robust throughput frontier. The confounding biases originate due to either the correlation between observed operational condition factors and unobserved operational condition factors in random error, the correlation between inefficiency and random error, or both. To the best of our knowledge, this study presents the first application of *causal statistical modelling* in empirical estimation of runway capacity. Further, we apply the proposed model to three major airports around the world by making use of data on their day-to-day operations in 2018 as maintained by the Airport Benchmarking Group (ABG) within the Transport Strategy Centre (TSC) at Imperial College London Airport Benchmarking Group (n.d.). The corresponding weather records are sourced from Weather Underground Weather Underground (n.d.), Visual Crossing Visual Crossing (n.d.) and ECMWF Reanalysis v5 Copernicus Climate Change Service (n.d.). Based on these high-granular and large-scale data, the estimates of parameters for operational factors provide insights into how these factors contribute to changes in runway capacity. Additionally, the reliability of our runway capacity estimates is validated by comparing them against reported data from table lookup and spreadsheet in FAA’s Advisory Circular Report 150/5060-5 FAA (1983) and Eurocontrol’s Airport Corner Eurocontrol (n.d.); and testing statistically via censored regression Kim & Hansen (2010). Such accurate estimates of runway capacity under specific operational conditions facilitate air traffic controllers a better understanding of the capacity under these conditions and further decisions.

2 DATA AND VARIABLES

Data

To estimate the runway capacity, we use large-scale and high-granular operational data for three congested airports provided by the ABG of TSC at Imperial College London. Because of data confidentiality, these three airports are anonymous in this study, denoted as A, B, and C. For each airport, the data provide detailed records for all aircraft movements in 2018, including scheduled and actual arrival and departure time, aircraft type, allocated runway and gate, and number of passengers. Such high-granular data including individual flight records ensure the metrics for operational factors are as close to what happens in actuality as possible Kim et al. (2015). The historical weather data for these airports are sourced from Weather Underground Weather Underground (n.d.), Visual Crossing Visual Crossing (n.d.) and ECMWF Reanalysis v5 Copernicus Climate Change Service (n.d.).

Based on these raw data, we construct panel data for each airport such that cross-sectional unit $i = 1, 2, \dots, N$ is defined as a time interval with 15 minutes length in a week and temporal unit $t = 1, \dots, T$ is the weeks in 2018. These time intervals are particularly sampled from the peak hours (10 am - 8 pm) on weekdays in a week since night hours, some holidays, and other off-peak periods are typically times of low traffic demand, which should be excluded for capacity estimation and delay performance evaluation Gelhausen et al. (2013). Therefore, we calculated the relevant variables on the basis of all individual flights that are recorded to be served in the runway system during each quarter-hour.

Variables

In this section, we introduce the dependent variable (throughput), frontier variables (factors determining capacity), and environmental variables (factors determining inefficiency) for the confounding-adjusted SFA, summarized in Table 1. The frontier is the capacity, determined by various frontier variables. These factors are categorized into airport design factors (such as active runway layout and the number of runway exits), aircraft movement characteristics (such as arrival rate and fleet mix), environmental factors (such as visibility, ceiling, precipitation, density altitude, crosswind speed, headwind speed, tailwind speed), and air traffic control factors (such as separation gap, ATM procedures, and air traffic controllers’ behaviors) Ashford et al. (2011). Inefficiency, the deviation from the frontier (capacity) to dependent variable (throughput) is explained with two environmental variables, delay, and demand. Delay has a positive effect on inefficiency since the runway system cannot be fully utilized as scheduled when scheduled flights are delayed Diana (2021). Throughput is the minimum value of capacity and demand. When demand is low during off-peak hours or in less congested airports, airport efficiency, and throughput decrease as follows Hansen (2004).

Table 1: Variables for confounding-adjusted SFA

Category	Factor	Definition	Data source
Output	Throughput	The number of aircraft that are served in the runway system	ABG
Frontier variables	Arrival rate (%)	The ratio of actual arrivals to actual departures	ABG
	Fleet mix (%)	The percent of large aircraft plus three times the percent of heavy aircraft	ABG
	Runway layout	An ordinal variable to represent the active runway layout	ABG
	Number of runway exits	The average number of runway exits of this active runway system	ABG
	Separation gap (s)	The average spacing time between leading and trailing aircraft	ABG
	Precipitation (mm)	The product of the condensation of atmospheric water vapor that falls under gravitational pull from cloud	Visual Crossing
	Visibility (miles)	The distance at which an object or light can be clearly discerned	Visual Crossing
	Ceiling (ft)	The height of the lowest clouds that cover more than half of the sky	ECMWF Reanalysis v5
	Density altitude (ft)	The pressure altitude corrected for temperature	Weather underground
	Crosswind speed (mph)	The average crosswind speed	Weather underground
Headwind speed (mph)	The average headwind speed	Weather underground	
Tailwind speed (mph)	The average tailwind speed	Weather underground	
Environmental variables	Delay (min)	The average flight delay for both arrival and departure flights	ABG
	Demand	The number of aircraft that are scheduled to be served	ABG

However, some frontier variables, such as the behaviors of air traffic controllers and the regulation rules they decided on are difficult to observe and estimate in empirical data. Therefore, the complex correlations between these unobserved frontier variables and other observed frontier/environmental variables result in confounding biases. In Figure 1, we use the level of air traffic control (ATC) to represent the air traffic controllers’ preferences and regulation rules. For example, first-come-first-serve (FCFS) is a common discipline to serve the aircraft in the runway system, while quite often, air traffic controllers loosen (decrease) the level of air traffic control and process a sequence of arrivals first and insert departures without disturbing the arrivals flow, which increases the runway capacity De Neufville et al. (2013). Therefore, when the level of ATC is tight (increase), the runway capacity decreases and arrivals would be served as a priority. The downward bias observed in the effect of arrival rate on runway capacity is due to confounded with the effect of ATC on runway capacity. The fleet mix also has a negative effect on runway capacity, while the increase in ATC might cause either a decrease or an increase in the fleet mix depending on the behaviors of air traffic controllers. Therefore, the effect of fleet mix on runway capacity would have either an upward or downward bias. The positive effect of delay on inefficiency is explained by non fully utilized runway system when scheduled flights are delayed. The inefficiency also increases when ATC becomes strict. Therefore, the estimated causal effect of delay on inefficiency experiences an upward bias when the unobserved ATC cannot be handled properly.

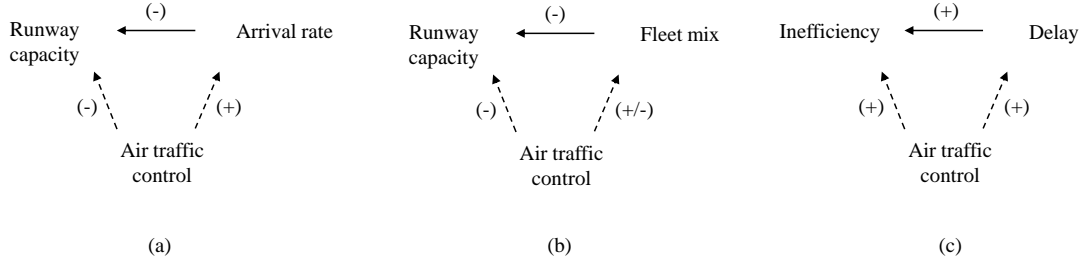


Figure 1: Causal relationships between variables

3 METHODOLOGY

To address the confounding biases discussed above, we employ a confounding-adjusted SFA, which allows the existence of a correlation between frontier variables/environmental variables and random error Karakaplan & Kutlu (2017); Karakaplan (2022).

$$y_{it} = c_{it} - u_{it} + v_{it}, \quad \text{where } c_{it} = \mathbf{x}_{fit}\boldsymbol{\alpha} \quad (1)$$

$$u_{it} = \exp(\mathbf{x}_{uit}\boldsymbol{\beta})u_i \quad (2)$$

$$\mathbf{x}_{it}^{en} = \mathbf{z}_{it}\boldsymbol{\gamma} + \boldsymbol{\epsilon}_{it} \quad (3)$$

The confounding-adjusted SFA under a panel data specification is developed in Equation 1, where unit $i = 1, \dots, N$ is the set of all 15 minutes length time intervals in a week; time $t = 1, \dots, T$ is the set of weeks. y_{it} is observed throughput for time interval i at week t . The deviation of throughput y_{it} from latent capacity c_{it} is the sum of negative inefficiency $-u_{it}$ and random error v_{it} . As we review in Section 2, runway capacity is determined by multiple airport operational condition factors. Therefore the latent runway capacity c_{it} is expressed as a function of a vector of airport operational condition factors, frontier variables, \mathbf{x}_{fit} and $\boldsymbol{\alpha}$ is a vector of unknown parameters to be estimated. v_{it} follows a normal distribution with time-invariant variance $v_{it} \sim N(0, \sigma_v^2)$. u_{it} is explained by a vector of environmental variables \mathbf{x}_{uit} . Delay and demand are introduced as environmental variables to determine inefficiency. u_i follows a non-negative normal distribution $u_i \sim N_+(\mu, \sigma_u^2)$. \mathbf{x}_{it}^{en} represents a vector of frontier variables and environmental variables that are confounding with error term; \mathbf{z}_{it} is a vector of instrumental variables for \mathbf{x}_{it}^{en} . The confounding biases are introduced by the correlation between $\boldsymbol{\epsilon}_{it}$ and v_{it} , as Equation 4.

$$\begin{bmatrix} \tilde{\boldsymbol{\epsilon}}_{it} \\ v_{it} \end{bmatrix} = \begin{bmatrix} \Omega^{-1/2}\tilde{\boldsymbol{\epsilon}}_{it} \\ v_{it} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_m & \sigma_v\rho \\ \sigma_v\rho' & \sigma_v^2 \end{bmatrix}\right) \quad (4)$$

where Ω is the variance-covariance matrix of $\boldsymbol{\epsilon}_{it}$, σ_v^2 is the variance of v_{it} , and ρ is a vector of correlation between $\tilde{\boldsymbol{\epsilon}}_{it}$ and v_{it} . By adopting Cholesky decomposition of the variance-covariance matrix of $\begin{bmatrix} \tilde{\boldsymbol{\epsilon}}_{it} \\ v_{it} \end{bmatrix}$, we have

$$\begin{bmatrix} \tilde{\boldsymbol{\epsilon}}_{it} \\ v_{it} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ \sigma_v\rho' & \sigma_v\sqrt{1-\rho'\rho} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\epsilon}}_{it} \\ \tilde{w}_{it} \end{bmatrix} \quad (5)$$

where $\tilde{\boldsymbol{\epsilon}}_{it} \sim N(0, 1)$ and $\tilde{w}_{it} \sim N(0, 1)$ are independent. Therefore, Equation 1 is expressed as

$$y_{it} = \mathbf{x}_{fit}\boldsymbol{\alpha} - u_{it} + w_{it} + \sigma_v\rho'\tilde{\boldsymbol{\epsilon}}_{it} = \mathbf{x}_{fit}\boldsymbol{\alpha} + (\mathbf{x}_{it}^{en} - \mathbf{z}_{it}\boldsymbol{\gamma})'\boldsymbol{\eta} + e_{it} \quad (6)$$

where $e_{it} = w_{it} - u_{it}$, $w_{it} = \sigma_v\sqrt{1-\rho'\rho}\tilde{w}_{it}$ and $\boldsymbol{\eta} = \sigma_v\Omega^{-1/2}\rho$. Therefore, e_{it} is conditionally independent from frontier variables \mathbf{x}_{fit} given \mathbf{x}_{it}^{en} and \mathbf{z}_{it} . Therefore, the log-likelihood function for each panel i , constructing by all T_i time periods for unit i , is given by Equation 7.

$$\ln L_i = \ln L_{i,y|x} + \ln L_{i,x} \quad (7)$$

$$L_{i,y|x} = -\frac{1}{2}(T_i \ln(2\pi\sigma_w^2) + \frac{e'_i e_i}{\sigma_w^2} + (\frac{\mu^2}{\sigma_u^2} - \frac{\mu_{i*}^2}{\sigma_{i*}^2})) + \ln(\frac{\sigma_{i*} \Phi(\frac{\mu_{i*}}{\sigma_{i*}})}{\sigma_u \Phi(\frac{\mu}{\sigma_u})}) \quad (8)$$

$$\ln L_{i,x} = -\frac{1}{2} \sum_{t=1}^{T_i} (\ln(|2\pi\Omega|) + \epsilon'_{it} \Omega^{-1} \epsilon_{it}) \quad (9)$$

where $\mu_{i*} = \frac{\sigma_w^2 \mu - \sigma_u^2 e'_i h_i}{\sigma_u^2 h'_i h_i + \sigma_w^2}$; $\sigma_{i*}^2 = \frac{\sigma_w^2 \sigma_u^2}{\sigma_u^2 h'_i h_i + \sigma_w^2}$; $\sigma_w = \sigma_v \sqrt{1 - \rho' \rho}$; $h_i = (h_{i1}, h_{i2}, \dots, h_{iT_i})$; $h_{it}^2 = \exp(x'_{uit} \beta)$; $e_{it} = y_{it} - x'_{fit} \alpha - \epsilon'_{it} \eta$; $\epsilon_{it} = x_{it}^{en} - z_{it} \gamma$; and Φ is the standard normal cumulative distribution function. The formula to predict inefficiency is $EFF_i = \exp(-u_i)$. Based on the $\hat{\alpha}$ obtained from Maximum Likelihood Estimation, the expected runway capacity estimation is $x_{fit} \hat{\alpha}$.

4 RESULTS AND DISCUSSION

Estimation results

In this section, we apply confounding-adjusted SFA to Airport A, B, and C respectively. Table 2 provides insights into how these airport operational factors contribute to the change of runway capacity and environmental factors affect inefficiency. For those exogenous factors, their parameters are in line with the intuition. The positive effect of the number of runway exits shown in airport A is in accordance with the intuition that sufficient runway exits shorten runway occupancy time and thus increase runway capacity. However, insignificant and even counter-intuitive results shown in other airports might be due to little variations in the number of runway exits for those airports with a symmetric parallel runway layout. Higher visibility and ceiling mean a greater flexible flight rule and more runway capacity. They are less significant in some airports because these European airports prefer to declare a more robust capacity, usually closer to the capacity under instrument meteorological condition (IMC), and thus capacity experiences less change when visibility and ceiling decrease Gulding et al. (2010). The results show that the increase in precipitation, crosswind speed, headwind speed, and even tailwind speed causes a significant reduction in runway capacity. Although headwind shortens runway occupancy time for both takeoffs and landings, such improvement is offset by a long time traveling through the terminal maneuvering area with a low speed due to the headwind.

In addition to these exogenous variables, three variables (arrival rate, fleet mix, and delay) are identified as endogenous. The endogeneity of these three variables is detected as statistically significant in terms of both joint significance η and individual significance η_1, η_2, η_3 . Therefore, the proposed model corrects these endogenous variables with their associated lagged difference as instruments $\Delta x_{it}(x_{it} - x_{it-1})$ and $\Delta x_{it-1}(x_{it-1} - x_{it-2})$ Arellano & Bover (1995). The performance of Model EN is reliable since these instruments pass the tests for exclusion (exogeneity) and inclusion restriction (relevance), which are the J test for over-identifying instruments (less than critical value $\chi_{3,0.95}^2 = 7.81$) and reduced from regression (greater than 10 based on the rule of thumb) respectively. The results show that arrival rate and fleet mix have negative effects on runway capacity, and delay has a positive effect on the inefficiency term, which is in accordance with intuitive signs.

Validation results

To further assess the reliability of the proposed confounding-adjusted SFA, we aim to validate our capacity estimates by simply comparing them with estimates from other empirical methods and testing them via a statistical model. Table 3 displays hourly capacity estimates under visual meteorological condition (VMC) from different empirical sources and methods, including table lookup FAA (1983), spreadsheet FAA (1983), Airport Corner in Eurocontrol Eurocontrol (n.d.), and our proposed confounding-adjusted SFA. To make this comparison practical, the operational condition across methods should be as consistent as possible, although the required inputs (operational condition factors) in each method are different.

The estimates from confounding-adjusted SFA in airport B and C appear to be low compared with the direct estimates from table lookup and spreadsheet methods while being within a reasonable range with 90 % of those estimates (in parentheses). Such a 10% reduction is suggested to make estimates comparable with the actual flow during peak hours since estimates from the previous

Table 2: Estimation Results

	Airport A	Airport B	Airport C
<i>Dep.var:throughput</i>			
constant	24.401*** (0.280)	20.806*** (0.068)	23.751*** (0.312)
arrival rate	-0.247*** (0.055)	-0.448*** (0.033)	-0.513*** (0.047)
fleet mix	0.074 (0.061)	-0.225*** (0.042)	-0.156** (0.049)
runway layout	0.614*** (0.046)	0.136*** (0.031)	0.409*** (0.032)
number of runway exits	0.617*** (0.046)	-0.019 (0.026)	-0.052* (0.030)
separation gap	-0.302** (0.093)	-0.567*** (0.076)	-0.070 (0.082)
precipitation	-0.080* (0.034)	-0.011 (0.026)	-0.103** (0.033)
visibility	-0.015 (0.041)	0.070* (0.029)	0.106** (0.034)
ceiling	-0.010 (0.042)	0.068* (0.029)	-0.044 (0.032)
density altitude	0.464*** (0.040)	0.192*** (0.028)	0.323*** (0.034)
crosswind speed	-0.115*** (0.033)	-0.069** (0.027)	-0.080** (0.031)
headwind speed	-0.149*** (0.036)	-0.114*** (0.029)	-0.099** (0.032)
tailwind speed	-0.047* (0.034)	-0.054* (0.027)	-0.100** (0.031)
<i>Dep.var:lnσ_u^2</i>			
constant	3.456*** (0.126)	0.312* (0.149)	3.671*** (0.127)
demand	-0.504*** (0.28)	-0.529*** (0.140)	-0.542*** (0.027)
delay	0.032*** (0.009)	0.221*** (0.038)	0.029** (0.009)
<i>Dep.var:lnσ_w^2</i>			
constant	2.632*** (0.012)	2.241*** (0.012)	2.458*** (0.012)
η_1 (arrival rate)	-0.458*** (0.089)	-0.102* (0.056)	-0.190* (0.075)
η_2 (fleet mix)	0.235* (0.095)	-0.051 (0.063)	-0.226** (0.09)
η_3 (delay)	-0.263*** (0.063)	-0.212*** (0.054)	-0.322*** (0.062)
η (joint endogeneity)	x2=48.76 p=0.000	x2=19.61 p=0.000	x2=40.09 p=0.000
exogeneity	0.55	3.21	0.13
relevance	34668.57	52558.43	37160.94
observations	14000	13999	13999
log likelihood	-85190.00	-77971.76	-82539.24
mean technical efficiency	0.0430	0.4419	0.0316
median technical efficiency	0.060	0.4295	0.0031

Notes: Standard errors in parenthesis. Asterisks indicate significance at 0.1% (***), 1% (**) and 5% (*) levels. All inputs are demeaned.

two empirical methods are calculated based on minimum separation gap TRB (2012). The large deviation in airport A is largely caused by mis-specifying runway layout in the first two methods without considering the runway incursion in actual runway operation that arrivals crossing the departure runway for taxiing in. The performance of our proposed method is also validated in airport B and C by comparing it with the declared capacity from Eurocontrol’s Airport Corner. Such declared capacity for the whole system is calculated by summing up the declared capacity for each runway as provided, which is valid when runways are independent, such as airports B and C. However, in airport A, the dependent parallel runways and incursion problem make actual operation cannot attain the maximum movement of each runway. Our estimate from the proposed method is still reasonable and acceptable since the percentile 99 of the throughput per hour observed in 2019 is only 108 from Eurocontrol’s performance review report Commission (1997). Other slight differences, within the 10% or even less variance range, might be due to different specifications in operational condition factors across methods or measurement errors.

Table 3: Validation via comparison

Airport	Runway layout	Capacity estimation method	Fleet mix (%)	Maximum capacity (VMC, hourly)
A	Dual independent parallel runway	Table lookup	121-180	189 (170)
		Spreadsheet	121-180	187 (168)
		Eurocontrol	–	148
		Confounding-adjusted SFA	149	98
B	Independent parallel runway	Table lookup	121-180	103 (93)
		Spreadsheet	121-180	95 (86)
		Eurocontrol	–	88
		Confounding-adjusted SFA	174	83
C	Independent parallel runway	Table lookup	81-120	111 (100)
		Spreadsheet	81-120	110 (99)
		Eurocontrol	–	90
		Confounding-adjusted SFA	116	95

Although the reliability of our estimates has been validated via simple comparison with other empirical methods, it is just a point estimate for the capacity under normal condition. Our proposed method is able to provide the estimates of runway capacity for each 15 minutes time interval under different runway operational conditions. Therefore, we further adopt censored regression to validate the proposed method statistically Kim & Hansen (2010).

$$y_i^* = \beta_1 x_i + \epsilon_i \quad (10)$$

$$y_i = \begin{cases} y_i^* & y_i^* \leq c_i \\ c_i & y_i^* > c_i \end{cases} \quad (11)$$

In this model, y_i^* is the latent capacity for a 15 minutes time interval i , y_i is the throughput (observed capacity), and c_i is the upper censoring limit, demand. x_i is the estimated capacity from the confounding-adjusted SFA. ϵ_i follows the independent identically distributed normal distribution with mean 0 and variance σ_0^2 . If estimates from our proposed method are reliable, $\hat{\beta}_1 \rightarrow 1$ could be yielded from the estimation of censored regression. Therefore, we calculate t-statistic $\frac{\hat{\beta} - \beta}{SE(\hat{\beta})}$ for the coefficients with hypothesis $H_0 : \beta_1 = 1$ against $H_1 : \beta_1 \neq 1$. Table 4 shows the results of censored regression based on the whole data set with 14000 time intervals. The estimates for β_1 are reasonably and acceptably close to the expected value in practice. Although the null hypothesis is rejected in terms of the t-statistic value, this is always a statistical problem that the null hypothesis will be rejected when the sample size is extremely large. Overall, the estimates from confounding-adjusted SFA are found to be reliable and compared favorably with the other capacity estimation method.

Table 4: Validation via censored regression

	Airport A			Airport B			Airport C		
	Estimate	Error	t-statistic	Estimate	Error	t-Statistic	Estimate	Error	t-Statistic
β_1	0.9239	0.0034	-22.38235	1.0683	0.0031	22.03226	0.8861	0.0033	-34.51515
σ_0	6.9568	0.0605	—	5.2932	0.0463	—	6.8202	0.0571	—

5 CONCLUSION

The contribution of this study is bifold, (1) first developing the idea of empirically estimating the runway capacity based on the SFA framework along with available large-scale and high-granular data, and (2) further demonstrating the added value by handling the confounding problems properly and obtaining unbiased estimates of runway capacity and the parameters of their operational condition factors via a confounding-adjusted SFA. This confounding-adjusted SFA is applied to three airports' day-to-day operation data respectively. The estimation results show the parameters for those exogenous variables are in line with the intuition. The endogeneities for those endogenous variables are statistically significant, and their parameters after instrument correction via the proposed method are also in accordance with the intuitive signs. Moreover, the runway capacity estimates from our proposed method are validated by comparing them with estimates from other empirical methods and testing via a statistical model. The capacity estimates from our proposed method are within 10% or less variance range and even nearly the same as the estimates from other methods in terms of the point estimate under the normal operational condition. The results from the statistical test also demonstrate the estimated capacity is reasonably and acceptably close to the true capacity in practice. Therefore, such unbiased estimates of parameters for associated operational condition factors facilitate airport operators' better understanding of how these factors contribute to the dynamic change of runway capacity. Based on unbiased parameters, runway capacity during short time intervals is estimated accurately which allows air traffic controllers to manage the daily demand on the runway by allocating optimal runway capacity effectively and modeling airport delay and delay propagation within airport networks. Moreover, the proposed method developed to estimate runway capacity is also capable to evaluate the latent capacity in other transport modes.

ACKNOWLEDGEMENTS

The authors thank Airport Benchmarking Group (ABG) in Transport Strategy Centre (TSC) for providing the data used in the analysis.

REFERENCES

- Airport Benchmarking Group. (n.d.). *Airport benchmarking at the transport strategy centre*. <https://www.imperial.ac.uk/transport-studies/transport-strategy-centre/applied-research/benchmarking-groups/airports>.
- Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1), 29–51. doi: 10.1016/0304-4076(94)01642-D
- Ashford, N. J., Mumayiz, S., & Wright, P. H. (2011). *Airport engineering: planning, design, and development of 21st century airports*. John Wiley & Sons.
- Barrer, J., Kuzminski, P., & Swedish, W. (2005). Analyzing the runway capacity of complex airports. In *Aiaa 5th atio and 16th lighter-than-air sys tech. and balloon systems conferences* (p. 7354). doi: 10.2514/6.2005-7354
- Blumstein, A. (1959). The landing capacity of a runway. *Operations Research*, 7(6), 752–763. doi: 10.1287/opre.7.6.752

- Bubalo, B., & Daduna, J. R. (2011). Airport capacity and demand calculations by simulation—the case of berlin-brandenburg international airport. *NETNOMICS: Economic Research and Electronic Networking*, 12(3), 161–181. doi: 10.1007/s11066-011-9065-6
- Cheung, W. L., Desai, J., Prakash, R., & Lim, H. S. (2017). An improved macroscopic analytical model for estimating runway capacity. In *Iie annual conference. proceedings* (pp. 567–572).
- Cheung, W. L., Piplani, R., Alam, S., & Bernard-Peyre, L. (2021). Dynamic capacity and variable runway configurations in airport slot allocation. *Computers & Industrial Engineering*, 159, 107480. doi: 10.1016/j.cie.2021.107480
- Commission, P. R. (1997). *Airport capacity imbalance* (Tech. Rep.).
- Copernicus Climate Change Service. (n.d.). *Era5 hourly data on pressure levels from 1959 to present*. <https://cds.climate.copernicus.eu>.
- De Neufville, R., Odoni, A. R., Belobaba, P. P., & Reynolds, T. G. (2013). *Airport systems: Planning, design, and management*. McGraw-Hill Education. Retrieved from <https://www.accessengineeringlibrary.com/content/book/9780071770583>
- Diana, T. (2021). Doing more with less: An assessment of capacity utilisation using stochastic frontier and spectral analysis models in the case of atlanta hartsfield-jackson international airport. *Journal of Airport Management*, 16(1), 87–104. Retrieved from https://www.ingentaconnect.com/contentone/hsp/cam/2021/00000016/00000001/art00009?crawler=true&mimetype=application/pdf&casa_token=MRTuYR_c3NoAAAAA:6qw3n_T3ZCQV5cWJnDA04kgqXUhhwNow07vWZHhnJ_migr1aPcGPQv4aaIS1-js4a3AK40G2U8X37ZCW
- Eurocontrol. (n.d.). *Public airport corner homepage data services*. https://ext.eurocontrol.int/airport_corner_public/.
- FAA. (1983). Airport capacity and delay. *Federal Aviation Administration, Advisory Circular AC 150/5060-5*. Retrieved from https://www.faa.gov/documentLibrary/media/Advisory_Circular/150_5060_5.pdf
- Gelhausen, M. C., Berster, P., & Wilken, D. (2013). Do airport capacity constraints have a serious impact on the future development of air traffic? *Journal of Air Transport Management*, 28, 3–13. doi: 10.1016/j.jairtraman.2012.12.004
- Gilbo, E. P. (1993). Airport capacity: Representation, estimation, optimization. *IEEE Transactions on control systems technology*, 1(3), 144–154. doi: 10.1109/87.251882
- Gulding, J., Knorr, D., Rose, M., Chen, X., Enaud, P., & Hegendoerfer, H. (2010). Us/europe comparison of atm-related operational performance. *Air Traffic Control Quarterly*, 18(1), 5–27.
- Hansen, M. (2004). Post-deployment analysis of capacity and delay impacts of an airport enhancement: case of a new runway at detroit. *Air Traffic Control Quarterly*, 12(4), 339–365. doi: 10.2514/atcq.12.4.339
- Karakaplan, M. U. (2022). Panel stochastic frontier models with endogeneity. *The Stata Journal*, 22(3), 643–663. doi: 10.1177/1536867X221124539
- Karakaplan, M. U., & Kutlu, L. (2017). Endogeneity in panel stochastic frontier models: an application to the japanese cotton spinning industry. *Applied Economics*, 49(59), 5935–5939. doi: 10.1080/00036846.2017.1363861
- Karakaplan, M. U., & Kutlu, L. (2019). School district consolidation policies: endogenous cost inefficiency and saving reversals. *Empirical Economics*, 56(5), 1729–1768. doi: 10.1007/s00181-017-1398-z
- Kim, A., & Hansen, M. (2010). Validation of runway capacity models. *Transportation research record*, 2177(1), 69–77. doi: 10.3141/2177-09
- Kim, A., Rokib, S., & Liu, Y. (2015). Refinements to a procedure for estimating airfield capacity. *Transportation Research Record*, 2501(1), 18–24.

- Kuzminski, P. C. (2013). An improved runwaysimulator—simulation for runway system capacity estimation. In *2013 integrated communications, navigation and surveillance conference (icns)* (pp. 1–11). doi: 10.1109/ICNSurv.2013.6548575
- Mascio, P. D., Rappoli, G., & Moretti, L. (2020). Analytical method for calculating sustainable airport capacity. *Sustainability*, *12*(21), 9239. doi: 10.3390/su12219239
- Ojo, T., & Baiyegunhi, L. (2020). Impact of climate change adaptation strategies on rice productivity in south-west, nigeria: An endogeneity corrected stochastic frontier model. *Science of The Total Environment*, *745*, 141151. doi: 10.1016/j.scitotenv.2020.141151
- O’Flynn, S. (2016). Airport capacity assessment methodology—acam manual. *Network Manager, EUROCONTROL, Tech. Rep, 1*. Retrieved from <https://www.eurocontrol.int/publication/airport-capacity-assessment-methodology-acam-manual>
- Pyrgiotis, N., Malone, K. M., & Odoni, A. (2013). Modelling delay propagation within an airport network. *Transportation Research Part C: Emerging Technologies*, *27*, 60–75. doi: 10.1016/j.trc.2011.05.017
- TRB. (2012). Evaluating airfield capacity. *Transportation Research Board, National Academy of Sciences*. doi: 10.17226/22674
- Visual Crossing. (n.d.). *Visual crossing historical weather data & weather forecast data*. <https://visualcrossing.com/weather-data>.
- Weather Underground. (n.d.). *Weather underground weather history and data archive*. <https://www.wunderground.com/history>.
- Xu, M., Tan, R., & He, X. (2022). How does economic agglomeration affect energy efficiency in china?: Evidence from endogenous stochastic frontier approach. *Energy Economics*, *108*, 105901. doi: 10.1016/j.eneco.2022.105901

The Effect of Incentives on the Actions Transit Riders Make in Response to Crowding

Bogdan Kapatsila*¹, Dea van Lierop², Francisco J. Bahamonde-Birke³, Emily Gris ⁴

¹ Ph.D. Candidate, School of Urban and Regional Planning, University of Alberta, Canada

² Assistant Professor, Human Geography and Spatial Planning, Utrecht University, Netherlands

³ Assistant Professor, School of Social and Behavioral Sciences, Tilburg University, Netherlands

⁴ Assistant Professor, School of Urban and Regional Planning, University of Alberta, Canada

SHORT SUMMARY

Public transit crowding influences riders' satisfaction and needs to be tackled using both demand and supply management approaches. In this study, we focus on the policy response to public transit crowding using customer incentive schemes. We used statistical tests and an Integrated Choice and Latent Variable model to analyze data collected in Metro Vancouver, Canada, during the COVID-19 pandemic. Our findings suggest that people who favor incentives tend to be more likely to change their travel behavior in response to crowding and that incentives that reduce the cost of travel have more potential to shift riders' travel time, while other incentives have a more pronounced effect on the decision to travel via a less crowded route. These findings are aimed at public transit agencies interested in employing policy instruments to manage transit crowding and researchers seeking to advance the knowledge about the influence of personal attitudes on travel behavior.

Keywords: Incentives, Transportation Demand Management, Transit Crowding, Transportation Behavior, Discrete Choice Modelling

1. INTRODUCTION

Overcrowded public transit impacts customer satisfaction and can lead some riders to opt for other modes (Cho & Park, 2021; de Oña & de Oña, 2015; dell'Olio et al., 2011; Eboli & Mazzulla, 2007; Haywood et al., 2017). Accordingly, effective strategies must be utilized for public transit crowding management that tackle the issue both quickly and efficiently. The traditional approach of adding system capacity offers a long-term solution to the challenges of transit crowding, however, such an approach is usually a prolonged and expensive endeavor that requires years of planning and execution. On the other hand, managing demand on public transit using policy tools might be an equally feasible intervention, able to provide much faster and more affordable congestion relief. Some cities, including Washington D.C., Melbourne, Sydney, Tokyo, and Hong Kong, use pre-peak hour free fares, discounts at off-peak hours, and fee increases during rush hours to manage the demand among public transit riders. More elaborate approaches attempted to use the knowledge about the human tendency to gamble (Anselme & Robinson, 2013) and engaged riders via smartphone games that offer opportunities to win prizes more valuable than a discounted or free fare, though this approach did not become widespread. To better equip public transit agencies with guidance regarding the incentives schemes that can engage riders to avoid the most congested routes or travel at less congested times, this study aims to systematically assess the riders' preferences for various incentives in the context of crowding reduction and investigate

whether the favorable view of incentives increases the likelihood of behavioral change necessary to reduce system crowding.

2. METHODOLOGY

This study pursued two objectives. To understand the differences in preferences between the offered incentives, we compared the values of collected indicators, as well as disaggregated them by income groups to gain further insights. Given the nonparametric nature of the attitudes measured on a Likert scale, the significance of the differences was evaluated using the Wilcoxon T-test (Siegel, 1956).

To achieve the second objective, we investigated the influence of the attitudes toward incentives on the decision to either change travel time or public transit route using an Integrated Choice and Latent Variable (ICLV) approach (Ben-Akiva et al., 2002). This modeling technique allows connecting the choices individuals make and attitudes they express via unobservable constructs (i.e. latent variables) and understanding the strength of the effect that attitudes have on the choices. The final integrated likelihood function for the estimated model comprised the likelihood of a selected outcome, the likelihood of observing the considered attitudinal indicators, and the distribution of the latent variable (LV). It took the following form:

$$L_q = \int_{\eta} P(y|X_q, \eta_q; \beta_x, \varepsilon_q) \cdot P(I_q | \eta_q; \gamma_{\eta}, \zeta_q) \cdot f(\eta_q | X_q, Y_q, \alpha_y, v_q) \cdot dv \quad (1)$$

There is no closed-form expression to the equation above, so it is commonly solved via numerical techniques, like a maximum simulated likelihood estimation (Ben-Akiva et al., 2002). We performed the modeling using the Apollo package (Hess & Palma, 2019) in the R statistical software (R Core Team, 2013). A 1000 Sobol draws (Sobol', 1967) were used to approximate the integration distribution and multiple starting values were tested to avoid obtaining the results for only a local optimum.

3. RESULTS AND DISCUSSION

The analysis was performed using data collected by the means of two waves of a survey disseminated in December 2020 and May 2021. Hard age and gender quotas were used to recruit a sample of respondents representative of Metro Vancouver from the panel managed by a marketing research company. Given the public transit focus of the survey, we only kept respondents who frequently commuted to work or education via transit before the COVID-19 pandemic. The final sample used for the analysis includes 1,201 respondents, the majority of whom (57.1%) did not stop using public transit during the pandemic. On top of the demographics of the individuals, we also recorded their attitudes toward incentives and actions in response to crowding using a 5-point Likert scale. Admittedly, government restrictions remained unchanged between the two waves of the survey, though the general shift towards remote employment and more private vehicle use has been observed (Kapatsila et al., 2022).

Figure 1 reveals that a fare discount is the type of incentive that had the highest support in our sample (median=4, IQR=2), followed by a 20\$ credit for a monthly pass (median=3, IQR=2) and a free coffee, or a discount coupon for a meal (median=3, IQR=3). The other options like a discount for other modes, the opportunity to participate in a raffle, or make a donation to a charity seem to be less preferable, with a median score of 2 and an equal spread. At the same time playing a smartphone game with an opportunity to win points and exchange them for a cash reward seems

to be appealing at least to some respondents. Though the median score for it is also 2, the inter-quartile range is as high as observed for the food coupons/discounts - 3. Lastly, an advantage over peers on a leadership board was the least preferable incentive (median=1, IQR=2), though a comparison to other options spread indicates that some people might consider it as well. All differences described above were found to be statistically significant.

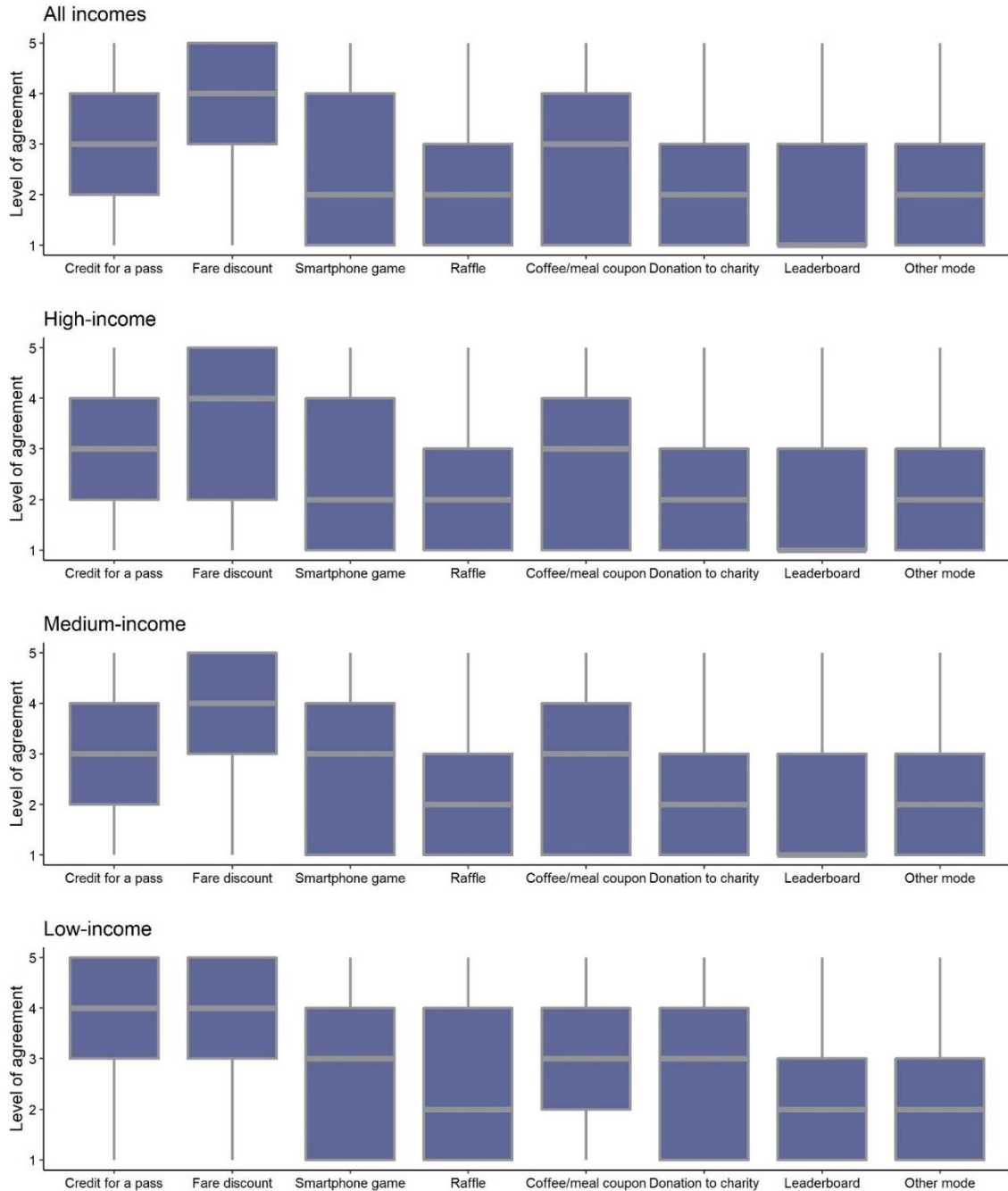


Figure 1: Attitudes towards incentives by income

Comparing the preferences towards incentives by different income groups provides additional insights. Although a fare discount remains the top choice across all income groups, the high-income earners (those making more than \$200,000 annually) display a larger range, suggesting

that some of them (most likely those at the top of the category) have a comparatively low preference for incentives in general. This, of course, is of no surprise, as it is expected that small rewards would have lower benefits for those with higher incomes. Another finding that stands out is that both medium- (making between \$50,000 and \$100,000 a year) and low-income (those earning less than \$50,000 annually) earners have a higher preference for winning points in a smartphone game when compared to high-income ones, and that difference is statistically significant ($p=0.074$ and $p=0.021$ respectfully).

In the second stage of the analysis, we simultaneously estimate two ICLV models, one evaluating the probability of changing the travel start time, and the other the probability of changing the public transit route, with both being subject to the influence of the identified LVs that captured attitudes towards incentives. Given the similar nature of the dependent variables, we introduced a normally distributed error term for both outcomes to capture the correlation effect of the parameters that could not be included in the model (e.g. social norms, trip context). The diagrammatic representation of the selected model is visualized in Figure 2.

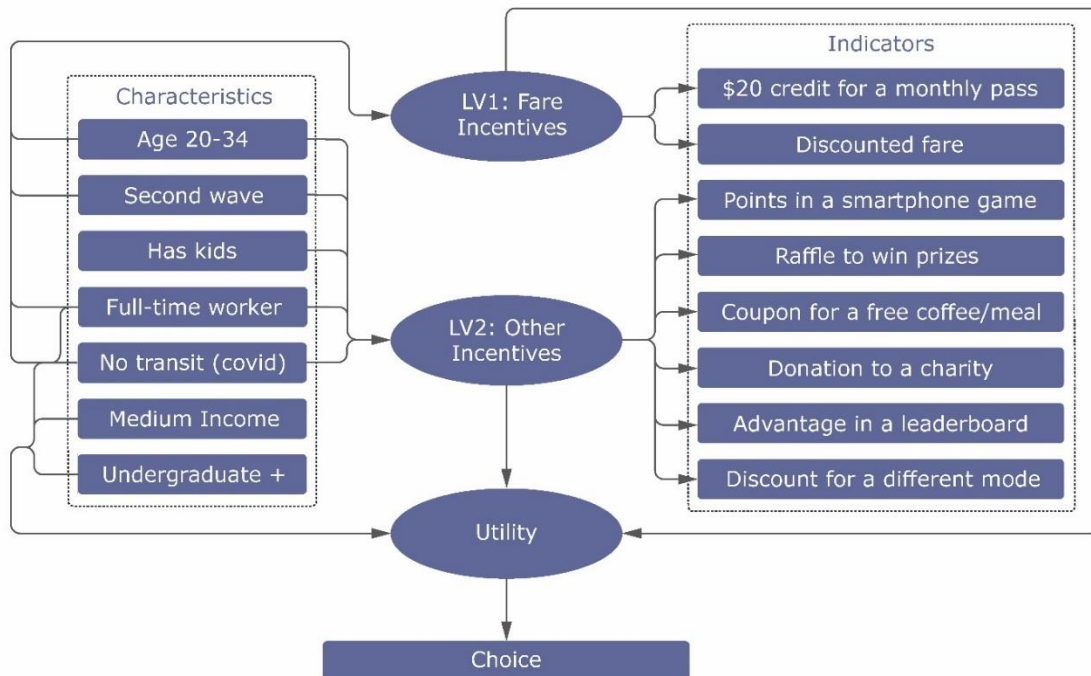


Figure 2: Diagrammatic representation of the selected ICLV model

The results of the model are presented in Table 1. Inspection of the structural equations estimates highlights the influence of several individual characteristics. Importantly, for both LVs they are nearly identical, with the only difference being individuals with kids influencing LV Other Incentives. As for other characteristics, individuals in the 20-34 age group are generally more likely to favor incentives, which goes along the lines of findings from other studies that pointed engagement with incentives to go down with aging (Dhingra et al., 2020). It is also natural that full-time workers are less likely to respond to consider incentives as they are caught between professional and domestic responsibilities and have little flexibility for any changes. The fact that people who stopped using public transit during the COVID-19 pandemic are less likely to favor incentives on transit is also fairly intuitive. It is hard to imagine for people who abandoned public transit out of concern or necessity to see incentives to change travel behavior on public transit in a positive light. The ebb and flow of the pandemic tide can also explain the more positive view of incentives

that respondents from the second wave of the survey had. In May 2021 Metro Vancouver saw a gradual increase in vaccination and a decline in COVID-19 hospitalizations (British Columbia Provincial Health Services Authority & BC Centre for Disease Control, 2022), which most likely improved the uneasiness towards public transit in general, and incentives on it as well.

Table 1: Results of the Model

Variable	Equation	Estimate	SD	t-test
Age 20-34	S.E. LV1: Fare Incentives	0.165	0.066	2.498
Full-time worker		-0.194	0.064	-3.045
No transit use (pandemic)		-0.181	0.063	-2.855
Second wave of the survey		0.133	0.063	2.122
Age 20-34	S.E. LV2: Other Incentives	0.364	0.066	5.506
Full-time worker		-0.134	0.062	-2.133
No transit use (pandemic)		-0.415	0.063	-6.642
Second wave of the survey		0.111	0.062	1.801
Has kids		0.306	0.069	4.416
ASC Change Route	Utility Change Route	2.396	0.221	10.844
Medium income		0.414	0.179	2.315
Undergraduate degree or higher		0.250	0.189	1.325
Full-time worker		-0.050	0.187	-0.272
No transit use (pandemic)		-0.262	0.184	-1.422
LV 1: Fare Incentives		0.255	0.118	2.158
LV 2: Other Incentives		0.663	0.117	5.678
Threshold 1		0	-	-
Threshold 2		1.504	0.102	-
Threshold 3		3.675	0.149	-
Threshold 4	5.504	0.194	-	
ASC Change Route	Utility Change Travel Time	3.102	0.236	13.131
Medium income		0.436	0.180	2.423
Undergraduate degree or higher		0.472	0.190	2.486
Full-time worker		-0.652	0.189	-3.447
No transit use (pandemic)		0.895	0.187	4.780
LV 1: Fare Incentives		0.722	0.123	5.862
LV 2: Other Incentives		0.472	0.115	4.109
Threshold 1		0	-	-
Threshold 2		1.530	0.120	-
Threshold 3	3.638	0.162	-	
Threshold 4	6.069	0.213	-	
Correlation Change Route & Change Travel Time		2.325	0.111	20.89
Number of observations: 1201				
Number of parameters: 70				
Log-likelihood of the whole model: -15422.72				

Shifting focus from the LVs to their impact on the choices, we can see that both LVs have a positive influence on the likelihood of either changing travel time or public transit route in response to crowding. This confirmation is a piece of encouraging evidence suggesting that at least in the stated preference design setting people who are more likely to respond to incentives and change their travel habits also tend to have a higher probability of changing travel behavior in response to crowding. Another insight worth noting is the size of the effect each LV has on the choices. Looking at the choice to change the public transit route, we can see that it is more likely to pertain to the individuals favoring other incentives since the respective LV has a higher impact than LV Fare Incentives on that choice. On the other hand, the reverse is true for the choice to change travel time. One explanation for this difference can be the familiarity of respondents with the fare price change in Metro Vancouver where it is more expensive to travel on light rail and ferries between the 3 zones at peak hours (TransLink, n.d.). As for the higher influence of LV Other Incentives on the likelihood to change the public transit route, several explanations can be hypothesized. There might be a correlation in the skills and preferences needed to both opt for another transit line and to play a game on a smartphone to win points, as both can be achieved using a smartphone (e.g. getting navigation via a route planning mobile application in the case of the former), however, the latter is impossible without a smartphone. Similarly, there is potentially a positive relationship between the propensity to switch to other public transit routes and responding to a discount for the use of other modes, as both require a change in the usual means of commuting.

4. CONCLUSIONS

This study investigated the differences in preferences towards various incentive schemes on public transit and assessed the relationship between the riders' eagerness to modify their travel patterns in response to crowding and the likelihood to respond to incentives that influence them to do the same. We found that people who favor incentives tend to be more likely to change their travel behavior in response to crowding and that incentives that reduce the cost of travel on public transit have more potential to shift riders' travel time, while other incentives have a more pronounced effect on the decision to travel via a less crowded public transit route. Similarly, we identified the incentive schemes that received the highest support and the demographics of potential users who favor those. Nevertheless, this study was subject to several limitations that should be acknowledged. First, the analysis was performed using a stated choice survey, which does not necessarily mean that the opinions respondents expressed would reflect their actual behavior. Similarly, some people might be highly favorable to incentives but have very limited options to change their travel time or route in practice. As such, future research should explore the opportunities to analyze the revealed choices of public transit riders when it comes to incentives. Secondly, both waves of the survey data were collected during the COVID-19 pandemic, and this time of heightened attention to public health and fewer systematic professional and personal travel needs could have affected the results obtained.

ACKNOWLEDGEMENTS

The authors are grateful for the financial support and cooperation provided by TransLink's New Mobility Lab that made this research possible. We are also thankful to Professor Dick Ettema for his thoughtful review of and feedback on the survey design.

REFERENCES

- Anselme, P., & Robinson, M. (2013). What motivates gambling behavior? Insight into dopamine's role. *Frontiers in Behavioral Neuroscience*, 7. <https://www.frontiersin.org/articles/10.3389/fnbeh.2013.00182>
- Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T., & Polydoropoulou, A. (2002). Integration of Choice and Latent Variable Models. In *In Perpetual Motion* (pp. 431–470). Elsevier. <https://doi.org/10.1016/B978-008044044-6/50022-X>
- British Columbia Provincial Health Services Authority, & BC Centre for Disease Control. (2022, September 1). *B.C. COVID-19 Dashboard*. British Columbia COVID-19 Dashboard. <https://experience.arcgis.com/experience/a6f23959a8b14bfa989e3cda29297ded>
- Cho, S.-H., & Park, H.-C. (2021). Exploring the Behaviour Change of Crowding Impedance on Public Transit due to COVID-19 Pandemic: Before and After Comparison. *Transportation Letters*, 13(5–6), 367–374. <https://doi.org/10.1080/19427867.2021.1897937>
- de Oña, J., & de Oña, R. (2015). Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches. *Transportation Science*, 49(3), 605–622. <https://doi.org/10.1287/trsc.2014.0544>
- dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy*, 18(1), 217–227.
- Dhingra, I., Zhang, S., Zhornitsky, S., Le, T. M., Wang, W., Chao, H. H., Levy, I., & Li, C.-S. R. (2020). The effects of age on reward magnitude processing in the monetary incentive delay task. *NeuroImage*, 207, 116368. <https://doi.org/10.1016/j.neuroimage.2019.116368>
- Eboli, L., & Mazzulla, G. (2007). *Service Quality Attributes Affecting Customer Satisfaction for Bus Transit*. <https://doi.org/10.5038/2375-0901.10.3.2>
- Haywood, L., Koning, M., & Monchambert, G. (2017). Crowding in public transport: Who cares and why? *Transportation Research Part A: Policy and Practice*, 100, 215–227. <https://doi.org/10.1016/j.tra.2017.04.022>
- Hess, S., & Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, 32, 100170. <https://doi.org/10.1016/j.jocm.2019.100170>
- Kapatsila, B., Gris , E., Lierop, D. van, & Bahamonde-Birke, F. J. (2022). From Riding to Driving: The Effects of the COVID-19 Pandemic on Public Transit in Metro Vancouver. *Findings*. <https://doi.org/10.32866/001c.33884>
- R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Siegel, S. (1956). Wilcoxon T-test. In *Non-parametric statistics for the behavioral sciences* (pp. 75–83). McGraw-Hill.
- Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4), 86–112. [https://doi.org/10.1016/0041-5553\(67\)90144-9](https://doi.org/10.1016/0041-5553(67)90144-9)
- TransLink. (n.d.). *Pricing and Fare Zones* [Government Agency]. Retrieved August 13, 2022, from <https://www.translink.ca/transit-fares/pricing-and-fare-zones>

The impact of covid-19 on modal shift in long-distance travel

Nejc Geržinič*¹, Maurizio van Dalen², Barth Donners³, Oded Cats⁴

¹ PhD candidate, Department of Transport and Planning, TU Delft, Netherlands

² MSc student, Department of Transport and Planning, TU Delft, Netherlands

² Supply Chain Specialist, Albert Heijn, Netherlands

³ Consultant, Royal HaskoningDHV, Netherlands

⁴ Full Professor, Department of Transport and Planning, TU Delft, Netherlands

SHORT SUMMARY

This research aims to analyse the perception of covid-19 infection risk in long-distance travel in Europe and how it impacts mode choice and travel behaviour. We make use of an HII variant type experiment and model it by means of a latent class choice model, where we uncover four distinct user groups. For infection risk perception, we apply a novel approach in the field, utilising a weighted least squares regression, to obtain segment-specific regression functions, based on their respective probabilistic segment allocations. Some segments exhibit risk-aversion behaviour that is time-based (longer journeys perceived as more risky), whereas others see it as time-independent. With respect to modal preferences, the four segments either show a strong preference or aversion to one of the two land-based modes: car-loving, car-averse (using train or air), train-loving and train-averse (using car and air).

Keywords: COVID-19, Discrete choice modelling, Hierarchical information integration, Long-distance travel, Risk perception, Travel behaviour

1. INTRODUCTION

In recent years, long-distance travel has become an increasingly prolific topic of both scientific literature and political discussion. With many new proposals, projects and service launches, a fair amount of research has been carried out in order to evaluate their impacts. Sun et al. (2017), provide an extensive literature review on research concerning long-distance travel, with a focus on high-speed rail (HSR). Perhaps their most relevant finding is that many papers within this domain report conflicting conclusions with regard to the consequences of introducing HSR. These differences can presumably be attributed to varying implementations, service patterns, policies and regulation of the air, rail and road markets in different contexts.

With respect to long-distance travel behaviour, most studies analysed passengers' perception of in-vehicle time, access/egress time, frequency/headway/waiting time, reliability and comfort. In-vehicle time was found to be valued at 10-30 €/h, with travellers being 1-2 times as sensitive to access/egress time as opposed to in-vehicle time (Bergantino & Madio, 2018; Ortúzar & Simonetti, 2008; Román et al., 2014; Román & Martín, 2010). The perception of other attributes differed even more, due to the different survey contexts.

There are many other potential trip characteristics that travellers may consider (transfer characteristics, luggage, reliability,...), making it challenging to capture everything. An approach that can help alleviate this is Hierarchical Information Integration (HII), first proposed by Louviere (1984). In it, respondents firstly evaluate groups of attributes (safety, reliability, comfort,...) one at a time. Secondly, a bridging experiment includes all attribute groups in a single discrete choice task, requiring respondents to trade-off the values of the different groups. Since its introduction, different versions of HII have been proposed, summarised by Molin & Timmermans (2009) as Conventional HII (Louviere, 1984), HII variant (Bos et al., 2004) and Integrated Choice experiment (Oppewal et al., 1994). Recently, the HII variant approach was applied onto two topics relating to long-distance travel, namely to the perception of airline safety (Molin et al., 2017) and the perception of night trains (Heufke Kantelaar et al., 2022).

Another aspect impacting long-distance travel behaviour, which has been at the forefront of travel behaviour research in recent years, is the COVID-19 pandemic and the associated (travel) behaviour changes. While many studies look into the impact of the pandemic on everyday life and the perception of risk on commuters' behaviour (Currie et al., 2021; de Haas et al., 2020; Shamshiripour et al., 2020; Shelat et al., 2022; Shortall et al., 2022; Tirachini & Cats, 2020), its impact on long-distance travel behaviour has, to the best of our knowledge, not been studied.

Molin et al. (2017) and Heufke Kantelaar et al. (2022) both utilised a Panel Mixed Logit (ML) model for the bridging experiment to capture respondent heterogeneity. Another common approach for capturing heterogeneity is the Latent Class Choice Model (LCCM). While ML allows parameters to vary within the sample, LCCM probabilistically allocates respondents to a discrete number of Multinomial Logit (MNL) models, each with its own set of taste parameters. LCCMs allow for a more straightforward interpretation, as the market segments can be clearly distinguished based on their different trade-offs. Additionally, attitudinal and socio-demographic information can be used to predict an individual's probability of belonging to a specific segment.

Applying LCCMs to HII data is challenging, as the perception of the subjective ratings in the bridging experiment will result in different parameters for each market segment. If the rating experiment is modelled with a regression function (as is common practice), it results in an equal perception of factors across segments. Intuitively, this subjective perception should differ between segments, yet to the best of our knowledge, latent class segmentation has not been attempted in HII.

One possible way of capturing segments' different perceptions of rating factors is by estimating separate regression models for each segment. However, since the class allocation of individuals in LCCM is probabilistic, this would have to be translated into the regression models as well. One approach that could utilise the allocation probabilities is Weighted Least Squares (WLS), which associates a weight with each data point, indicating the accuracy/importance of said data point for the regression model.

The contributions of this paper are twofold. Firstly, we evaluate the perception of various COVID-19 measures aimed at limiting the spread of the COVID-19 virus, through an HII variant type SP survey. The rating experiment includes eight attributes associated with the perception of infection risk. This is then carried into the bridging experiment, along with travel cost, travel time and travel class, where respondents choose their preferred travel mode for a long-distance trip of approximately 500km and 1000km. Secondly, upon modelling the bridging experiment by means of an LCCM, we estimate several WLS regression models to uncover different perceptions of infection risk as experienced by different population segments obtained from the LCCM.

2. METHODOLOGY

We design an HII variant type experiment, comprised of two sections: the rating experiment and the bridging experiment. Both experiments have their own design and modelling approach, all of which is presented in this section. We first present both survey designs, followed by both modelling approaches.

Survey design

In the rating experiment, respondents are presented with a variety of factors pertaining to COVID-19 and their perceived risk of infection. Shelat et al. (2022) performed an HII experiment for COVID-19 risk perception for train route choice in the Netherlands. In the rating experiment, they analysed on-board crowding, number of transfers, mask policy, sanitisation, infection rate and lockdown status. Crowding has also been recognised as a major influencing factor on mode choice by Currie et al. (2021). We consider most of the listed factors as relevant for long-distance travel. We do not consider the number of transfers, as in long-distance travel, these can vary substantially and would be very difficult to present to respondents in a consistent and understandable way. We expand the lockdown status into two categories, one considering the originating country (travel advice) and the other considering the destination country (entry requirements). Finally, we add vaccination rate to the design, as vaccination had become widespread and the concept of herd immunity may make travellers feel more at ease.

As we do not have priors for all factors, we design an orthogonal (fractional factorial) design in Ngenie (ChoiceMetrics, 2018). The design has 20 rows, divided into four blocks, resulting in five choice tasks per respondent. Based on the levels of the following eight factors, respondents had to indicate their perceived level of risk on a 5-point Likert-scale :

- Mask policy (type of mask required)
- Air circulation (ventilation, air-conditioning)
- Cleaning policy
- Government travel advice (in the origin country for the destination country)
- Visiting country entry requirements (proof of vaccination/recovery/negative test)
- Infection rate at the destination
- Vaccination rate at the destination

Next, respondents are confronted with the bridging experiment, where the perceived risk of infection is presented alongside travel time, travel cost and travel class (first/business or second/economy class). Each attribute is associated with three travel modes: car, train, aircraft. Based on past research, these seem to be the most frequently used modes and the most relevant attributes in long-distance travel (Bergantino & Madio, 2018; Ortúzar & Simonetti, 2008; Román et al., 2014; Román & Martín, 2010). For both travel time and cost, we consider them to be door-to-door, meaning they include the access and egress time to the airport/train station and the time spent there.

We employ a Bayesian D-efficient design, using priors from literature. A value-of-time of 10 €/h is used, based on a detailed study carried out for the Dutch government (Kouwenhoven et al., 2014). For first/business class, we assume an additional willingness-to-pay (WtP) of 50€ (Ortúzar & Simonetti, 2008). For perceived risk, we take the value of 5€ per risk level (Shelat et al., 2022). We make no assumptions on mode-specific constants and leave them at 0. A Bayesian efficient design includes standard errors of all priors, indicating the level of certainty. We set these standard errors at half the value of the prior. Given the assumed normal distribution, this results in a 0.975

certainty that the prior has the correct sign (negative for travel time, cost and perceived infection risk, positive for comfort).

Model estimation

The bridging experiment in HII is a regular discrete choice experiment (DCE) and can thus be analysed using choice modelling techniques. We assume that respondents make decisions by maximising their expected utility (Random Utility Maximisation) (McFadden, 1974). As outlined in the Introduction, we apply a Latent Class Choice Model (Greene & Hensher, 2003) on the bridging experiment. We also include socio-demographic information in the class allocation function and iteratively exclude them one-by-one, until only significant predicting socio-demographics remain (significant for at least one of the classes).

Based on the parameter estimates of the LCCM, each respondent is assigned a probability of belonging to each of the market segments. These probabilities are included in the modelling of the rating experiment by means of a WLS regression. WLS is very similar to regular Ordinary Least Squares (OLS) regression, with the weighted sum of squares (WSS) assigned an extra variable – weight (π) – determining the importance of each data point. As a higher probability of belonging to a certain class is analogous to the importance of a data point, we adopt this approach. Equation 1 presents the calculation of WSS, where x represents the attribute level, y the observed perceived risk value and β the estimated parameter, determining the impact of an attribute on the level of perceived infection risk.

$$WSS_s = \sum_{n=1}^N \left(\pi_{n,s} \cdot \left(y_n - \sum_{k=1}^K x_{n,k} \cdot \beta_{s,k} \right)^2 \right) \quad (1)$$

3. RESULTS AND DISCUSSION

Preliminary results, based on 705 valid responses collected through the Dutch railways’ panel (NS, 2020), indicate that our sample can be segmented into four distinct groups with respect to long-distance international trips in Europe. Based on their travel preferences, we denominate the four segments as follows:

1. Time-sensitive travellers
2. Prudent travellers
3. Frequent train-loving travellers
4. Cautious car travellers

Time sensitive travellers are, as the name implies, highly time sensitive (WtP > 70 €/h, compared to 40€/h for the full sample). They prefer the train for shorter trips and flying for longer ones, avoiding the use of car, indicating they may also value the time they have while travelling (not having to drive themselves). With respect to infection risk, they perceive it as a function of time, not as a fixed penalty irrespective of travel time. **Prudent travellers** exercise the most trading-off behaviour, showing the highest variation in mode choice, switching modes quickly when circumstances change. Unlike the previous segment, they see risk as a fixed penalty, irrespective of travel time (but different per mode). **Frequent train-loving travellers** strongly prefer the train, even for very long trips. For them, risk is also seen as a fixed penalty, not based on the duration of travel. Finally, the **Cautious car travellers** will often choose to travel by car. They have the lowest WtP (< 20 €/h) and see infection risk as both a function of time and a fixed penalty. To better highlight the modal share of each segment for different risk levels and trip distances, we

simulate trips varying between 100km and 1500km. The results are presented in a ternary graph in Figure 1.

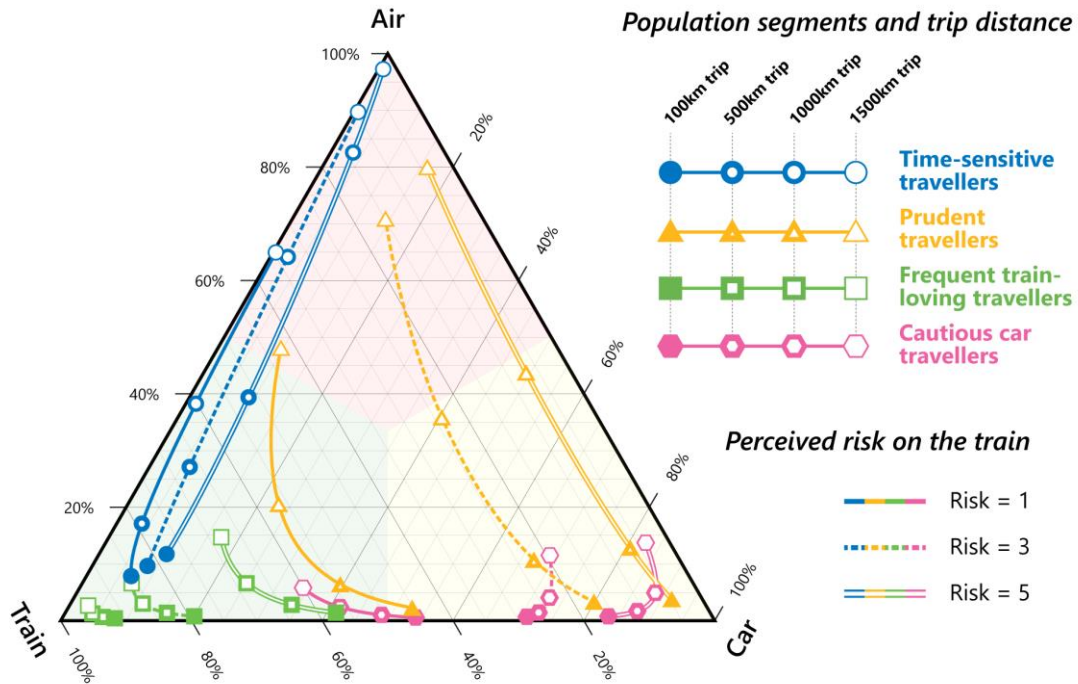


Figure 1. Ternary graph, indicating the modal split between car, train and air, for distances of 100km to 1,500km, for risk levels of 1, 3 and 5 for all four market segments

4. CONCLUSIONS

In this study we present a novel approach of capturing sample heterogeneity in an HII experiment through a latent class choice model, by utilising a weighted least squares regression approach, as opposed to the conventionally used ordinary least squares, when the full sample is analysed as a whole. We apply the model on survey data investigating the impact of COVID-19 infection risk on long-distance travel in Europe.

We identify four distinct user groups. Overall, risk for air travel is rarely seen as time-dependent due to the short flight time and small differences in flight time when travelling in Europe. For trains, some segments see risk as fixed, whereas others as time-dependent (increasingly risky with a longer travel time).

Through the rating experiment, we are also able to uncover which policy measures make travellers more at ease and reduce their perceived risk. This will be investigated further but preliminary outcomes show that some groups prefer to evaluate the data on their own (rate of infections, vaccination rate), whereas others rely on government policies and advice and trust those.

This gives operators and policymakers important knowledge on how to react in situations of increased risk of respiratory and other diseases. It enables them to make informed decisions on which measures to take, how to adjust their policies on masks, cleaning, pricing etc.

ACKNOWLEDGEMENTS

We would like to thank the Dutch railways for providing us help and support in the use of their panel, as well as to the panellist of the Dutch railways' panel for taking part in our survey.

REFERENCES

- Bergantino, A. S., & Madio, L. (2018). High-Speed Rail, Inter-Modal Substitution and Willingness-to-Pay. A Stated Preference Analysis for the “Bari-Rome.” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3091537>
- Bos, I. D. M., Van der Heijden, R. E. C. M., Molin, E. J. E., & Timmermans, H. J. P. (2004). The choice of park and ride facilities: An analysis using a context-dependent hierarchical choice experiment. *Environment and Planning A*, 36(9), 1673–1686. <https://doi.org/10.1068/a36138>
- ChoiceMetrics. (2018). *Ngene1.2 User Manual & Reference Guide*. Retrieved from www.choice-metrics.com
- Currie, G., Jain, T., & Aston, L. (2021). Evidence of a post-COVID change in travel behaviour – Self-reported expectations of commuting in Melbourne. *Transportation Research Part A: Policy and Practice*, 153, 218–234. <https://doi.org/10.1016/j.tra.2021.09.009>
- de Haas, M., Faber, R., & Hamersma, M. (2020). How COVID-19 and the Dutch ‘intelligent lockdown’ change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives*, 6, 100150. <https://doi.org/10.1016/j.trip.2020.100150>
- Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), 681–698. [https://doi.org/10.1016/S0191-2615\(02\)00046-2](https://doi.org/10.1016/S0191-2615(02)00046-2)
- Heufke Kantelaar, M., Molin, E. J. E., Cats, O., Donners, B., & Wee, B. van. (2022). Willingness to use night trains for long-distance travel. *Travel Behaviour and Society*, 29, 339–349. <https://doi.org/10.1016/j.tbs.2022.08.002>
- Kouwenhoven, M., de Jong, G. C., Koster, P., van den Berg, V. A. C., Verhoef, E. T., Bates, J., & Warffemius, P. M. J. (2014). New values of time and reliability in passenger transport in The Netherlands. *Research in Transportation Economics*, 47(1), 37–49. <https://doi.org/10.1016/j.retrec.2014.09.017>
- Louviere, J. J. (1984). Hierarchical Information Integration: A new Method for the Design and Analysis of Complex Multiattribute Judgment Problems. *Advances in Consumer Research*, 11, 148–155. Retrieved from <https://www.acrwebsite.org/volumes/6233/volumes/v11/NA-11/full>
- McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4), 303–328. [https://doi.org/10.1016/0047-2727\(74\)90003-6](https://doi.org/10.1016/0047-2727(74)90003-6)
- Molin, E. J. E., Blangé, J., Cats, O., & Chorus, C. (2017). Willingness to pay for safety improvements in passenger air travel. *Journal of Air Transport Management*, 62, 165–175. <https://doi.org/10.1016/j.jairtraman.2017.04.002>
- Molin, E. J. E., & Timmermans, H. J. P. (2009). Hierarchical information integration experiments and integrated choice experiments. *Transport Reviews*, 29(5), 635–655. <https://doi.org/10.1080/01441640902829470>
- NS. (2020). NS Panel. Retrieved December 6, 2022, from <https://nspanel.nl/>
- Oppewal, H., Louviere, J. J., & Timmermans, H. J. P. (1994). Modeling Hierarchical Conjoint Processes with Integrated Choice Experiments. *Journal of Marketing Research*, 31(1), 92–105. <https://doi.org/10.1177/002224379403100108>
- Ortúzar, J. de D., & Simonetti, C. (2008). Modelling the demand for medium distance air travel

- with the mixed data estimation method. *Journal of Air Transport Management*, 14(6), 297–303. <https://doi.org/10.1016/j.jairtraman.2008.08.002>
- Román, C., & Martín, J. C. (2010). Potential demand for new high speed rail services in high dense air transport corridors. *International Journal of Sustainable Development and Planning*, 5(2), 114–129. <https://doi.org/10.2495/SDP-V5-N2-114-129>
- Román, C., Martín, J. C., Espino, R., Cherchi, E., Ortúzar, J. de D., Rizzi, L. I., ... Amador, F. J. (2014). Valuation of travel time savings for intercity travel: The Madrid-Barcelona corridor. *Transport Policy*, 36, 105–117. <https://doi.org/10.1016/j.tranpol.2014.07.007>
- Shamshiripour, A., Rahimi, E., Shabanpour, R., & Mohammadian, A. (Kouros). (2020). How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. *Transportation Research Interdisciplinary Perspectives*, 7, 100216. <https://doi.org/10.1016/j.trip.2020.100216>
- Shelat, S., Van De Wiel, T., Molin, E., van Lint, J. W. C., & Cats, O. (2022). Analysing the impact of COVID-19 risk perceptions on route choice behaviour in train networks. *PLoS ONE*, 17(3 March), e0264805. <https://doi.org/10.1371/journal.pone.0264805>
- Shortall, R., Mouter, N., & Van Wee, B. (2022). COVID-19 passenger transport measures and their impacts. *Transport Reviews*, 42(4), 441–466. <https://doi.org/10.1080/01441647.2021.1976307>
- Sun, X., Zhang, Y., & Wandelt, S. (2017, May 25). Air transport versus high-speed rail: An overview and research agenda. *Journal of Advanced Transportation*. Hindawi Limited. <https://doi.org/10.1155/2017/8426926>
- Tirachini, A., & Cats, O. (2020). COVID-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, 22(1), 1–34. <https://doi.org/10.5038/2375-0901.22.1.1>

Day-to-day delivery demand management: Evaluation based on routing efficiency and customer satisfaction

Ryota Okazaki^{*1}, Yuki Oyama², Naoto Imura³, and Katsuhiko Nishinari⁴

¹Master Student, Department of Civil Engineering, Shibaura Institute of Technology, Japan

²Associate Professor, Department of Civil Engineering, Shibaura Institute of Technology, Japan

³Project Professor, Research Center for Advanced Science and Technology, University of Tokyo, Japan

⁴Professor, Research Center for Advanced Science and Technology, University of Tokyo, Japan

SHORT SUMMARY

As demand for online shopping and home delivery increases rapidly, courier companies often offer services focusing on customer satisfaction. This places strong constraints on the planning of delivery routes for courier vehicles, making delivery routes inefficient. The objective of this study is to present a framework to evaluate demand management policies in terms of the balance between customer satisfaction and delivery efficiency. To this end, we first estimate a delivery option choice model using the stated choice data of e-commerce users. Then, based on day-to-day delivery demand simulated by the estimated model, we optimize a multi-period vehicle routing problem and evaluate the delivery efficiency. We implement two policies: a surcharge for morning delivery and an expansion of the time slot range. The results show that the former significantly reduces customer satisfaction, while the latter achieves higher customer satisfaction and delivery efficiency.

Keywords: E-commerce, demand management, discrete choice modeling, multi-period vehicle routing problem, last-mile delivery, operations research applications

1 INTRODUCTION

The COVID-19 pandemic has boosted demand for home delivery. Online-shopping customers are placing greater emphasis on speed of delivery, delivery fees, and their own time constraints, resulting in heavy use of express delivery services and demand concentration at certain times of the day. This also imposes severe constraints on delivery route planning, resulting in inefficient deliveries. While delivery demand management (e.g. pricing or slotting) could potentially improve delivery efficiency, changes in service levels could also lead to lower customer satisfaction and fewer orders (Rao et al., 2011; Marium & Arsalan, 2017). Practical demand management requires analyzing customer preferences for delivery options and identifying policies that improve delivery efficiency while retaining customer satisfaction.

E-marketplaces like Amazon often offer their customers several delivery options with different flexibility for delivery, including regular delivery, scheduled delivery where customers can specify the delivery date and time slot, and express delivery such as next-day or same-day delivery. Scheduled delivery may produce the demand concentration on a specific day or time, and express delivery imposes a hard time constraint, while regular delivery gives flexibility to delivery. As such, the delivery option choice behavior of customers highly impacts day-to-day delivery efficiency. Therefore, to analyze the effect of this option choice behavior, delivery efficiency has to be evaluated over multiple days with multiple time slots.

Regarding delivery demand management, the management of delivery time slots has been recently studied, as reviewed by Wałmuth et al. (2023). Agatz et al. (2011) and Köhler et al. (2020) indicated that longer time slots can reduce delivery costs by relaxing time window constraints. However, they did not analyze customers' preferences or model their option choices, and thus could not analyze the policy impact on customer behavior or satisfaction. Most of the studies on delivery time slot management have focused on scheduled delivery services for e-grocers, which require prompt and accurate in-person delivery, and the effect of delivery option choice behavior

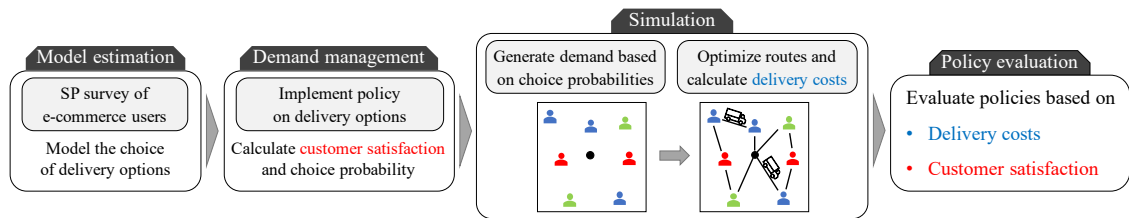


Figure 1: Framework of this study

on delivery efficiency has yet to be analyzed.

Delivery costs are generally calculated by optimizing delivery routes by solving a Vehicle Routing Problem (VRP). The planning period of a typical VRP is a single day, and delivery options such as next-day delivery or regular delivery cannot be considered in the problem. For this reason, it is necessary to apply a multi-period vehicle routing problem (Multi-period VRP) that extends the planning period to multiple days, e.g., Archetti et al. (2015). Most of the studies related to multi-period VRP focus on formulation and runtime, and there are no studies on analyzing delivery efficiency based on customer choice behavior.

The objective of this study is to evaluate and identify policies that achieve the balance between delivery routing efficiency and customer satisfaction. To this end, we first analyze customer choice behavior of delivery options among next-day, regular, and scheduled delivery, by estimating a discrete choice model using the stated choice data. Then, we formulate and optimize a multi-period VRP with due dates and time windows, given the delivery demand simulated by the estimated choice model, and evaluate the routing costs. Finally, based on routing efficiency and customer satisfaction, different scenarios of policies are evaluated. The framework of the study is summarized in Figure 1.

The contribution of the study lies in the following three items.

1. Behavioral analysis of different delivery options

To date, most analyses of delivery option choice have focused on specific delivery options, such as scheduled delivery. In contrast, this study analyzes customer choice among delivery options widely offered by Amazon and other e-marketplaces, including next-day, scheduled, and regular delivery.

2. Evaluating policies based on customer Satisfaction

Most previous studies that attempted to improve delivery efficiency did not take customer satisfaction into account. In this study, to consider more realistic demand management, customer satisfaction was calculated and evaluated as an indicator for policy considering the balance with routing efficiency.

3. Optimize day-to-day delivery routes based on customer choice of delivery options

There is no study that reflects customer choice behavior in the optimization of delivery plans by solving a multi-period VRP. The novelty of this study is that it optimizes day-to-day delivery plans using multi-period VRP to analyze the effect of delivery option choices on delivery efficiency.

2 METHODOLOGY

Stated choice data

We analyzed customer delivery option choice behavior based on responses from a stated preference (SP) survey conducted from April 30 to May 14, 2021. The respondents were among "Kuroneko Members," members of Yamato Holdings Co., Ltd. which has the largest share of parcel delivery in Japan. The survey was conducted on the assumption that respondents would purchase the same products again that they purchased online the previous time. Respondents were asked to select a delivery option among (1) next-day delivery, (2) scheduled delivery (2 to 8 days after order), and (3) regular delivery (date and time can not be specified). Respondents who selected scheduled delivery were additionally asked to select delivery date and time. Each respondent makes one

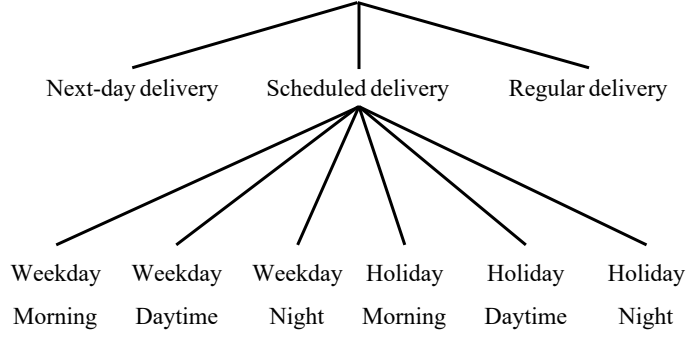


Figure 2: Nest Structure of Choice Model

choice for each of the five different choice scenarios. For each scenario, three attributes were varied: delivery fee, delivery time (the number of days required for delivery), and range of a time slot. For details on the survey design and data, see Oyama et al. (2022). The survey was completed by 4,872 respondents (response rate 4.87%), and given that five scenarios were presented to each respondent, the original sample size is 24360. Data cleaning was then performed based on the experience of online shopping, choice of dominated alternatives, resulting in the final sample size of 18,928.

Delivery option choice model

We then develop a delivery option choice model. In addition to next-day and regular deliveries, we assume six alternatives for scheduled delivery, which are a combination of delivery day (weekday and holiday) and time slot (morning, daytime, and nighttime). With these eight alternatives in total, we apply the Nested Logit (NL) Model with a nest of scheduled delivery based on the assumption that the utilities of the six options are correlated (Figure 2).

The utility V_{ni} of option i for individual n is defined as

$$V_{ni} = ASC_i + \beta_{\text{day}} \cdot \text{day}_i + \beta_{\text{fee}} \cdot \text{fee}_i + \beta_{\text{slot_range}} \cdot \text{slot_range}_i \cdot \delta_{i, \text{scheduled}} \quad (1)$$

where ASC_i is the alternative specific constant for i , day_i is the number of days required to be delivered when choosing i , fee_i is the delivery fee in JPY, and slot_range_i is the time slot range in hours, which implies how long the individual has to be at home to receive the item. The dummy variable $\delta_{i, \text{scheduled}}$ is 1 if i is an option for scheduled delivery and 0 otherwise, implying that slot_range_i only matters for scheduled delivery.

Multi-period VRP

The delivery cost is evaluated on a day-to-day basis, by simultaneously optimizing the delivery routes for multiple days based on a multi-period VRP. Let n be the number of customers, c_{ij} be the travel cost between points i, j (distance between two points), T be the planning horizon. Then, let u_i be the amount of cargo loaded when the vehicle leaves point i , t_i be the arrival time at point i , m be the capacity of the vehicle k , K be the actual number of vehicles in service, and the capacity of all vehicles is Q . For customer i , the holding cost of the package is h_i , the day the package arrives at the depot is o_i , the quantity demanded is q_i , the specified delivery time is e_i to l_i , and the specified delivery day is s_i to d_i ($s_i = d_i$ for next day delivery and scheduled delivery).

The decision variable of the problem is:

$$x_{ijk}^t = \begin{cases} 1 & \text{if vehicle } k \text{ travels from point } i \text{ to } j \text{ on day } t, \\ 0 & \text{otherwise,} \end{cases}$$

and the optimal route is found by solving the following problem:

$$\min_x z(x) \equiv \sum_{k=1}^K \sum_{i=0}^n \sum_{j=0}^n \sum_{t=1}^T c_{ij} x_{ijk}^t + \sum_{i=1}^n \sum_{t_1=o_i}^H h_i (1 - \sum_{t_2=o_i}^{t_1} \sum_{j=0}^n \sum_{k=1}^K x_{ijk}^{t_2}), \quad (2)$$

subject to

$$\sum_{k=1}^K \sum_{j=0}^n \sum_{t=s_i}^{d_i} x_{ijk}^t = 1 \quad (i = 1, 2, \dots, n) \quad (3)$$

Table 1: Delivery option choice

Delivery option (N=18928)	Next-day delivery	44.8%
	Scheduled delivery	38.3%
	Regular delivery	16.8%
Delivery time (N=7253)	Morning	45.9%
	Daytime	28.8%
	Night	25.3%
Delivery Date (N=7253)	Weekday	18.5%
	Holiday	81.5%

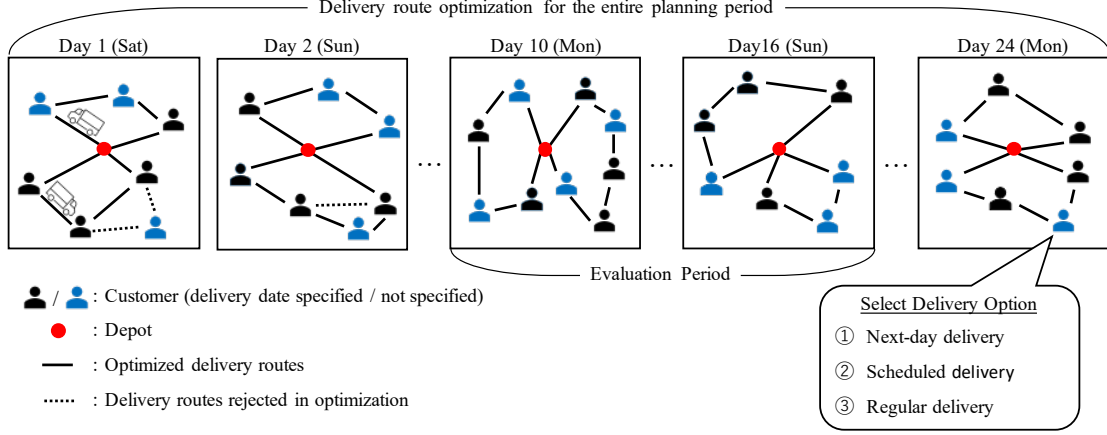


Figure 3: Overview of multi-period vehicle routing problem

$$\sum_{j=0}^n x_{ijk}^t = \sum_{j=0}^n x_{jik}^t \quad (i = 0, 1, 2, \dots, n), (k = 1, 2, \dots, K), (t = 1, 2, \dots, T), \quad (4)$$

$$\sum_{k=1}^K \sum_{j=1}^n x_{0jk}^t \leq m \quad (t = 1, 2, \dots, T), \quad (5)$$

$$u_i - u_j + Q \cdot x_{ijk}^t \leq Q - q_j \quad (t = 1, 2, \dots, T), (k = 1, 2, \dots, K), (i, j = 1, 2, \dots, n), i \neq j, \quad (6)$$

$$t_i - t_j + M \cdot x_{ijk}^t \leq M - c_{ij} \quad (t = 1, 2, \dots, T), (k = 1, 2, \dots, K), (i, j = 1, 2, \dots, n), i \neq j, \quad (7)$$

$$q_i \leq u_i \leq Q \quad (i = 1, 2, \dots, n), \quad (8)$$

$$e_i \leq t_i \leq l_i \quad (i = 1, 2, \dots, n). \quad (9)$$

The objective function (2) includes transportation and holding costs, respectively represented by the first and second terms. Constraint (3) ensures that the delivery is made within the specified date. Constraint (4) ensures that the number of vehicles arriving at and departing from customer i on each day is equal. Inequality (6) and constraint (7), which is called the potential (MTZ) formulation (Miller et al., 1960), are constraints on subtour elimination. Note that M in inequality (7) represents a sufficiently large constant.

Indicators for evaluation

To evaluate demand management policies, this study uses two indicators: (i) routing efficiency and (ii) customer satisfaction. The routing efficiency is calculated as the operation cost for delivery, directly obtained as $z(x^*)$ where x^* is the optimal plan for multi-period vehicle routing given a simulated demand. Customer satisfaction is based on the expected maximum utility (EMU), calculated with the estimated choice model. The EMU is defined as follows:

$$\text{EMU}_n = \frac{1}{\mu} \ln \left\{ \exp(\mu V_{n, \text{Next-day}}) + \exp(\mu V_{n, \text{Regular}}) + \frac{\mu}{\mu_g} \left(\sum_{i \in S} \exp(\mu_g V_{ni}) \right) \right\} \quad (10)$$

where S is the collection of six alternatives for scheduled delivery, and μ_g is a scale parameter associated with the subgroup.

Table 2: Parameter estimation results

Parameter	MNL		NL	
	Estimate	t-stat	Estimate	t-stat
$ASC_{\text{Next-day}}$	0.628	15.21**	0.638	16.68**
$ASC_{\text{Holiday-Morning}}$	0.427	5.68**	0.590	7.59**
$ASC_{\text{Holiday-Daytime}}$	-1.152	-1.50	0.150	1.76
$ASC_{\text{Holiday-Night}}$	0.172	2.19*	0.441	5.11**
$ASC_{\text{Weekday-Morning}}$	-1.054	-12.74**	-0.616	-6.02**
$ASC_{\text{Weekday-Daytime}}$	-1.190	-14.10**	-0.726	-6.87**
$ASC_{\text{Weekday-Night}}$	-0.922	-10.58**	-0.452	-4.21**
β_{day}	-0.221	-18.50**	-0.213	-19.87**
β_{fee}	-0.010	-81.91**	-0.009	-77.51**
$\beta_{\text{slot_range}}$	-0.068	-3.16**	-0.066	-3.08**
μ_g			1.231	28.03**
Sample size		18928		18928
Initial log likelihood		-39359.67		-39359.67
Final log likelihood		-25503.85		-25489.04
Adjusted rho-square		0.3520		0.3524
Likelihood-ratio test	$-2\{L(\hat{\beta}_{\text{MNL}}) - L(\hat{\beta}_{\text{NL}})\} = 29.62 > 3.84 = \chi^2_{1,0.05}$			

3 RESULTS AND DISCUSSION

Table 1 shows an aggregate result, the percentage of each delivery option being chosen. Approximately 45% of observations chose next-day delivery, and among those for scheduled delivery, 46% selected morning delivery. Note that the survey period coincided with "Golden Week" (a major holiday period in Japan), which resulted in approximately 81% of observations choosing a holiday to receive the ordered items.

The estimation result of the delivery option choice model is shown in Table 2, compared with the result of the standard multinomial logit (MNL) model. Both models suggest the same signs for delivery attributes, and the likelihood ratio test suggests the statistical preference of the NL model over the MNL model. The scale parameter μ_g was estimated to be 1.231, which gives $1/\mu_g = 0.812$, confirming an appropriate nest structure. The value of delivery time (VODT) for e-commerce users was calculated as approximately 23.67 JPY/day (0.17 EUR/day at the rate of 140 JPY to 1 EUR). This describes the customer's willingness to pay for a shorter delivery day, implying that, on average, customers are willing to wait one additional day if the delivery fee is about 24 JPY higher. The VODT has been analyzed in different contexts (e.g., Hsiao, 2009; Meister et al., 2023), and ours was calculated in the context of delivery option choice of e-commerce users as in Oyama et al. (2022). Moreover, our estimation result suggests that an increase in the range of time slots decreases the utility. This is because customers have to be at home for a longer period of time.

In this study, we evaluate the following two policies that can change the delivery option choice behavior of customers and have impacts on delivery efficiency.

1: Additional charge of 100 JPY for morning delivery

The SP survey results show that about 45% of those who chose scheduled delivery specified morning delivery, indicating that there is a high demand for morning time slots. Therefore, the number of vehicles in operation may be unevenly distributed depending on the time of day. The pricing is expected to reduce the demand for morning slots and thus may improve the routing efficiency.

2: Change of the time slot range from 2 to 4 hours.

The current service provided by Yamato Holdings Co., Ltd. has a 2-hour time slot for afternoon deliveries. In addition to this, we evaluate the cases with a time slot range of 3 or 4 hours. A longer time slot relaxes the time window constraints for delivery to be satisfied in finding the optimal route, thus improving the efficiency of the delivery route. However, on the other hand, longer slot ranges increase the uncertainty of the delivery time, which is

Table 3: Scenario evaluation result

Scenario	AM +100	Slot range	Cost ^a	Δ Cost ^b	Δ Satis ^c	Δ Cost/ Δ Satis ^d
Bm	×	2h	3300.54	—	—	—
1	○	2h	3200.67	3.03%	-0.129	23.54
2	×	3h	3246.39	1.64%	-0.029	56.75
3	○	3h	3170.72	3.93%	-0.154	25.56
4	×	4h	3202.21	2.98%	-0.057	52.22
5	○	4h	3162.10	4.19%	-0.179	23.50

^aAverage cost of delivery for 100 simulations

^bReduced delivery costs

^cChanges in customer satisfaction

^dCost savings divided by reduction in customer satisfaction

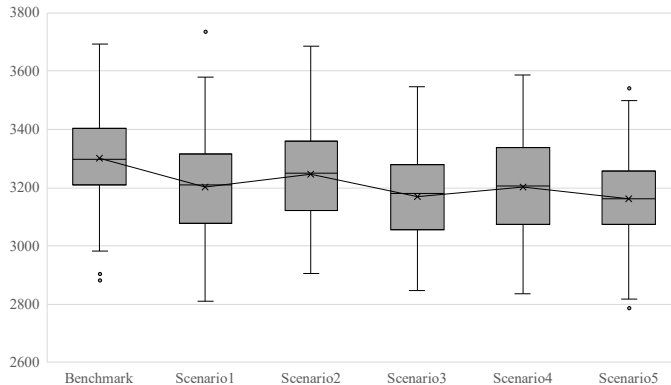


Figure 4: Delivery cost variation of different scenarios

expected to decrease customer satisfaction.

Regarding the multi-period VRP, we set the planning horizon to 24 days, with the first day being Saturday. The demand for home delivery is generated at a random point in a square with side lengths of 100 on each day of the planning horizon. Each customer’s choice of delivery options is simulated according to the choice probability of the NL model. Given the simulated delivery demand, we optimize the delivery schedule for a multi-day period and calculate the operation cost by solving the multi-period VRP, as shown in Figure 3. We implemented the optimization using the Gurobi Optimizer mathematical optimization solver. Since the demand that occurred before the planning period cannot be considered in the early stages of the planning horizon, the evaluation focuses on the seven days from the 10th day (Monday) to the 16th day (Sunday). We summed up the operation costs of the seven days, which is defined as the delivery cost and computed for different policies. For each scenario, we performed 100 runs of the demand simulation and routing optimization.

Table 3 reports the average indicator values over 100 runs, and Figure 4 shows the variation of operation costs, for different scenarios. As a result, both the policies implemented reduced the cost of delivery. In particular, Scenario 5 which introduced the morning slot pricing and expanded the time slot to four hours lowered costs the most, but it also caused the most significant decrease in customer satisfaction. Although charging an additional fee for morning delivery is more effective than expanding the range of time slots in terms of reducing costs, it is not the optimal policy because of the large decrease in customer satisfaction.

To evaluate the balance between delivery efficiency and customer satisfaction, we calculated the cost reduction rate divided by the decrease in customer satisfaction, which is shown in the right-most column of Table 3. The result shows that Scenario 2 has the best value. In conclusion, it is clear that the policy of expanding the range of time slots is more effective than the policy of charging an additional fee for morning deliveries.

4 CONCLUSIONS

This study evaluated policies for day-to-day delivery demand management based on the balance between delivery routing efficiency and customer satisfaction. For such evaluation, we first estimated a delivery option choice model and analyzed the customer preferences. Moreover, the multi-period VRP was formulated and optimized to evaluate the impact of the option choice behavior on the day-to-day delivery efficiency.

Specifically, two policies were considered: an additional charge of 100 JPY for morning deliveries, and the change in time slot range from 2 to 4 hours. While the surcharge policy for morning delivery reduced delivery costs, it also significantly decreased customer satisfaction. The best value was obtained in terms of the balance when the time slot range was 3 hours, regardless of whether the surcharge policy was implemented or not. From the above, we conclude that the optimal time slot range for this case study is 3 hours and that the morning delivery surcharge policy should not be implemented.

Future work includes the improvement of simulation setup conditions. This study assumed all demand during the planning horizon to be generated a priori and did not consider demand dynamically occurring on a day-to-day basis. Therefore, it is not possible to reflect a decrease in delivery efficiency due to sudden demand for next-day delivery. Therefore, a more advanced policy evaluation is needed in the future by incorporating the day-to-day dynamic nature of demand and optimizing the delivery plan with uncertainty.

REFERENCES

- Agatz, N., Campbell, A., Fleischmann, M., & Savelsbergh, M. (2011). Time slot management in attended home delivery. *Transportation Science*, *45*, 435-449.
- Archetti, C., Jabali, O., & Speranza, M. G. (2015). Multi-period vehicle routing problem with due dates. *Computers & Operations Research*, *61*, 122-134.
- Hsiao, M.-H. (2009). Shopping mode choice: Physical store shopping versus e-shopping. *Transportation Research Part E: Logistics and Transportation Review*, *45*(1), 86-95.
- Köhler, C., Ehmke, J. F., & Campbell, A. M. (2020). Flexible time window management for attended home deliveries. *Omega*, *91*, 102023.
- Marium, M. S., & Arsalan, N. (2017). Understanding the impact of service convenience on customer satisfaction in home delivery: evidence from pakistan. *International Journal of Electronic Customer Relationship Management*, *11*, 23-43.
- Meister, A., Winkler, C., Schmid, B., & Axhausen, K. (2023). In-store or online grocery shopping before and during the covid-19 pandemic. *Travel Behaviour and Society*, *30*, 291-301.
- Miller, C. E., Tucker, A. W., & Zemlin, R. A. (1960). Integer programming formulation of traveling salesman problems. *Journal of the Association for Computing Machinery*, *7*, 326-329.
- Oyama, Y., Fukuda, D., Imura, N., & Nishinari, K. (2022). E-commerce users' preferences for delivery options. *arXiv e-prints*.
- Rao, S., Goldsby, T. J., Griffis, S. E., & Iyengar, D. (2011). Electronic logistics service quality (e-lsq): Its impact on the customer's purchase satisfaction and retention. *Journal of business logistics*, *32*, 167-179.
- Waßmuth, K., Köhler, C., Agatz, N., & Fleischmann, M. (2023). Demand management for attended home delivery—a literature review. *European Journal of Operational Research*.

Estimating Household-Level Time-Use within a Week Activity Scheduling Framework – Application of the MDCEV Model

Anna Sophie Reiffer*¹, Peter Vortisch¹

¹Institute for Transport Studies, Karlsruhe Institute of Technology (KIT), Germany

SHORT SUMMARY

Activity-based approaches have become state-of-the-art in travel demand modelling due to their behavioural realism. While there have been great advances in modelling techniques, most studies do not consider the household context, and almost all are limited to the generation of single-day activity schedules. Therefore, we propose an activity generation and scheduling approach for one week, considering the household context. This study provides a general overview over the proposed framework, and further details the model used to generate household-level activity time-use for the period of one week.

The proposed model contributes to the current state of activity-based models as it goes beyond individual, single-day travel demand and allows for analysis of household-level decisions for the modelling period of one week.

Keywords: activity-based model, MDCEV, week activity schedules, time-use

1. INTRODUCTION

Activity-based approaches have become state-of-the-art in travel demand modelling due to their behavioural realism. While there have been great advances in modelling techniques, most studies do not consider the household context, and almost all are limited to the generation of single-day activity schedules. Therefore, we propose an activity generation and scheduling approach for one week, considering the household context. This study provides a general overview over the proposed framework, and further details the model used to generate household-level activity time-use for the period of one week.

Activity-based approaches can be categorised into rule-based and econometric models. Rule-based models rely on hard-coded rules and heuristics, which make them easier to implement. However, this limits their behavioural realism and the ability to generalise model results.

Econometric approaches mitigate these limitations by modelling individual decisions, not through rules and heuristics, but based on the principle of utility maximisation. Bowman and Ben-Akiva (Bowman & Ben-Akiva, 2001) presented the first disaggregate activity-based approach, which generates activity schedules by sequentially modelling individual decisions through (nested) logit models. Although the sequential model of decisions remains a popular approach in activity-based travel demand models, the method has some limitations. The sequence in which the analyst considers the decisions in the model claims that there is an order among the individual decisions. This possibly arbitrary order does not allow for consideration of trade-offs between all choices. This limitation has given rise to the development and application of the multiple discrete-continuous extreme value (MDCEV) model ((Bhat, 2005, 2008)). In this approach, individuals do not consider alternatives as perfect substitutes for each other but simultaneously as a

combination of different activities and the time allocated to them, subject to a time budget constraint. While the first formulation of the model only allowed for modelling aggregated time allocation to each activity type, more recent studies show that the model can also consider activity episodes (Palma et al., 2021) and their order (Saxena et al., 2022) .

Another approach to overcome some of the limitations of sequential models is to consider trade-offs between daily scheduling choices by formulating an optimisation problem ((Manser et al., 2022; Pougala et al., 2022)). In this approach, the objective is to maximise the utility of an individual's schedule through a mixed-integer linear program.

Although the presented approaches all improve state-of-the-art activity-based models, some limitations are worth noting. First, they only consider activities and their schedules for one day. However, past studies highlight the importance of considering multiple days for a more realistic simulation of travel behaviour within travel demand models ((Hilgert et al., 2017; Mallig & Vortisch, 2017)). Furthermore, all choices are considered on an individual level and disregard the context of the household. While this is sensible for some activities like work or work-related activities, the household context influences who conducts certain activities, such as shopping or escorting activities.

However, the proposed approaches cannot simply be transferred from the single-day to the 7-day context. Considering 7-day schedules and household context significantly increases the dimensions of the models, which will likely render the currently defined optimisation problems too large to find a solution within a sensible timespan. Furthermore, we challenge that the underlying assumptions regarding the choice situations of scheduling activities still hold in the 7-day context. In utility theory, we assume that individuals know all possible alternatives within a choice set and choose the one that maximises their utility. Manser et al. (Manser et al., 2022) elaborate on the issue concerning this assumption regarding modelling activity schedules and present a method to generate a feasible choice set. Although the authors propose a sensible approach for single-day activity schedules, it is arguable whether activity schedules of one week actually result from individuals comparing and choosing among a set of alternative schedules or rather from scheduling activities such that they meet a set of constraints.

In this study we combine the idea of activity generation through an MDCEV model and the scheduling using an optimisation approach. The rest of this paper is structured as follows. We will first provide an overview over the activity generation and scheduling framework. Subsequently, we describe the data used in our study and detail the model specification of the MDCEV model. We go on to present the estimation results and conclude our paper with future work and final remarks.

2. MATERIALS AND METHODS

Although the scope of this paper is focusing on the estimation of parameters to model household-level time-use, the motivation for the chosen model specification is rooted within the framework of modelling week activity schedules. This section, therefore, first provides an overview over the proposed framework. We subsequently provide a brief overview over the data used in this study, and finally specify the model.

Activity Generation Scheduling Framework

We propose an activity generation and scheduling approach for one week, considering the household context through a combination of the MDCEV model and a constraint satisfaction optimisation approach. The framework for activity generation and scheduling is illustrated in figure 1. The input data can consist of either 7-day travel diary data or time-use data. Additionally, multiday

data generated through pattern sampling based on single-day data (as proposed by Zhang et al. (Zhang et al., 2018)) is also possible.

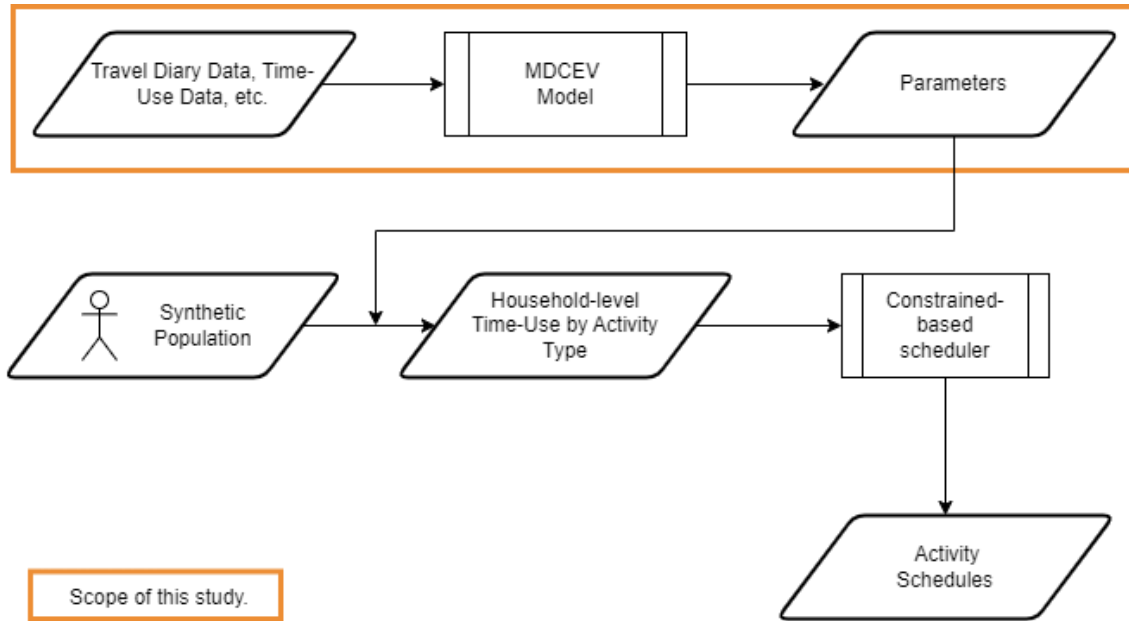


Figure 1 - Activity Generation and Scheduling Framework

Based on this data, we estimate an MDCEV model. The estimated parameters are then applied to the synthetic population of the model region. At this stage, we define the model according to Bhat's (Bhat, 2005) original formulation such that activities and the time allocated to them are predicted at an aggregate level. The model specification and results are scope of this study and detailed in the following sections.

Given the household-level activity types and times, the activity scheduler then considers each time slice of the activities and allocates it to a time slice within a household member's schedule. Similar to Pougala et al. (Pougala et al., 2022) and Manser et al. (Manser et al., 2022), we propose to define an optimisation problem to generate these schedules. However, instead of defining an objective function that needs to be maximised, we propose to solve a constraint satisfaction problem (CSP). In this case, no objective function is necessary. The objective of a CSP is to find a feasible solution subject to previously defined constraints. These constraints can either be defined as hard (hc) or soft constraints (sc). If a hard constraint is violated, the solution is rejected, whereas the violation of a soft constraint results in a penalty. We propose to use constraints that apply to household members and those that apply to the timing of activities:

- Work activities can only be assigned to employed household members (hc)
- Education activities can only be assigned to household members in education (hc)
- Shopping and escorting activities can only be assigned to household members over the age of 16 (hc) and should ideally be conducted by adult household members (sc)
- Work activities are subject to regional employment law, which constrains the total work time per week, the work time per day and required breaks (hc). Soft constraints could be added to, e.g. account for flextime or different durations for lunch breaks.
- Education activities can only be conducted during reasonable times (e.g. secondary school times start between 7:30 and 8 am and finish at the latest at 5:30 pm).
- Shopping can only be conducted during shop opening hours.

Additionally, we include constraints that ensure all activities have a reasonable minimal duration and that frequent switches between short activities are avoided.

Data

The data used in this study stems from the German Mobility Panel (MOP), a longitudinal survey that has been conducted annually since 1994. In the survey, participants report their trips in a 7-day travel diary in addition to providing personal and household information.

For this study, we used data from 2017 to 2019, which includes data on 4.564 households. As the data is collected using a travel diary and not a time-use diary, we had to prepare the data such that it reflects activity time-use. We set the start of each diary to midnight of the first survey period and assigned the time until the first trip to “home”. We repeated the same for the activity of the last trip of the week, setting the end of the diary to midnight on the last assigned survey day. We then determined the time-use for each activity per person and subsequently summarized the values at the household-level.

We differentiate eight alternative activities: home, work, business or work-related (meaning work activity outside the workplace), shopping, escorting someone, education, leisure, and other activities. Further, we have included parameters to account for household information on income (high vs. low), number of children in the household (yes/no), and household size.

Household-Level Time-Use Estimation

The household-level time use is estimated using a MDCEV model approach as it was first presented by Chandra Bhat (Bhat, 2005). The model is specified such that home activities are treated as an outside good. Integrating an outside good ensures the positive consumption of that alternative; in this case the specification results in all individuals conducting a home activity. The problem is defined by:

$$Max \sum_{k=1}^K \frac{\gamma_k}{\alpha} \Psi_k \left(\left(\frac{x_k}{\gamma_k} + 1 \right)^\alpha - 1 \right) \quad (1)$$

Subject to the budget constraint B

$$B = \sum_{k=1}^K x_k \quad (2)$$

where K is the number of considered activities, x_k is the amount of time spent on activity k. The budget of a household is the number of minutes per week (10.080) times the number of household members. The α and γ parameters determine the satiation. In our model, we specified α such that it does not vary over alternatives, while different γ parameters are determined for each alternative. The probability of an observed combination of activities including their duration is given by:

$$P(x_1^*, x_2^*, \dots, x_M^*, 0, \dots, 0) = \frac{1}{\sigma^{M-1}} \left(\prod_{m=1}^M f_m \right) \left(\sum_{m=1}^M \frac{p_m}{f_m} \right) \left(\frac{\prod_{m=1}^M e^{V_i/\sigma}}{(\sum_{k=1}^K e^{W_k/\sigma})^M} \right) (M-1)! \quad (3)$$

We estimated several models accounting for different parameters. The final model is the one in which all parameters were statistically significant.

3. RESULTS AND DISCUSSION

Table 1 provides the estimated parameters of the MDCEV model on household-level activity time use.

Table 1: Parameter Estimates of MDCEV model

	δ -coefficient (utility)	γ -coefficient (satiation)
Work		21.149
Intercept	-5.400	
High Income	0.378	
Business/work-related		5.157
Intercept	-6.546	
High Income	0.503	
Shopping		0.279
Intercept	-2.806	
Household size	-0.137	
Escorting someone		0.489
Intercept	-5.793	
Children under 10 y/0 in the household (yes/no)	1.335	
Education	-6.887	20.988
Leisure		3.470
Intercept	-3.767	
High Income	0.021	
Travel	3.293	0.00288

The results show that travelling has largest the δ parameter indicating that this is the most popular activity. This is not surprising as in our case, all activities (except home) are bound to travelling to a different location. On the other hand, considering the satiation parameter of travel, we can see that the least time is invested in travel. Escorting someone is the least popular activity, but the utility is increased when children under the age of 10 are living in the household. This is sensible, as smaller children are more likely to be escorted e.g., to childcare or school. The satiation parameter is, again, comparatively low meaning that not a lot of time is invested into the activity. Both work and work-related have similar coefficients, albeit work-related activities are slightly less popular. The utilities of both alternatives are increased in households with higher income. The two activities differ considerably regarding their satiation. Work activities have the highest overall satiation parameter indicating that most time is spent working. Although the satiation parameter for work-related activities is still relatively large, it is much smaller compared to the one for work activities at the workplace.

Compared to the other activities, shopping is rather popular. This reasonable, as almost all households conduct some shopping activity throughout the week. Interestingly, the utility of shopping decreases with increased household size indicating that larger households actually conduct fewer

shopping activities per week. While this seems counterintuitive at first, this can be explained by the fact that smaller households are less likely to buy in bulk and, thus, have to go shopping more frequently. The satiation parameter for shopping shows that time spent is rather small.

Leisure activities are less popular than shopping activities, indicating that households conduct fewer leisure activities compared to shopping, however, the satiation parameter shows that a comparatively large amount of time is invested.

There are few comparable studies that allow for a discussion of the model results in relation to other research. Our model is most similar to the aggregated time-use model presented by Palma et al. (Palma et al., 2021). In this study, the authors also find that travelling is considered the most popular activity, whereas work and escorting activities are less popular.

4. CONCLUSIONS

This study presents a household-level activity generation and scheduling framework with a focus on model estimation for the generation of activity time-use. We propose a combined MDCEV model and constrained satisfaction approach as a realistic representation of household level activity time-use and scheduling decisions.

The presented results of the MDCEV model are sensible and show that e.g. shopping activities need to be considered in the household context. It is somewhat difficult to relate the model results to the literature as there are only few studies conducted applying an MDCEV model to activity time use and especially because activities are often defined differently. In future work, we will relate the entire modelling framework to other published work as this allows for a more comprehensive discussion of our results.

At the current stage of the model, we only consider aggregate household-level activity time-use. In further modelling work, we will test to see if the consideration of some activities on the individual level (such as work activities) by specifying the activities as individual alternatives leads to better results.

The proposed model contributes to the current state of activity-based models as it goes beyond individual, single-day travel demand and allows for analysis of household-level decisions for the modelling period of one week.

REFERENCES

- Bhat, C. R. (2005). A multiple discrete–continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, 39(8), 679–707. <https://doi.org/10.1016/j.trb.2004.08.003>
- Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological*, 42(3), 274–303. <https://doi.org/10.1016/j.trb.2007.06.002>

- Bowman, J. L., & Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, 35(1), 1–28. [https://doi.org/10.1016/S0965-8564\(99\)00043-9](https://doi.org/10.1016/S0965-8564(99)00043-9)
- Hilgert, T., Heilig, M., Kagerbauer, M., & Vortisch, P. (2017). Modeling Week Activity Schedules for Travel Demand Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2666(1), 69–77. <https://doi.org/10.3141/2666-08>
- Mallig, N., & Vortisch, P. (2017). Measuring Stability of Mode Choice Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, 2664(1), 1–10. <https://doi.org/10.3141/2664-01>
- Manser, P., Haering, T., Hillel, T., Pougala, J., Krueger, R., & Bierlaire, M. (2022). Estimating flexibility preferences to resolve temporal scheduling conflicts in activity-based modelling. *Transportation*. <https://doi.org/10.1007/s11116-022-10330-8>
- Palma, D., Enam, A., Hess, S., Calastri, C., & Crastes dit Sourd, R. (2021). Modelling multiple occurrences of activities during a day: An extension of the MDCEV model. *Transportmetrica B: Transport Dynamics*, 9(1), 456–478. <https://doi.org/10.1080/21680566.2021.1900755>
- Pougala, J., Hillel, T., & Bierlaire, M. (2022). Capturing trade-offs between daily scheduling choices. *Journal of Choice Modelling*, 43, 100354. <https://doi.org/10.1016/j.jocm.2022.100354>
- Saxena, S., Pinjari, A. R., & Paleti, R. (2022). A multiple discrete-continuous extreme value model with ordered preferences (MDCEV-OP): Modelling framework for episode-level activity participation and time-use analysis. *Transportation Research Part B: Methodological*, 166, 259–283. <https://doi.org/10.1016/j.trb.2022.09.008>
- Zhang, A., Kang, J. E., Axhausen, K., & Kwon, C. (2018). Multi-day activity-travel pattern sampling based on single-day data. *Transportation Research Part C: Emerging Technologies*, 89, 96–112. <https://doi.org/10.1016/j.trc.2018.01.024>

Bayesian Networks for travel demand generation: An application to Switzerland

Aurore Sallard*¹ and Dr. Miloš Balać²

¹PhD student, Institute for Transport Planning and Systems, ETH Zürich, Switzerland

²Senior research assistant, Institute for Transport Planning and Systems, ETH Zürich, Switzerland

SHORT SUMMARY

Bayesian Networks (BNs) are probabilistic graphical models representing conditional dependencies existing between variables of interest. Recent studies have employed BNs for population synthesis and daily activity plan generation. Those studies highlight the ability of BNs to efficiently detect the causality links between variables in an easily interpretable way. This short paper aims to propose a further application of BNs for both population and daily activity plan synthesis in Switzerland. We show that understanding the dependency structure linking the population characteristics and its mobility behaviour is key to generating representative synthetic activity patterns. Furthermore, we lay the foundations for the development of temporally transferable travel demand models.

Keywords: Activity-based modeling; Bayesian networks; Synthetic Populations; Travel demand generation;

1 INTRODUCTION

According to Rasouli & Timmermans (2014), three main categories of activity-based models have been developed since their emergence in the late 60s (Chapin, 1968): constraint-based models (Jones et al., 1983), rule-based models (Arentze & Timmermans, 2000; Guan et al., 2003) and utility-maximization frameworks (Ben-Akiva & Bowman, 1998). Virtually all of them are dependent on a synthetic population: the quality of the outputs of the activity-based models is highly dependent on the quality of the input population. This is why Rasouli & Timmermans (2014) conclude their review paper by leading the research community towards the development of behaviourally rich models allowing the investigation of causality rules.

Approaches based on Markov processes, initially implemented for population synthesis (Farooq et al., 2013) are a step toward this direction. However, the approaches developed in this paper require the researchers to prepare manually the full set of conditional distributions. Sun & Erath (2015) address this issue and introduces Bayesian Networks (BNs) as an efficient tool for population synthesis. This first study was replicated and expanded by Joubert (2018). The first application of BNs for activity pattern generation was proposed by Joubert & De Waal (2020) and extended in de Waal & Joubert (2022). In these studies, the authors focus on the working population of Cape Town, South Africa, and show that their mobility behaviour is linked to the individuals' age, and employment status and to their owning a car and a driving license. All those studies show that BNs avoid over-fitting and scalability issues without being trapped by the curse of dimensionality. Moreover, they are easily interpretable and can detect complex dependency structures. They can create unobserved patterns, contrary to frequentist approaches, and allow the combination of multiple data sources into one single model. Thus, BNs appear as a promising approach in the domain of travel demand generation.

In this short paper, we propose to apply this methodology and open-source software to generate a synthetic population and its daily activity patterns. Our main contributions will be the following: first, develop a model linking population synthesis and activity chain generation using BNs; and, second, highlight the advantages of BNs compared to statistical matching in a "forecasting" experiment.

2 METHODOLOGY

Data

The Micro-census - Mobility and Transport (BFS, 2015) (referred to afterward as MZMV) is a travel survey conducted by the Swiss Federal Office of Statistics every five years. For this study, we focus on the 2010 and 2015 releases. For each edition, around 1% of the Switzerland resident population above the age of 6 is asked to report on their mobility on a certain day: for each of their trips, they have to provide information (among others) about the travel time and distance, the chosen transport mode and the trip purpose. Information about the respondents themselves and their households is collected too. Personal attributes include age, gender, driving license ownership, employment status and level of education, while the household description provide insights into the household size, structure and monthly income. Each observation (household, person, trip) is weighted. In the following, we consider the persons' weights as the focus lies on the individuals' complete activity chains. The trips data set allows us to reconstruct the individuals' activity chains.

After cleaning and removing incomplete trip data, the data set contains 50 576 personal records for 2015 and 57 087 for 2010. Those individuals reported 170 541 trips in 2015 and 190 308 in 2010 (average number of trips per respondent: 3.37 in 2015 and 3.33 in 2010). Both data sets contain around 3 000 distinct records of activity chains (2 880 in 2015 and 3 054 in 2010). The maximum activity chain length observed in the data set is 22, and 95.5% have a length lesser or equal to 7. Thus, to keep the BN structure concise and follow the approach of Joubert & De Waal (2020), we focus only on the observations where the activity chain length is not greater than 7.

Bayesian Networks

Bayesian Networks (Jensen et al., 1996) (BNs) are probabilistic graphical models consisting of two parts: a structure - one of a directed acyclic graph in which the vertices represent random variables, and the edges correspond to dependency links between the vertices - and parameters, which are joint probability tables encoding the probability distribution of the random variables. The structure and the parameters are estimated from the data or manually defined by an expert. The two approaches can complement each other. Learning the structure of a BN is an unsupervised learning problem. In this study, we use Python 3.9.7 and the library pgmpy (Ankan & Panda, 2015). An implementation of the Hill-Climb search algorithm (Selman & Gomes, 2006) based on the Bayesian Dirichlet equivalent uniform (BDEU) score (Heckerman et al., 1995) is used for the network structure estimation. The parameter learning is based on a maximum-likelihood estimator (White, 1982).

Learning the BN structure

Our goal is to compare the BN approach with the statistical matching algorithm (D'Orazio et al., 2006), which was implemented in the Switzerland eqasim scenario (Hörl & Balac, 2021a). Because of constraints inherent to this scenario, all activity chains not starting or not ending at home had to be removed from the data set. The base idea behind statistical matching is to link each agent with one activity chain record based on the weight and five socioeconomic attributes: age class, household size class, municipality type, sex, and marital status (Hörl & Balac, 2021b). Those attributes are usually obtained from national censuses, which only report a limited set of such socio-economic variables. For the BN estimation, three other attributes are included: the monthly household income, the respondent's employment status, and their ownership of a driver's license. Two main classes of variables are thus considered: the socioeconomic variables, among which are the five matching attributes, and the seven activities forming the activity chain.

Most of the BN's structure was estimated from the data, yet we imposed the following constraints. First, the five attributes mentioned above (age class, household size class, municipality type, sex, and marital status) are seen as "root" nodes in the network. This means that they cannot have parent variables. Second, we are interested in detecting how socioeconomic attributes influence activity chains. Consequently, we impose that an "activity" node can only influence the following activities. Finally, to sample a synthetic population from the BN, we generate a new activity chain for each observation - consisting of the set of socio-economic attributes - present in the training data set using the conditional probability tables estimated in the previous step.

3 RESULTS AND DISCUSSION

Replicating a given distribution

The first experiment aims to verify that the proposed travel demand generation approach is able to replicate a given activity chain distribution. Only the most recent release of the MZMV is used. The structure of the learned Bayesian Network is depicted in Figure 1. One can distinguish two “layers” in the network: on the top are socio-economic variables while the activities are on the bottom. The employment and the driving license ownership connect the two parts of the network, which is similar to the findings of Joubert & De Waal (2020). Figure 2a shows the comparison between the activity chain distributions. Dark blue bars represent the prevalence of activity chains sampled from the Bayesian Network while gray bars correspond to the distribution computed from the input data. Light blue bars represent the activity chain prevalence distribution obtained from statistical matching. The comparison shows that the statistical matching replicates almost perfectly the input distribution, which is confirmed by a Wasserstein distance (Vallender (1974))¹ between the two distributions of around 0.09. The distance between the distribution sampled from the BN and the input data is higher (around 0.12), which we can observe on Figure 2a by the larger gaps existing for some activity chains, such as the under-represented “h-w-s-h” or the over-represented “h-w-h-w-h”, using the abbreviations of activity names introduced in Figure 2. Still, those differences disappear when we focus on the aggregated count of activities, as represented in Figure 2b.

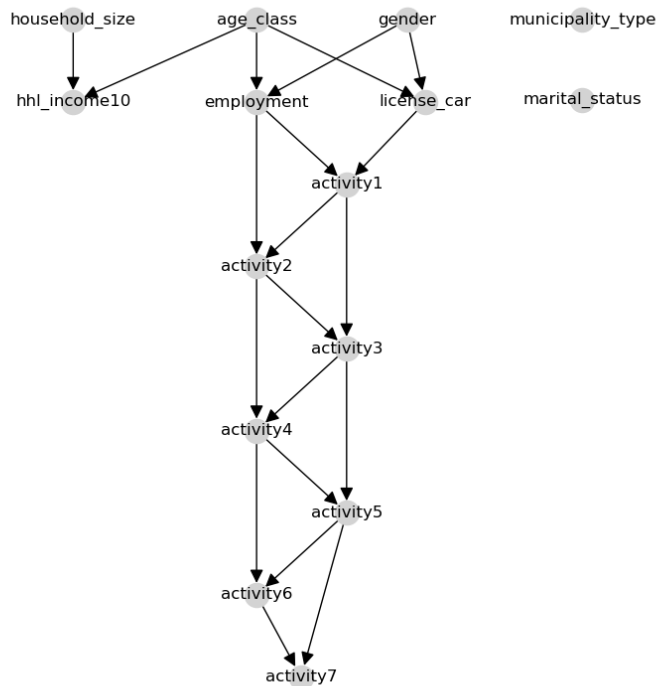
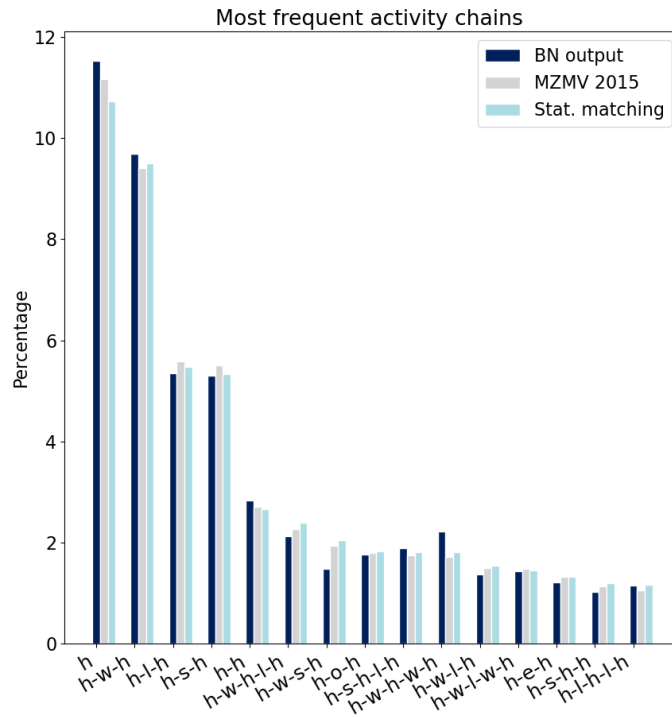


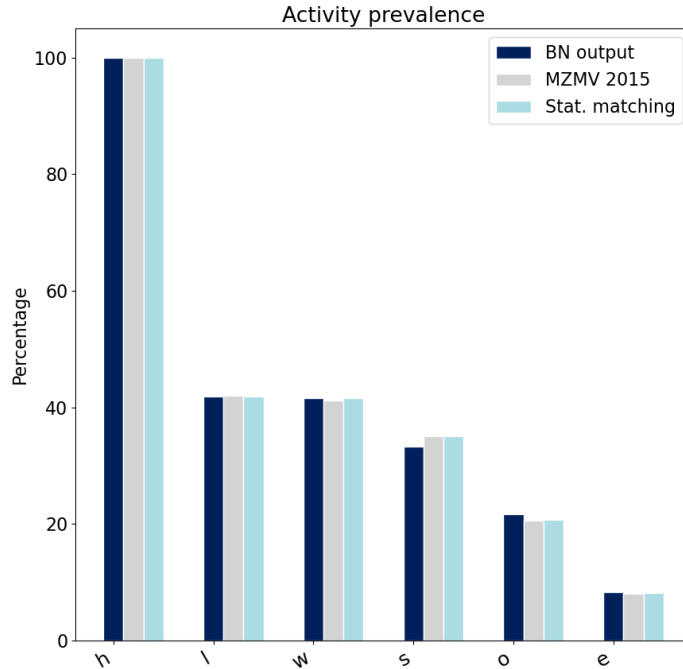
Figure 1: Bayesian Network structure.

Consequently, both Bayesian Networks and statistical matching are suitable methods when it comes to replicating a given distribution of activity chains. The relatively weaker performance of the BN, which we pointed out while computing the Wasserstein distances, can be explained by its ability to generate unobserved activity chains: looking at all the activity chains generated by the BN, regardless of their prevalence, 47.3% of them are absent from the training data set and were “created” by the BN. However, those activity chains are very rare and represent only 5.4% of the agents’ daily plans. An advantage in favor of the BN approach yet seems to stand out when one combines data sources from different time contexts, as the next experiment shows.

¹Here, instead of considering the entire range of activity chains, we take into account only the 50 most prevalent activity chains, so as to ensure that all activity chains are observed a minimal number of times.



(a) Most prevalent activity chains in the MZMV 2015, obtained from statistical matching and sampled from the BN.



(b) Percentage of activity chains containing at least one activity of each purpose.

Figure 2: Activity chain distribution and prevalence of each activity type. The six activities considered in this study are the following: home (h), work (w), education (e), leisure (l), shopping (s) and other (o).

Table 1: Accuracy, precision and F-score of the Bayesian Networks and of the Statistical matching approach

Chain	MZMV 2015 prevalence	BN prevalence	Statistical matching prevalence	BN			Statistical matching		
				Accuracy	Precision	F-score	Accuracy	Precision	F-score
h	10.7%	10.3%	10.6%	82.7%	18.0%	17.6%	82.2%	16.4%	16.3%
h-w-h	9.5%	8.5%	8.1%	85.0%	17.6%	16.6%	84.7%	14.5%	13.3%
h-l-h	5.5%	5.3%	5.5%	90.0%	6.9%	6.7%	89.8%	6.8%	6.8%
h-s-h	5.3%	5.7%	5.9%	90.1%	9.5%	9.9%	89.8%	8.1%	8.5%
h-h	2.6%	2.5%	2.7%	95.0%	2.8%	2.8%	94.9%	2.9%	2.9%
h-w-h-l-h	2.3%	2.0%	2.2%	95.9%	3.9%	3.6%	95.6%	3.4%	3.3%
h-w-h-w-h	1.8%	2.6%	2.4%	95.8%	2.7%	3.2%	95.9%	2.7%	3.1%
h-w-s-h	1.8%	1.3%	1.9%	97.0%	2.9%	2.3%	96.5%	4.7%	4.7%
h-s-h-l-h	1.7%	1.7%	1.7%	96.7%	3.5%	3.5%	96.7%	1.5%	1.5%
h-e-h	1.6%	1.4%	1.3%	97.3%	10.6%	9.9%	97.3%	7.3%	6.7%
Average				90.3%	9.8%	9.6%	90.1%	8.6%	8.4%

Towards temporally transferable travel demand generation models

In this second experiment, the BN is estimated using the older release of MZMV, dating back to 2010, while, as in the previous experiment, the population for which we generate the activity chains is the set of respondents of MZMV 2015. The obtained network has the same structure as the one depicted in Figure 1; however, the conditional probability tables changed as the training data set is different. To generate the corresponding data with the statistical matching algorithm, activity chains extracted from MZMV 2010 were matched to MZMV 2015 respondents.

Aggregated performance indicators: Similarly as before, the aggregated performance of both approaches can be measured with the Wasserstein distance. The distance between the activity chain distribution estimated from the BN and the one from the MZMV 2015 is 0.172, almost exactly the same as the distance from the statistical matching distribution and the reference data, which is 0.171. This shows that, although the BN method has a disadvantage because it is able to generate unseen activity patterns, it can compensate it by better reacting to changes in the socio-economic structure of the population, such as those one can observe between the two releases of MZMV.

Disaggregated performance indicators: Beyond the Wasserstein distance, disaggregated indicators such as accuracy, precision and F-score can be used to compare the performance of the two approaches. Those indicators are presented in Table 1 for the 10 most prevalent activity chains; the averages are related to the 15 most prevalent activity chains. The three indicators show similar values to the ones presented in Joubert & De Waal (2020) and de Waal & Joubert (2022). Moreover, they show that the BN approach outperforms in almost all cases the statistical matching algorithm.

Discussion

A detailed analysis shows that those better results are mostly linked to the fact that the BN captures with a higher accuracy the links between the population characteristics and its activity chains. More precisely, the statistical matching algorithm cannot be implemented using more than a few matching attributes, as explained in Hörl & Balac (2021b), and those attributes must be chosen by the researchers. Figure 1 highlighted that the employment status and the ownership of a driving license, which are not used for matching in Hörl & Balac (2021b), are directly influencing the mobility behavior. In the previous experiment, the fact that those variables were extracted from the most recent release of the MZMV thus led to improved results. To confirm this, a similar experiment was realized without sampling those two attributes from MZMV 2015: the distributions were kept unchanged compared to MZMV 2010. The average accuracy, precision and F-score are presented in Table 2. The table shows that, in this case, the BN approach is outperformed by statistical matching. Consequently, the BN identified which attributes are linking the population characteristics with its observed mobility behavior, which results in a more representative synthetic travel demand.

Table 2: Average accuracy, precision and F-score, over the 15 most prevalent activity chains, obtained with the statistical matching algorithm, the BN method using non-matching attributes and without using the non-matching attributes.

	Accuracy	Precision	F-score
BN with employment status and driving license	90.3%	9.8%	9.6%
BN without employment status and driving license	90.0%	8.4%	8.2%
Statistical matching	90.1%	8.6%	8.4%

4 CONCLUSIONS

This short paper presents an application of Bayesian Networks for synthetic population and travel demand generation. Contrary to Joubert & De Waal (2020), who focus on the employed population living in Cape Town, we can synthesize agents representative to the Swiss people, except for children below six years of age. This study is based on open-source software. We used both aggregated and disaggregated indicators to evaluate our approach. We showed that BNs could accurately replicate a given activity chain distribution and outperform the statistical matching algorithm when combining two data sources in a “forecasting” experiment. We highlighted that the ability of BN to identify the critical socio-economic attributes influencing the activity chain is of great help to generating representative synthetic population and travel demand. Several points indicate a potential for future research: this study will first be extended to include experiment results obtained from the 2005 release of MZMV. Moreover, the networks were estimated using a combination of data-driven, machine-learned methods and expert knowledge, as some constraints about the links’ directions were imposed. It would thus be relevant to conduct sensitivity analyses to evaluate the impact of imposing these constraints. A third possible future research direction is to estimate networks specific to given socio-economic categories. In this way, one could capture and represent the differences in dependency structures and ultimately contribute to the understanding of social mechanisms leading to heterogeneity in mobility behaviour.

ACKNOWLEDGEMENTS

This study is part of a National Research Program (NFP78) funded by the Swiss National Science Foundation.

REFERENCES

- Ankan, A., & Panda, A. (2015). pgmpy: Probabilistic graphical models using python. In *Proceedings of the 14th python in science conference (scipy 2015)*.
- Arentze, T., & Timmermans, H. (2000). *Albatross: a learning based transportation oriented simulation system*. Citeseer.
- Ben-Akiva, M. E., & Bowman, J. L. (1998). Activity based travel demand model systems. In *Equilibrium and advanced transportation modelling* (pp. 27–46). Springer.
- BFS. (2015). *Mikrozensus Mobilität und Verkehr*. <https://www.bfs.admin.ch/bfs/de/home/statistiken/mobilitaet-verkehr/erhebungen/mzmv.html>. (Accessed: 2022-05-03)
- Chapin, F. S. (1968). Activity systems and urban structure: A working schema. *Journal of the American Institute of Planners*, 34(1), 11–18.
- de Waal, A., & Joubert, J. W. (2022). Explainable Bayesian networks applied to transport vulnerability. *Expert Systems with Applications*, 209, 118348.
- D’Orazio, M., Di Zio, M., & Scanu, M. (2006). *Statistical matching: Theory and practice*. John Wiley & Sons.
- Farooq, B., Bierlaire, M., Hurtubia, R., & Flötteröd, G. (2013). Simulation based population synthesis. *Transportation Research Part B: Methodological*, 58, 243–263.

- Guan, J., Roorda, M. J., & Miller, E. J. (2003, January). Approximation of 24 Hour Travel Times in the Greater Toronto Area. *Presented at the 82nd Annual Meeting of the Transportation Research Board, Washington DC.*
- Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning bayesian networks: The combination of knowledge and statistical data. *Machine learning*, 20(3), 197–243.
- Hörl, S., & Balac, M. (2021a). Introducing the eqasim pipeline: From raw data to agent-based transport simulation. *Procedia Computer Science*, 184, 712–719.
- Hörl, S., & Balac, M. (2021b). Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transportation Research Part C: Emerging Technologies*, 130, 103291.
- Jensen, F. V., et al. (1996). *An introduction to Bayesian networks* (Vol. 210). UCL press London.
- Jones, P. M., Dix, M. C., Clarke, M. I., & Heggie, I. G. (1983). *Understanding travel behaviour* (No. Monograph).
- Joubert, J. W. (2018). Synthetic populations of South African urban areas. *Data in brief*, 19, 1012–1020.
- Joubert, J. W., & De Waal, A. (2020). Activity-based travel demand generation using Bayesian networks. *Transportation Research Part C: Emerging Technologies*, 120, 102804.
- Rasouli, S., & Timmermans, H. (2014). Activity-based models of travel demand: promises, progress and prospects. *International Journal of Urban Sciences*, 18(1), 31–60.
- Selman, B., & Gomes, C. P. (2006). Hill-climbing search. *Encyclopedia of Cognitive Science*, 81, 82.
- Sun, L., & Erath, A. (2015). A Bayesian network approach for population synthesis. *Transportation Research Part C: Emerging Technologies*, 61, 49–62.
- Vallender, S. (1974). Calculation of the Wasserstein distance between probability distributions on the line. *Theory of Probability & Its Applications*, 18(4), 784–786.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, 1–25.

Simulated Annealing in a Co-Evolutionary, Agent-Based Transport Modeling Framework - The Example of Ride-pooling Driver Supply Optimization

Nico Kuehnel^{*1}, Shivam Arora², Felix Zwick¹, and Qin Zhang³

¹MOIA GmbH, Germany

²Department of Civil Engineering, IIT Roorkee, India

³School of Engineering and Design, Technical University of Munich, Germany

SHORT SUMMARY

This paper introduces an integrated simulated annealing optimization method within the co-evolutionary agent-based transport modeling framework MATSim, using a small illustrative ride-pooling service as an example to optimize driver shift supply for a given and static demand. Simulated annealing is a metaheuristic optimization algorithm that has already been employed in a wide range of problems and domains. MATSim makes use of a co-evolutionary design in which individual agents try to optimize their daily schedule by finding optimal transport options. The iterative nature of both simulated annealing and MATSim’s co-evolutionary design makes the implementation straightforward and compatible. The outcomes validate the feasibility of the approach in optimizing specific components of the transport model and indicate its potential for future use in comparable applications. The presented case of driver supply optimization may help to design scenarios for new services and to better assess the efficiency and costs of such a service.

Keywords: agent-based transport model, demand-supply-matching, MATSim, on-demand mobility, operations research, simulated annealing.

1 INTRODUCTION

Transport systems are complex and involve various stakeholders, multiple modes of transportation, and numerous decision-making processes. Agent-based transport models, such as MATSim (Multi-Agent Transport Simulation, Horni et al., 2016), enable researchers to simulate and analyze the behavior of individual travelers and their interaction in the context of the wider transport system. The central feature of MATSim’s is a co-evolutionary algorithm that enables individual agents to optimize their daily activity schedules through the identification of the most optimal travel options. Consequently, the decisions made by one agent can have implications for every other agent, often impacting shared resources such as road or bus capacity.

In numerous simulation studies, the supply side of the transportation system is treated as static while searching for a stable demand equilibrium. Typically, the effects of alterations to the supply side are analyzed *across* multiple simulations rather than *within* a single simulation. While this approach is suitable for many use cases, there are situations where the supply side must also react dynamically within the simulation. One example is the simulation of competing minibus operators that do not operate on fixed schedules, but are instead demand-driven and can be modeled using an evolutionary algorithm (Neumann, 2014). Other examples include public transport pricing and supply planning (Kaddoura et al., 2015), the implementation of traffic actuated or traffic adaptive transport signals that dynamically respond to current traffic flows (Kühnel et al., 2018), the optimization of charging infrastructure placement (Fadranski et al., 2023) or tour planning in freight applications (Zilske & Joubert, 2016). An in-depth discussion about optimization problems in (iterative) and stochastic simulation frameworks is given in (Flötteröd, 2017).

A prime example of an inherent necessity for supply-side optimization can be found in recent demand-responsive transport (DRT) systems such as online ride-hailing and ride-pooling, wherein an operator must dynamically respond to passenger requests. To address such cases, a central dispatcher optimizes the fleet supply using various algorithms such as insertion heuristics (Maciejewski, 2016) or integer linear programming (Alonso-Mora et al., 2017). A recent study high-

lighted the need for explicit simulation of operational aspects such as charging, hub facilities, and driver shift supply to yield realistic results for current human-operated fleet operations in which the drivers are employees of the operator (Zwick et al., 2022). This underscores the importance of generating realistic driver shift plans that minimize operating costs while maintaining high-quality service. For simulation studies without available historical data, or those subject to new policies or other changing conditions, the driver shift plans must be scheduled to match the anticipated demand. Given that these scheduling problems are typically NP-hard (Chuin Lau, 1996), simulated annealing (SA) has proven to be a valuable tool for addressing this challenge (Thompson, 1996).

SA is a metaheuristic optimization algorithm that has been applied across a range of domains, including transportation, to address complex optimization problems. SA can be integrated with co-evolutionary algorithms to optimize distinct facets of the transport system, as both approaches share important similarities and fall into the category of *general iterative algorithms* (Youssef et al., 2001). In this study, we provide a case illustration of SA employed within MATSim to optimize the supply side of DRT systems in the form of driver shift planning, based upon initial findings (Arora, 2021).

We demonstrate that SA is well-suited for the iterative design of a co-evolutionary transport model framework and can generate favorable outcomes for optimizing specific simulation components, as exemplified by driver shift planning. The versatility of the SA approach lies in its adaptability to various mobility-related optimization problems, including but not limited to DRT stop or hub placement, charging strategies, traffic signal plans, or fleet sizes. Consequently, the generic SA implementation within the MATSim framework will be made available as an open-source feature to encourage further optimization research in the field.

2 METHODOLOGY

Based on the DRT extension by Maciejewski (2016) including its default re-positioning strategy (Bischoff & Maciejewski, 2020) and the operational aspects described in Zwick et al. (2022) we simulate a human-operated ride-pooling service in MATSim and use SA to optimize the driver shift plan.

The general outline of SA is as follows:

1. Initialize the system with an initial solution λ_0 and set the initial temperature T_0 (to a high value).
2. Choose a candidate solution λ_i for iteration i by making a perturbation to the current solution.
3. Calculate the energy (or cost) difference between the candidate solution cost $c(\lambda_i)$ and the current accepted solution cost $c(\lambda_a)$.
4. If the energy difference is negative, accept the candidate solution as the new current solution.
5. If the energy difference is positive, accept the candidate solution with a probability $P_i(\lambda_i)$ that depends on the current temperature T_i and the energy difference. This probability decreases as the temperature decreases and is designed to allow the algorithm to escape from local minima.
6. Decrease the temperature according to a cooling schedule.
7. Repeat steps 2-6 until the stopping criterion is met (e.g., a maximum number of iterations is reached).

The acceptance probability in step 5 is calculated as:

$$P_i(\lambda_i) = e^{-\left(\frac{k \cdot (c(\lambda_i) - c(\lambda_a))}{T_i}\right)} \quad (1)$$

Multiple cooling schedules exist to adjust temperature T_i . In this study we use an exponential multiplicative schedule:

$$T_i = T_0 * \alpha^i, \quad (2)$$

with a constant $0 < \alpha < 1$.

In our application of optimizing driver shifts, we assume an infinite pool of drivers and limit the optimization to a single day. Shifts s are characterized by their start and end times and are of a fixed duration $t_{d,s}$ of either 5 or 8 hours. An 8-hour shift requires a mandatory break of duration $t_{b,s} = 60$ min, which must occur no earlier than 3.5 hours and no later than 5.5 hours into the shift. The solution λ is defined as a shift plan, i.e., the set of n shifts $\lambda = \{s_1, s_2, \dots, s_n\}$.

For the cost function, we propose a function that

1. sums up the driving hours of all shifts (shift durations $t_{d,s}$ minus optional break durations $t_{b,s}$) and multiplies it with the cost per operational hour θ ,
2. subtracts from the costs the sum of revenues ϵ_r of all served rides $r \in R(\lambda)$ served with solution λ ,
3. adds a penalty δ for each time bin t in which the rejection rate η_λ was greater than a predefined threshold η_{max} :

$$c(\lambda) = \theta \cdot \sum_{s \in \lambda} t_{d,s} - t_{b,s} - \sum_{r \in R(\lambda)} \epsilon_r + \sum_t \Gamma(t, \lambda) \cdot \delta, \quad (3)$$

$$\text{with } \Gamma(t, \lambda) = \begin{cases} 1 & \eta_\lambda(t) > \eta_{max}, \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

For the revenue of a ride, we propose a generic cost function that consists of a base fare β_0 and a distance-dependent price per kilometer β_{km} :

$$\epsilon_r = \beta_0 + \beta_{km} \cdot d_r, \quad (5)$$

where d_r is the distance of ride r .

For step 2 of the algorithm, multiple perturbations were defined to allow an extensive but guided exploration of possible solutions:

Add shift This strategy randomly adds a new shift to the shift plan by drawing from a weighted random distribution of possible time spans. Using a sliding window approach, each possible time span over time bins t gets a weight that relates to the request rejection rates $\eta(t)$ of the iteration of the last accepted solution. The more rejections a possible time span covers, the higher the probability of being selected. The time spans have fixed durations of either 5 or 8 hours, which are the two possible shift durations employed in this study.

Remove shift This strategy randomly removes a shift from the plan. Similar to the *add shift* strategy, a weighted selection from existing shifts in the plan is performed. Here, the weight is calculated by the efficiency of the shift, defined as the ratio of revenue earned over the cost of the shift (duration times the cost per hour) during the last accepted solution's iteration.

Move shift This strategy randomly moves the start of a shift forwards or backwards in time. The time difference is randomly drawn from a uniform distribution and respects the service times of the service.

Duplicate shift This strategy randomly duplicates an existing shift, by drawing from a weighted distribution. Similar to the removal of shifts, the weights are defined by the efficiency of a shift during the last accepted solution's iteration, with more effective shifts being more likely of being duplicated.

Change shift duration This strategy changes the duration of an existing shift by randomly choosing between a 5- and an 8-hour shift.

The SA is implemented in parallel to MATSim's usual iterative cycle as shown in figure 1. The mobility simulation is used in both, the SA algorithm and MATSim's standard demand co-evolution and represents the joint environment to allow the evaluation of the solution (set). The actual evaluation (i.e. scoring/cost updates) and preparation of new solutions (replanning/solution update) are performed in separate cycles. The solution update includes the cooling schedule, the decision for accepting the latest solution and perturbing the accepted solution. For the present study, we

assume the demand to be static and only optimize the driver shift plan given a fixed demand to show the applicability of the approach. Thereby, the simulation framework mimics an iterative traffic assignment model for fleet simulation only, ignoring any additional modes such as private cars.

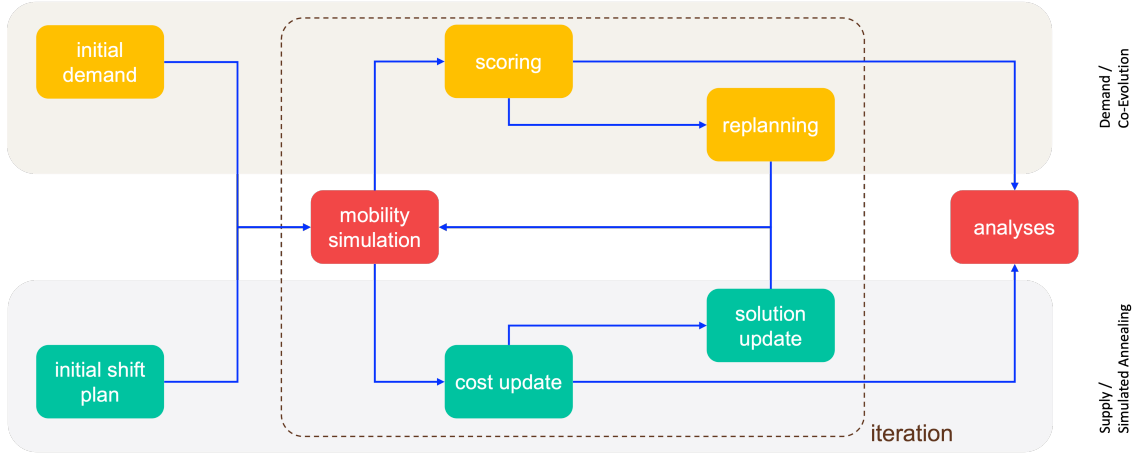


Figure 1: Updated MATSim cycle adapted from Horni et al. (2016) to include the implemented simulated annealing cycle which runs in parallel to the original MATSim cycle.

To test our implementation, we make use of a small existing scenario for the city of Holzkirchen in southern Bavaria, Germany. The scenario has been described by Zwick et al. (2021) and is available open source¹. The temporal distribution of DRT requests is shown in figure 2.

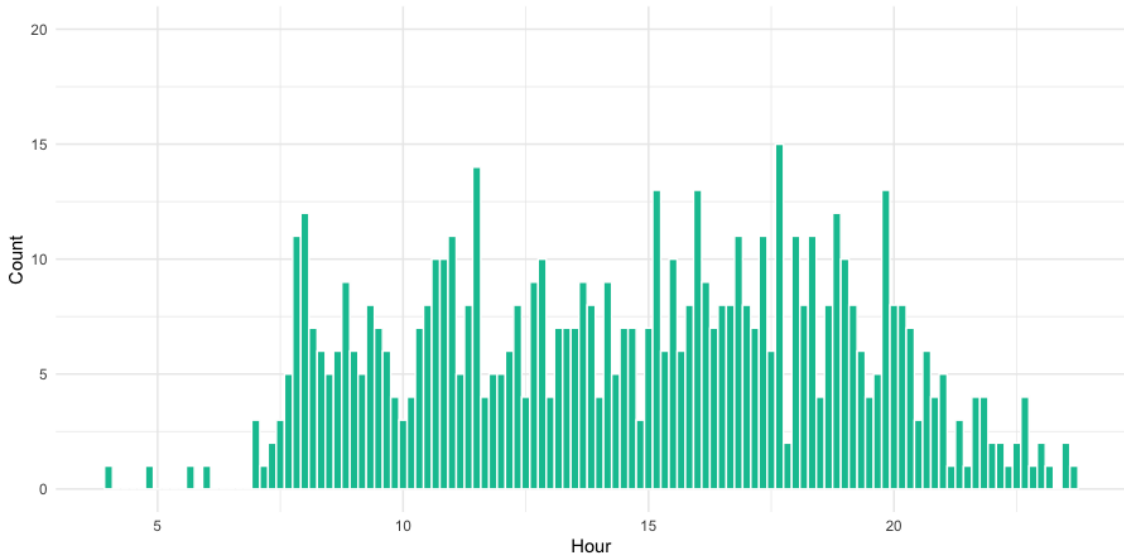


Figure 2: DRT requests over the time of day in the Holzkirchen scenario (Zwick et al., 2021).

In total, we let the simulation run for 300 MATSim iterations. As the rebalancing algorithm of the DRT also needs to adapt to previous iterations, we apply a ratio of 2 MATSim iterations per SA iteration. In addition, we choose to have 3 SA iterations per cooling cycle (making it 6 MATSim iterations per cycle). The best solution found over all iterations will be fixed for the last 3 iterations and serves as the final output. The initial temperature T_0 is set to 1,500 and α to 0.85. In every SA iteration, we randomly choose between 1 and 10 perturbations from the strategies defined above. The cost per driver hour θ is set to EUR 30, to assume somewhat realistic costs of bus drivers,

¹<https://github.com/matsim-org/matsim-libs/tree/master/examples/scenarios/holzkirchen>

including some overhead (Frank et al., 2008). The cost penalty δ equals EUR 50, which has been found to be a good value to ensure minimum service level, with the rejection rate threshold η_{max} set to 0.15. On the revenue side, we choose $\beta_0 = 5 \text{ EUR}$ and $\beta_{km} = 0.7 \text{ EUR/km}$. The values given above do not reflect a coherent business case and are chosen for illustrative purposes. The shift breaks are not part of the optimization in this study, although perturbations on their time windows would also be possible. The initial shift plan is a manually created plan with a lot of oversupply (see iteration 0 in figures 4 and 5). In total, a maximum of 20 DRT vehicles may be employed in the simulation.

3 RESULTS AND DISCUSSION

The figures presented below illustrate the progression of solution quality and associated costs. Figure 3 displays the initial, accepted, current, and best overall costs across iterations, as well as the temperature curve that depicts the cooling schedule. The initial solution is characterized by high costs resulting from a considerable oversupply. In the first few iterations, the costs improve considerably and converge towards a minimum of approximately -1334 EUR. The erratic behavior of the current cost curve indicates that the algorithm is searching around the accepted solution space. The accepted cost curve reveals that, particularly around iteration 100 when the temperature remains high, the current accepted solution may be allowed to be inferior to a prior solution from earlier iterations. In general, the asymptotic nature of the curves suggests a relatively stable and optimized solution, which the algorithm reached in iteration 232.

It is worth noting that negative costs indicate a profitable service, as revenues exceed driver costs. However, the cost and revenue factors presented here are for illustrative purposes only and may not accurately reflect actual scenarios. Additionally, the fixed demand in the Holzkirchen scenario was estimated based on the assumption of an autonomous service with lower cost factors.

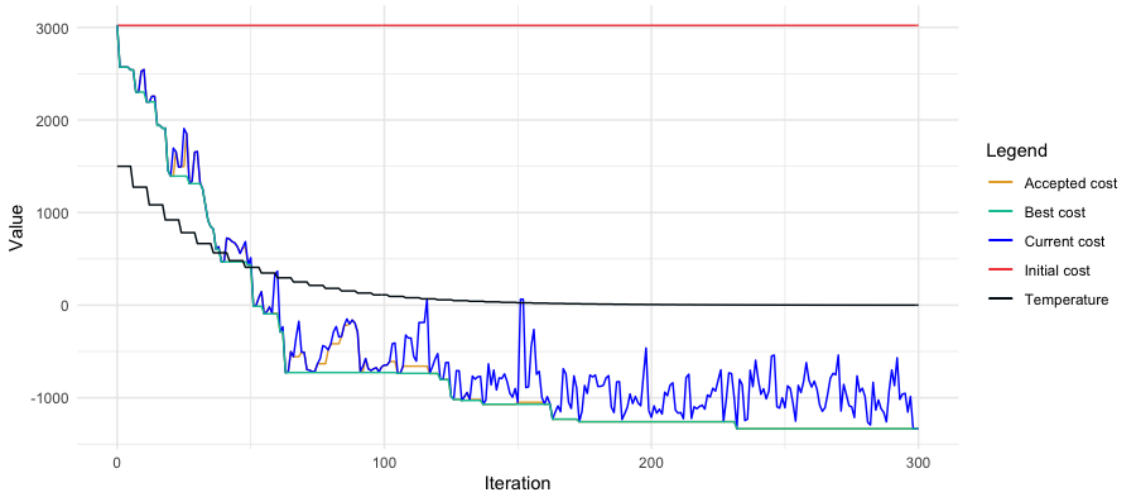


Figure 3: The different cost values (in EUR) and temperature over the course of the simulation.

Figure 4 depicts six plots of shift histograms and rejections at different iterations during the simulation. The algorithm aims to minimize driver hours by limiting the number of rejections to below a specified threshold. In the initial iteration, shifts are evenly distributed throughout the day, resulting in few rejections. Subsequent shift histograms exhibit a more detailed shift schedule with fewer total shifts. The final shift setup displays a significantly reduced shift histogram with an acceptable rejection rate (overall rejection rate of 4 %). The plots illustrate that the algorithm occasionally adds meaningless shifts randomly, such as shifts starting at midnight in iteration 60, which are subsequently eliminated. Additionally, the final shift plan is consistent with the demand pattern in figure 2, with a noticeable peak around 5 pm. In this hypothetical scenario, it shows that a fleet size of around 10 vehicles may be sufficient.

Figure 5 illustrates the vehicle occupancy at the same six iterations as the shift histograms in

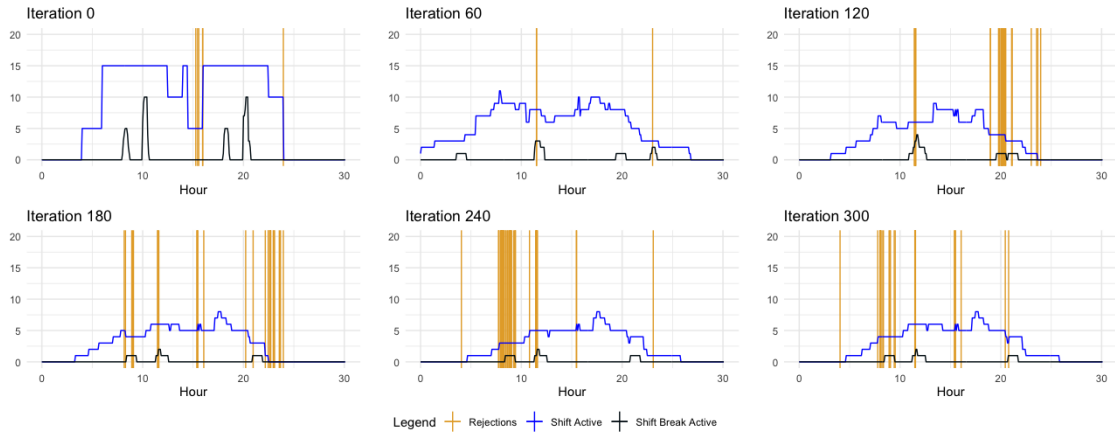


Figure 4: Shift histograms and rejections over the iterations 0, 60, 120, 180, 240 and 300. Each vertical line represents a rejected request. The blue and black lines depict the number of active shifts and shift breaks, respectively.

figure 4. The occupancy plot for iteration 0 reveals that too many shifts were scheduled given the demand, with "STAY" indicating vehicles on shift without tasks. A similar yet less pronounced pattern is present in iteration 60. In contrast, the shift schedules in iterations 240 and 300 demonstrate an efficient utilization of shifts based on demand, indicating that the algorithm has likely found a satisfactory solution to the problem.

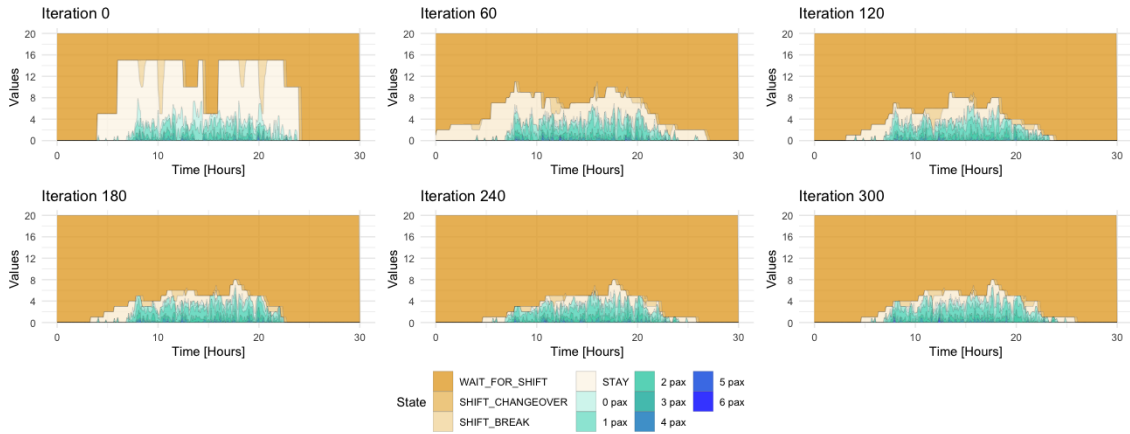


Figure 5: Occupancy plots over the iterations 0, 60, 120, 180, 240 and 300. It shows the distribution of vehicle states over the time of day, including the number of boarded passengers.

4 CONCLUSIONS

In conclusion, the results presented in this study demonstrate the applicability of the simulated annealing algorithm for the given problem. However, it is important to note that these results are only illustrative and require further refinement of input parameters, particularly in relation to driver costs, revenues, and demand. In addition, the cost optimization could be extended to also include other operating costs such as the distance-dependent costs of electricity for electric vehicles. Given that one can also infer the maximum amount of simultaneously operating vehicles one can also estimate the number of required vehicles in the fleet. Given the unpredictable nature of demand, an approximate and valid solution of supply is sufficient for the problem, which supports the idea of using heuristic approaches such as SA. Future research should focus on exploring adaptive and co-evolving demand, testing larger and realistic scenarios, and optimizing other components of the simulation to improve the accuracy of the results. The findings presented here may also be transferable to other iterative transport models.

DECLARATION OF INTERESTS

We acknowledge that Nico Kuehnel and Felix Zwick are employed at the ride-pooling operator MOIA.

REFERENCES

- Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., & Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(3), 462–467. doi: 10.1073/pnas.1611675114
- Arora, S. (2021). *Optimization of driver shift (and break) schedule using simulated annealing in ride-pooling services* (Master’s thesis, School of Engineering and Design, Technical University of Munich). doi: 10.13140/RG.2.2.16415.64162
- Bischoff, J., & Maciejewski, M. (2020). Proactive empty vehicle rebalancing for demand responsive transport services. *Procedia Computer Science*, *170*, 739–744. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1877050920306220> doi: <https://doi.org/10.1016/j.procs.2020.03.162>
- Chuin Lau, H. (1996). On the complexity of manpower shift scheduling. *Computers Operations Research*, *23*(1), 93–102. Retrieved from <https://www.sciencedirect.com/science/article/pii/0305054894000940> doi: [https://doi.org/10.1016/0305-0548\(94\)00094-0](https://doi.org/10.1016/0305-0548(94)00094-0)
- Fadranski, D., Syré, A. M., Grahle, A., & Göhlich, D. (2023). Analysis of charging infrastructure for private, battery electric passenger cars: Optimizing spatial distribution using a genetic algorithm. *World Electric Vehicle Journal*, *14*(2). Retrieved from <https://www.mdpi.com/2032-6653/14/2/26> doi: 10.3390/wevj14020026
- Flötteröd, G. (2017). A search acceleration method for optimization problems with transport simulation constraints. *Transportation Research Part B: Methodological*, *98*, 239–260.
- Frank, P., Friedrich, M., & Schlaich, J. (2008). Betriebskosten von Busverkehren schnell und genau ermitteln. *Der Nahverkehr*, *11*.
- Horni, A., Nagel, K., & Axhausen, K. (Eds.). (2016). *Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press. doi: 10.5334/baw
- Kaddoura, I., Kickhöfer, B., Neumann, A., & Tirachini, A. (2015). Agent-based optimisation of public transport supply and pricing: impacts of activity scheduling decisions and simulation randomness. *Transportation*, *42*, 1039–1061.
- Kühnel, N., Thunig, T., & Nagel, K. (2018). Implementing an adaptive traffic signal control algorithm in an agent-based transport simulation. *Procedia Computer Science*, *130*, 894–899. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1877050918304484> doi: <https://doi.org/10.1016/j.procs.2018.04.086>
- Maciejewski, M. (2016). Dynamic Transport Services. In *The Multi-Agent Transport Simulation MATSim* (pp. 145–152). Andreas Horni, Kai Nagel and Kay W. Axhausen. Retrieved from <https://matsim.org/the-book>
- Neumann, A. (2014). A paratransit-inspired evolutionary process for public transit network design. Retrieved from https://depositonce.tu-berlin.de/bitstream/11303/4393/1/neumann_andreas.pdf
- Thompson, G. M. (1996). A simulated-annealing heuristic for shift scheduling using non-continuously available employees. *Computers Operations Research*, *23*(3), 275–288. Retrieved from <https://www.sciencedirect.com/science/article/pii/0305054895000127> doi: [https://doi.org/10.1016/0305-0548\(95\)00012-7](https://doi.org/10.1016/0305-0548(95)00012-7)
- Youssef, H., M. Sait, S., & Adiche, H. (2001). Evolutionary algorithms, simulated annealing and tabu search: a comparative study. *Engineering Applications of Artificial Intelligence*, *14*(2), 167–181. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0952197600000658> doi: [https://doi.org/10.1016/S0952-1976\(00\)00065-8](https://doi.org/10.1016/S0952-1976(00)00065-8)
- Zilske, M., & Joubert, J. W. (2016). Freight Traffic. In *The Multi-Agent Transport Simulation MATSim* (pp. 155–156). Andreas Horni, Kai Nagel and Kay W. Axhausen. Retrieved from <https://matsim.org/the-book>

- Zwick, F., Kuehnel, N., & Hörl, S. (2022). Shifts in perspective: Operational aspects in (non-)autonomous ride-pooling simulations. *Transportation Research Part A: Policy and Practice*, 165, 300-320. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0965856422002294> doi: <https://doi.org/10.1016/j.tra.2022.09.001>
- Zwick, F., Kuehnel, N., Moeckel, R., & Axhausen, K. W. (2021). Ride-pooling efficiency in large, medium-sized and small towns -simulation assessment in the munich metropolitan region. *Procedia Computer Science*, 184, 662-667. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1877050921007195> doi: <https://doi.org/10.1016/j.procs.2021.03.083>

Analyzing Network-wide Energy Consumption of Electric Vehicles in a Multimodal Traffic Context: Insights from Drone Data

Sijia Sun^{*1,2}, Hossam M. Abdelghaffar², Sérgio F. A. Batista², Mónica Menéndez²,
and Yuanqing Wang¹

¹Key Laboratory of Transport Industry of Management, College of Transportation Engineering, Chang'an University, PO Box 487, South 2nd Ring Road., Xi'an, China, 710064

²Division of Engineering, New York University Abu Dhabi, Saadiyat Marina District PO Box 129188 - Abu Dhabi, United Arab Emirates

SHORT SUMMARY

The environmental benefits and driving range of electric vehicles are closely related to their energy consumption. In this paper, we analyze the energy consumption characteristics of electric mobility systems in a multimodal urban traffic context by establishing the aggregated relationships between macroscopic fundamental diagram (MFD) dynamics and network-wide energy consumption. To do this, we utilize a data-based approach, combining vehicle trajectories collected by a swarm of drones in the downtown areas of Athens, Greece, during the pNEUMA experiment with microscopic energy consumption models. We assume all the trajectories are driven by electric vehicles yet maintain the same behavior observed in the pNEUMA dataset. Preliminary results show well-defined relationships between aggregated traffic parameters and energy consumption at a network level. The total energy consumption of electric cars and buses in the network increases linearly with vehicle accumulation under uncongested traffic conditions. At the same time, the energy consumption per distance traveled by electric buses significantly decreases as the spatial mean speed increases. While for electric cars, the impact of spatial mean speed on energy consumption is marginal, especially when the average speed is above 10 *km/h*.

Keywords: Electric mobility, Energy consumption, pNEUMA dataset, Macroscopic fundamental diagram.

1 INTRODUCTION

Climate change, mainly caused by carbon dioxide emissions from human activities, severely threatens human health and the planet's ecosystem (Zhang et al., 2020). The transportation sector could play an essential role in climate change mitigation, as the sector is responsible for the highest energy consumption in 40% of countries globally and contributes to approximately 15% of total greenhouse gas emissions (IEA, 2022). It is worth mentioning that road transportation is the largest source of transport emissions, accounting for 69% of the sector's overall emissions (IPCC, 2022). This situation will be even more alarming in the decades to come as the trend toward motorization continues (Gao & Newman, 2018).

The electrification of vehicle fleets has been widely recognized as a crucial path to decarbonizing and alleviating fossil fuel dependency in the road transportation sector. Battery electric vehicles (EVs) represent an advanced and promising technology that offers an opportunity to increase energy efficiency and achieve 'zero emissions' compared to their traditional fossil fuel-powered counterparts (Xie et al., 2020). However, EVs are not truly 'zero emissions' from the life cycle perspective as they consume electrical energy, and the indirect emissions produced by the electricity generation are non-negligible, especially where carbon-intensive grids operate. This highlights that the environmental benefits provided by EVs are directly dependent on their energy consumption. Moreover, energy consumption determines vehicle driving range, and the limited driving range remains one of the significant barriers to the massive adoption of EVs. In this context, optimizing EVs' energy consumption plays a vital role in advancing the development of a more sustainable transportation system while concurrently alleviating concerns surrounding range anxiety for EVs.

Most studies so far have focused on minimizing the energy consumption of EVs from two perspectives. One is adopting an eco-driving strategy, which provides drivers with recommendations for modifying their driving behavior to avoid the high energy consumption caused by aggressive

driving patterns (Y. Zhang et al., 2022; Donkers et al., 2020; Bingham, 2012). The other is developing an eco-routing strategy, which incorporates the energy-saving potentials when planning routes for electric vehicle operation (Ahn et al., 2021; Basso et al., 2019; Fiori et al., 2018). Although these strategies are efficient and reliable, their real-world application is limited in terms of scope, as they typically apply to a few routes or trips and a single transportation mode. While several studies have investigated the impact of eco-routing and eco-driving strategies on network-wide energy consumption, they mostly resorted to traffic simulators and used the simplest energy consumption models (e.g., energy consumption is linear to distance) for algorithm simplicity (see e.g., Rakha et al., 2012; Hiermann et al., 2019)). Moreover, to the best of our knowledge, no study yet has analyzed the energy consumption characteristics of EVs in a multimodal traffic network. In this paper, we address this gap, combining real multimodal traffic data with microscopic energy consumption to analyze the network-wide energy consumption of electric mobility systems. Existing vehicle energy consumption models can generally be classified as either macroscopic or microscopic (Othman et al., 2019). Macroscopic models use a single value (i.e., energy per unit distance or time) to roughly calculate the energy demand of vehicles. On the other hand, microscopic models provide a more accurate estimation of energy consumption based on high-resolution driving profile data; however, such data are not easily obtained, especially on a large scale. To overcome this, some studies have used macroscopic fundamental diagram (MFD)-based traffic models to estimate network-wide vehicle environmental externalities (e.g., CO_2 emissions). The MFD describes the well-defined relationships between network production, accumulation, and speed (Geroliminis & Daganzo, 2008). Shabihkhani & Gonzales (2014) proposed an analytical model to estimate the network emissions leveraging the relationship between MFD and the driving cycle. They further evaluated this model in an idealized homogeneous network. Saedi et al. (2020) developed a network-wide emission modeling framework by combining the network fundamental properties with the microscopic emission model. This framework was applied to an urban network through simulation. Recently, Barmounakis et al. (2021) combined large-scale drone data with the MOVES emission model to establish the relationships between network accumulation, speed, and vehicle emissions. They referred to this relationship as the emission-MFD. However, these studies only focused on traditional fossil fuel-powered vehicles, and the impact of electrified technology on network-scale energy consumption is still unclear.

In this paper, we focus on analyzing the network-wide energy consumption characteristics of electric vehicles in a multimodal urban traffic context. We do so by utilizing a data-based approach, combining high-resolution vehicle trajectory data collected by a swarm of ten drones in the central business district of Athens, Greece, during the pNEUMA experiment (Barmounakis & Geroliminis, 2020) with microscopic energy consumption models. This allows us to estimate the large-scale vehicular energy consumption and further investigate the aggregated relationship between network-wide energy consumption and macroscopic fundamental diagram dynamics. We refer to this aggregation relationship as energy consumption-MFD, following the naming method proposed in Barmounakis et al. (2021). The analysis conducted in this paper paves the foundation for optimizing the energy consumption and environmental footprint of electric vehicles in multimodal traffic networks.

The remainder of this paper is organized as follows. In Sect. 2, we describe the pNEUMA dataset and the data processing method. We also briefly introduce the microscopic energy consumption models utilized for different vehicle types. In Sect. 3, we discuss the preliminary results on the aggregated energy consumption at the multimodal urban network. In Sect. 4, we draw the main findings of this paper.

2 METHODOLOGY

In this section, we first introduce the pNEUMA dataset and describe the data pre-processing method. We then present the energy consumption models utilized for electric vehicles.

Data source and pre-processing

The pNEUMA experiment was conducted in the central business district of Athens, Greece, in October 2018 (Barmounakis & Geroliminis, 2020). This experiment collected nearly half a million naturalistic vehicle trajectories in a $1.3 [km^2]$ urban area using a swarm of ten drones during morning peak hours (8:00 - 10:30) over four weekdays. Figure 1 shows the overview of the whole study area and the subareas flown by each drone. The pNEUMA dataset records



Figure 1: Study area of the pNEUMA experiment and drone-assigned subareas and flight routes (Barmounakis & Geroliminis, 2020).

vehicle trajectory information in 0.04-second time intervals, including longitude, latitude, speed, longitude acceleration, latitude acceleration, and timestamp. Due to the multimodal urban traffic characteristics in the selected study area, six vehicle types are recorded in the dataset: car, taxi, bus, motorcycle, medium vehicle, and heavy vehicle. In this paper, we focus on three vehicle types, i.e., car, taxi, and bus.

Extensive pre-processing of the empirical dataset is necessary because measurement errors were detected during the observation period of some drones. We follow the pNEUMA dataset pre-processing method proposed by Hamm et al. (2022), removing the records in the last 2 minutes of each drone flight from the dataset. In addition, we filter some unreasonable records based on the mechanical properties of vehicles and the real-world driving conditions in downtown areas. For example, bus records with instantaneous acceleration greater than $3.5 [m/s^2]$, car records with average travel speeds higher than $80 [km/h]$, and vehicle records with zero instantaneous speed and acceleration throughout the whole trajectory (probably are parked vehicles).

Microscopic energy consumption modeling

According to the information reported in (Barmounakis et al., 2021), during the pNEUMA experiment in Athens, the fuel type of taxis and buses was diesel, and the fuel type of cars was gasoline. For the analysis in our paper, we assume all three vehicle types are electric-powered, yet they maintain the same behavior as that observed in the pNEUMA dataset. We adopt the VT-CPEM (Virginia Tech Comprehensive Power-based Energy consumption Model) to calculate the energy consumption of electric cars/taxis (Fiori et al., 2016). For electric buses, we use the microscopic power-based energy consumption model (Ma et al., 2021). Both models belong to the microscopic backward-looking longitudinal dynamic models, which estimate vehicles' energy consumption based on the calculation of tractive force. In particular, these models produce the energy consumption in units of $[kwh/km]$ and the instantaneous energy consumption in units of $[kw]$ using the instantaneous speed profile as the input. Such input data can readily be provided by the pNEUMA dataset. Previous studies have widely utilized these models and demonstrated their accuracy in estimating the energy consumption of vehicles in the urban traffic context (Ahn et al., 2020; Ma et al., 2021). For the mathematical details of these models, interested readers could refer to the above references.

It is also worth mentioning that road grade has a significant influence on the energy consumption of electric vehicles (Liu et al., 2017). The city of Athens is surrounded by mountains, resulting in relatively large terrain fluctuations. Therefore, the impact of road grade on vehicle energy consumption should not be omitted. In this paper, we use the Shuttle Radar Terrain Mission (SRTM)

digital elevation model to obtain road elevation information and then calculate the road slope between every two consecutive records in the dataset (Farr et al., 2007).

3 RESULTS AND DISCUSSION

In this section, we describe the empirical results regarding two aggregated relationships in the network, specifically, (i) the relationship between accumulation and network-wide total energy consumption; and (ii) the relationship between spatial mean speed and network-wide energy consumption per distance traveled. Before showing the results, we discuss how we calculate the network fundamental properties and network-wide energy consumption.

In this paper, we use vehicle trajectory data collected from 8:30 to 11:00 on October 24th. After pre-processing the dataset as described in the previous section, we gathered records of 36282 vehicle trajectories. Among them are 34820 trajectories of private cars and taxis and 1462 trajectories of buses. We consider 1 minute as the time interval T for aggregating the MFD dynamics and energy consumption results. For each time interval, the accumulation n_r and the spatial mean speed of buses v_r (including private cars and taxis) in the network r are determined as:

$$n_r = \frac{\sum_{i=1}^{N_{car}} tt_i}{T} \quad (1)$$

$$v_r = \frac{\sum_{i=1}^{N_{car}} td_i}{\sum_{i=1}^{N_{car}} tt_i} \quad (2)$$

where N_{car} [veh] is the number of cars circulating in the network during the given time interval; tt_i [s] is the time spent by car i in the network during the time interval, and td_i [m] is the distance traveled by car i during the time interval.

The total energy consumption of car traffic in the network for each time interval is determined as:

$$EC_r = \sum_{i=1}^{N_{car}} ec_i \quad (3)$$

where ec_i [kWh] is the energy consumption of car i during the time interval.

The energy consumption per distance traveled [kWh/veh.km] is calculated with Eq. 4:

$$ECD_r = \frac{EC_r}{\sum_{i=1}^{N_{car}} td_i} \cdot 1000 \quad (4)$$

Regarding the traffic dynamics and energy consumption of buses, we use the same equations for calculations.

Macroscopic relationship between accumulation and total energy consumption

Figure 2 depicts the total energy consumption of cars and buses in the network as a function of accumulation (i.e., energy consumption-MFD). The value of each blue or green data point in the figure represents the aggregated energy consumption of cars or buses in the network over a given period (i.e., 1 minute). For electric cars, we observe that when the accumulation is smaller than 1600 [veh], the total energy consumption increases roughly linearly with the increase in accumulation. This is because when the car traffic in the network is not heavy, the total energy consumption of the system increases correspondingly with the number of vehicles in the network. However, when the traffic conditions become congested, the additional effects of congestion make the relationship between car accumulation and energy consumption non-linear (refer to the blue points in Figure 2 (a) when the accumulation is larger than 1600 [veh]). This is because heavy traffic leads to congestion and lower speeds, which means that cars spend more time traveling the same distance. As a result, the energy consumed by cars in the network further increases. We also observe that for electric buses, the total energy consumption shows a growing trend with the increase in accumulation. Considering the empirical dataset has a limited range of observations, especially for public transport vehicles, our empirical energy consumption-MFD may only represent

the aggregated relationship between accumulation and total energy consumption during a part of the network’s loading and unloading cycles.

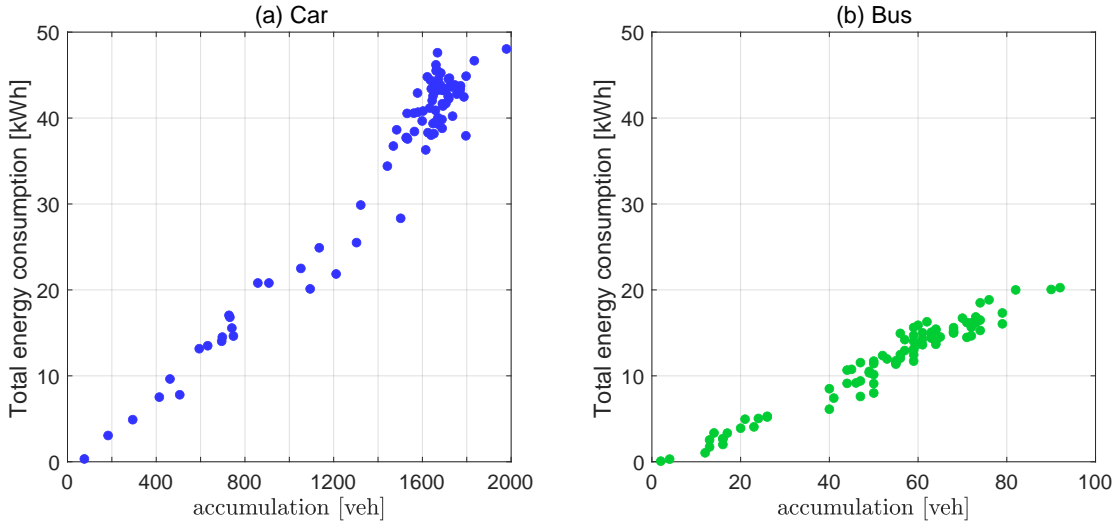


Figure 2: Correlation between total energy consumption and accumulation.

Macroscopic relationship between mean speed and energy consumption

Figure 3 depicts the aggregated relationships between average speed and the energy consumption normalized for distance on a network scale. We observe that when the average speed is lower than 10 km/h, the energy consumption of electric cars decreases with the increase in average speed. However, this decreasing trend tends to be minimal, and the energy consumption of electric cars is basically constant when the average speed is higher than 10 km/h. In contrast, the energy consumption of electric buses significantly decreases as the average speed increases (at least within the typical speed range for buses in the pNEUMA dataset). The different relationships between mean speed and the energy consumption of electric cars and electric buses could be attributed to their differences in vehicle configurations, such as motor power, vehicle mass, drag resistance coefficient and rolling resistance coefficient. Electric cars are lightweight and aerodynamically efficient, which results in approximately constant energy consumption over a wide range of speeds. In comparison, electric buses are heavy and have high rolling resistance. As the speed increases, the rolling resistance reduces, while the aerodynamic drag resistance only slightly increases, leading to decreasing energy consumption.

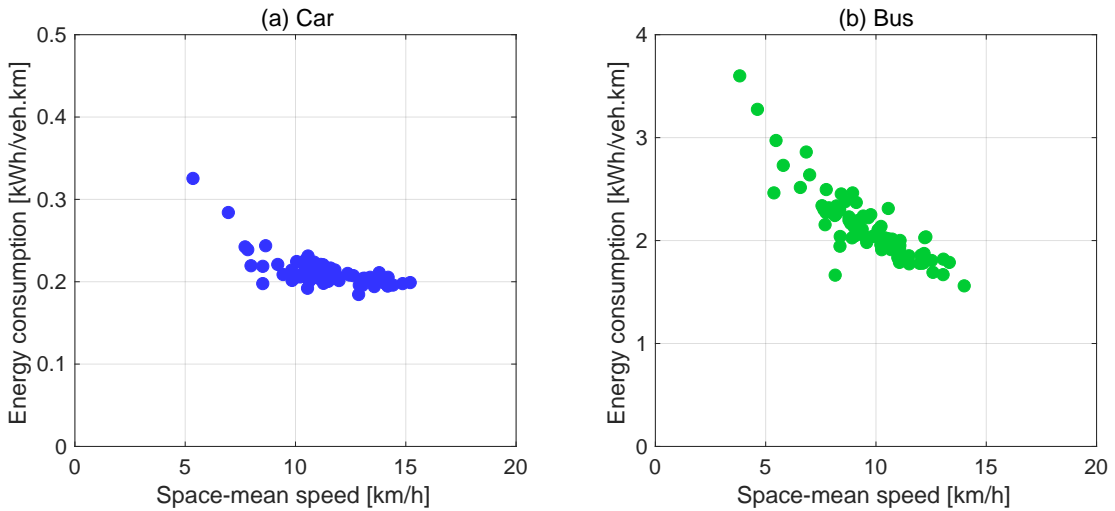


Figure 3: Correlation between energy consumption per distance traveled and spatial mean speed.

4 CONCLUSIONS

This paper extends the macroscopic fundamental diagram (MFD) to the energy consumption-MFD by investigating the aggregated relationships between network traffic dynamics and the energy consumption of electric mobility systems in a multimodal urban traffic context. We use a data-based approach, combining naturalistic vehicle trajectories collected by a swarm of drones during the pNEUMA experiment with microscopic energy consumption models. We assume all the trajectories belong to electric vehicles yet maintain the same behavior observed in the pNEUMA dataset. Preliminary results show that well-defined relationships exist between MFD parameters and the energy consumption of electric vehicles (electric cars or buses) at a network level. The total energy consumption of electric cars and buses in the network follows a linear relationship with accumulation under uncongested traffic conditions. When the accumulation exceeds a certain limit (e.g., 1600 [veh] for car traffic in the pNEUMA dataset), the additional effects of congestion make the relationship between accumulation and energy consumption non-linear. We also show that the energy consumption of electric cars decreases as the average speed increases when the average speed is lower than 10 km/h and then tends to be relatively constant even though the average speed further increases. For electric buses, their energy consumption exhibits an obvious decreasing trend with the increase in average speed. These findings provide valuable insights into understanding the network-wide energy consumption characteristics of electric vehicles. In the next phase of our research, we propose to comprehensively analyze the energy consumption distribution of EVs and traditional fossil fuel-powered vehicles across the entire network. On this basis, we could accurately identify the energy consumption hotspots in the network and track the origin of these hotspots. Furthermore, we will also leverage a simulation-based approach to generalize our empirical results and explore viable strategies for mitigating energy consumption hotspots in networks with both electric and fossil fuel-powered vehicles.

ACKNOWLEDGEMENTS

Sijia Sun and Yuanqing Wang acknowledge support from the National Natural Science Foundation of China (NO. 51178055), the Fundamental Research Funds for the Central Universities of the Ministry of Education of China (Grant No. 300102341302) and (Grant No. 300102341303), and the Natural Science Foundation of Shaanxi Province (NO.2021JQ-293). Hossam M. Abdelghafar, Sérgio F. A. Batista, and Mónica Menéndez acknowledge support by the NYUAD Center for Interacting Urban Networks (CITIES) and by the NYUAD Arabian Center for Climate and Environmental Sciences (ACCESS), funded by Tamkeen under the NYUAD Research Institute Awards CG001 and CG009, respectively. We acknowledge the data source is pNEUMA, <https://open-traffic.epfl.ch/>.

REFERENCES

- Ahn, K., Bichiou, Y., Farag, M., & Rakha, H. A. (2021). Multi-objective eco-routing model development and evaluation for battery electric vehicles. *Transportation Research Record*, 2675(12), 867-879. doi: 10.1177/03611981211031529
- Ahn, K., Park, S., & Rakha, H. A. (2020). Impact of intersection control on battery electric vehicle energy consumption. *Energies*, 13(12). doi: 10.3390/en13123190
- Barpounakis, E., & Geroliminis, N. (2020). On the new era of urban traffic monitoring with massive drone data: The pneuma large-scale field experiment. *Transportation Research Part C: Emerging Technologies*, 111, 50-71. doi: <https://doi.org/10.1016/j.trc.2019.11.023>
- Barpounakis, E., Montesinos-Ferrer, M., Gonzales, E. J., & Geroliminis, N. (2021). Empirical investigation of the emission-macroscopic fundamental diagram. *Transportation Research Part D: Transport and Environment*, 101, 103090. doi: <https://doi.org/10.1016/j.trd.2021.103090>
- Basso, R., Kulcsár, B., Egardt, B., Lindroth, P., & Sanchez-Diaz, I. (2019). Energy consumption estimation integrated into the electric vehicle routing problem. *Transportation Research Part D: Transport and Environment*, 69, 141-167. doi: 10.1016/j.trd.2019.01.006

- Bingham, C. (2012, March). Impact of driving characteristics on electric vehicle energy consumption and range. *IET Intelligent Transport Systems*, 6, 29-35(6).
- Donkers, A., Yang, D., & Viktorović, M. (2020). Influence of driving style, infrastructure, weather and traffic on electric vehicle performance. *Transportation Research Part D: Transport and Environment*, 88, 102569. doi: <https://doi.org/10.1016/j.trd.2020.102569>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., ... Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2). doi: <https://doi.org/10.1029/2005RG000183>
- Fiori, C., Ahn, K., & Rakha, H. A. (2016). Power-based electric vehicle energy consumption model: Model development and validation. *Applied Energy*, 168, 257-268. doi: <https://doi.org/10.1016/j.apenergy.2016.01.097>
- Fiori, C., Ahn, K., & Rakha, H. A. (2018). Optimum routing of battery electric vehicles: Insights using empirical data and microsimulation. *Transportation Research Part D: Transport and Environment*, 64, 262-272. (The contribution of electric vehicles to environmental challenges in transport. WCTRS conference in summer) doi: <https://doi.org/10.1016/j.trd.2017.08.007>
- Gao, Y., & Newman, P. (2018). Beijing's peak car transition: Hope for emerging cities in the 1.5 °c agenda. *Urban Planning*, 3(2), 82-93. doi: [10.17645/up.v3i2.1246](https://doi.org/10.17645/up.v3i2.1246)
- Geroliminis, N., & Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9), 759-770. doi: <https://doi.org/10.1016/j.trb.2008.02.002>
- Hamm, L. S., Loder, A., Tilg, G., Menendez, M., & Bogenberger, K. (2022). Network inefficiency: Empirical findings for six european cities. *Transportation Research Record*, 2676(8), 99-111. doi: [10.1177/03611981221082588](https://doi.org/10.1177/03611981221082588)
- Hiermann, G., Hartl, R. F., Puchinger, J., & Vidal, T. (2019). Routing a mix of conventional, plug-in hybrid, and electric vehicles. *European Journal of Operational Research*, 272(1), 235-248. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0377221718305605> doi: <https://doi.org/10.1016/j.ejor.2018.06.025>
- IEA. (2022). *Tracking transport 2020*.
- IPCC. (2022). *Climate change 2022: Mitigation of climate change*.
- Liu, K., Yamamoto, T., & Morikawa, T. (2017). Impact of road gradient on energy consumption of electric vehicles. *Transportation Research Part D: Transport and Environment*, 54, 74-81. doi: <https://doi.org/10.1016/j.trd.2017.05.005>
- Ma, X., Miao, R., Wu, X., & Liu, X. (2021). Examining influential factors on the energy consumption of electric and diesel buses: A data-driven analysis of large-scale public transit network in beijing. *Energy*, 216, 119196. doi: <https://doi.org/10.1016/j.energy.2020.119196>
- Othman, B., De Nunzio, G., Di Domenico, D., & de Wit, C. C. (2019). Ecological traffic management: A review of the modeling and control strategies for improving environmental sustainability of road transportation. *Annual Reviews in Control*, 48, 292-311. doi: <https://doi.org/10.1016/j.arcontrol.2019.09.003>
- Rakha, H. A., Ahn, K., & Moran, K. (2012). Integration framework for modeling eco-routing strategies: Logic and preliminary results. *International Journal of Transportation Science and Technology*, 1(3), 259-274. doi: <https://doi.org/10.1260/2046-0430.1.3.259>
- Saedi, R., Verma, R., Zockaie, A., Ghamami, M., & Gates, T. J. (2020). Comparison of support vector and non-linear regression models for estimating large-scale vehicular emissions, incorporating network-wide fundamental diagram for heterogeneous vehicles. *Transportation Research Record*, 2674(5), 70-84. doi: [10.1177/0361198120914304](https://doi.org/10.1177/0361198120914304)
- Shabihkhani, R., & Gonzales, E. J. (2014). Macroscopic relationship between networkwide traffic emissions and fundamental properties of the network. *Transportation Research Circular E-C197, Symposium Celebrating 50 Years of Traffic Flow Theory*.

- Xie, Y., Li, Y., Zhao, Z., Dong, H., Wang, S., Liu, J., . . . Duan, X. (2020). Microsimulation of electric vehicle energy consumption and driving range. *Applied Energy*, *267*, 115081. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0306261920305936> doi: <https://doi.org/10.1016/j.apenergy.2020.115081>
- Zhang, Pan, S.-Y., Li, H., Cai, J., Olabi, A. G., Anthony, E. J., & Manovic, V. (2020). Recent advances in carbon dioxide utilization. *Renewable and Sustainable Energy Reviews*, *125*, 109799. doi: <https://doi.org/10.1016/j.rser.2020.109799>
- Zhang, Y., Fu, R., Guo, Y., & Yuan, W. (2022). Environmental screening model of driving behavior for an electric bus entering and leaving stops. *Transportation Research Part D: Transport and Environment*, *112*, 103464. doi: <https://doi.org/10.1016/j.trd.2022.103464>

Application of a Metamodel-Based Optimization Approach for Toll Optimization and its comparison with Metaheuristics-based Model Optimization via a Case Study.

Gaurav Malik*¹, Chris M.J. Tampère²

¹ FWO PhD SB Research Fellow, Department of Mechanical Engineering, KU Leuven, Belgium

² Professor, Department of Mechanical Engineering, KU Leuven, Belgium

SHORT SUMMARY

First, a concept of metamodel-based optimization, in which a transport economics inspired model acts as a metamodel over an underlying set-up of directly interfaced transport models, is discussed. Then, a toll optimization scenario including a city and its neighboring rural municipalities is developed and a case study concerning its cooperative version is presented. The metamodel for this case study involves the player(s) optimizing their objective based on a schematic network, and simplified cost and demand functions, whereas the underlying set-up is a Static Traffic Assignment over the physical network with physical origin-destination elastic demand. This new metamodel-based optimization is then compared with traditional metaheuristics-based optimization. Results show that the new approach not only leads to lower computational expense but even outperforms metaheuristics-based optimization in terms of optimality.

Keywords: Game theoretical interactions in mobility, Pricing and capacity optimization, Transport economics and policy, Metamodel-based optimization, Transportation network modelling

1. INTRODUCTION

Road pricing is an important topic for many transportation stakeholders. Comprehensive analysis of tolling schemes should consider interactions between transportation subsystems/stakeholders (travelers, mobility service providers, and local network operators) and other systems (neighboring governments' network operators, housing market, urban design/land-use).

Thus, for a particular tolling scheme, the challenge for modelers is to 1) identify relevant (sub)systems/stakeholders and their interactions. 2) develop models that take into account these interactions in a more elaborate way than fixed inputs or unidirectional influences. 3) develop mechanisms for computation of consistent impacts on all interrelated (sub)systems/stakeholders. There have been attempts to approach solutions based on three major approaches:

1. **All-encompassing micro-models:** These are highly detailed and disaggregate models in which every relevant player is modelled at a micro level e.g., the Multi-Agent Transport Simulation (MATSim) framework (Horni, Nagel, & Axhausen, 2016). Such frameworks can provide information at a very disaggregate level to the stakeholders allowing analysis from equity as well as efficiency perspective. However, developing, extending,

calibrating, maintaining, and interpreting such models require substantial effort, and game-theoretical analyses involving multiple stakeholders with different objectives can be prohibitively expensive.

2. **Tailor-made simplified conceptual models:** On the other end of the spectrum lie simplified conceptual models which are often used by transport economists (B. De Borger, Dunkerley, & Proost, 2007; Bruno De Borger & Proost, 2021). These are comparatively easier to develop, calibrate, maintain, and interpret. It is also easy to tailor them and focus only on the relevant stakeholders/(sub)systems and their interactions for a particular problem. However, they involve extensive simplifications of the underlying (sub)systems; thus, they only provide highly aggregate and schematic results which is, usually, not sufficient for aiding actual decisions.
3. **Tailor-made directly interfaced traditional mono-disciplinary models:** Traditionally, a toll optimization problem is solved by using a bi-level optimization framework in which toll is altered in an outer loop around the Static Traffic Assignment (STA) (Ekström, Rydergren, & Sumalee, 2014). Nowadays, to account for interactions with other (sub)systems, dedicated interfaces between relevant mono-disciplinary transportation models are being built within the inner loop e.g., the connection between activity-based demand model and STA in Strategisch Personen Model Vlaanderen (Vanderhoydonc & Borremans, 2020). This approach may provide flexibility and produce disaggregate results. However, solving complex optimization problems involving several stakeholders can still be extremely sluggish because: a) As the number of models increase, attaining consistency via bi-directional interfaces becomes computationally quite expensive. b) Due to the possibility of only marginal steps in the optimization variables, the risk of getting stuck in local stationary points is quite high.

It can be appreciated that none of these three approaches offers the combination of scalability, detail, and flexibility required by contemporary transportations problems.

Inspired from the complementary characteristics of the 2nd and 3rd approaches mentioned above, **we aim to use a transport economics inspired conceptual model as a metamodel to find an optimal toll for an underlying set-up of directly interfaced traditional transport models.** In this way, the underlying set-up only has the computational load related to achieving consistency between the directly interfaced models, whereas the computational load for optimization lies completely at the metamodel level. At every iteration, the underlying set-up is used to (re-)calibrate the metamodel which includes a simpler and more aggregate version of all the relevant stakeholders/(sub)systems. Toll optimization is performed for the meta-model and optimal tolls are transmitted to the underlying set-up. At the new tolls, the underlying set-up is evaluated again, and the meta-model is recalibrated at the new point. This sequence is repeated until a certain level of convergence is achieved in the optimal toll values.

The objective of this paper is to present the development and results of a **proof of this concept** and thereby, determine 1) the optimality and 2) the computation speed of this framework.

2. METHODOLOGY

The fictional problem considered for the proof of concept is as follows: the city municipality is looking to impose two non-discriminatory cordon tolls i.e., an entry toll each for radial and ring roads with the intention of curbing the use of city infrastructure by transit traffic. To avoid the re-routing of transit traffic to their infrastructure, rural municipalities come together to charge an entry toll for the neighboring rural territory.

The framework has three main parts: 1) The underlying Set-up 2) the metamodel 3) the calibration interface between underlying set-up and the metamodel. **Figure 1** shows a block diagram representing a basic instance of this framework.

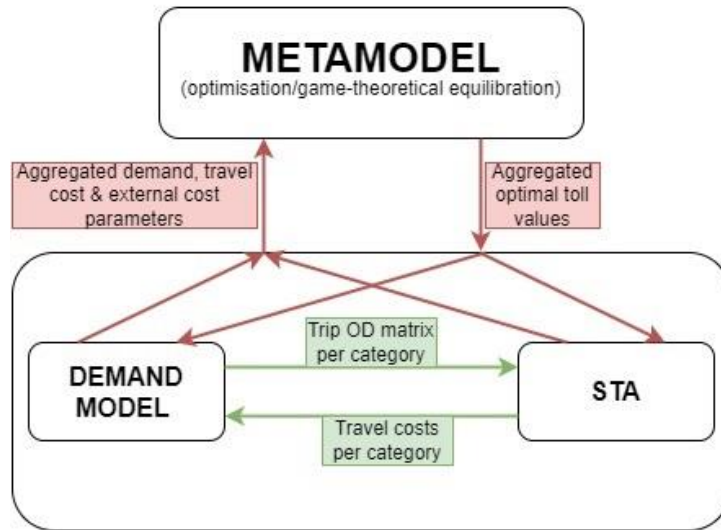


Figure 1: Basic instance of framework

Underlying Set-up

The underlying set-up, for this exercise, consists of an STA for a fictional city and a simple demand model with linear elasticities per OD.

Static Traffic Assignment (STA):

Details of STA are mentioned in **Table 1**.

Figure 2 and **Figure 3** show the network and the zoning respectively.

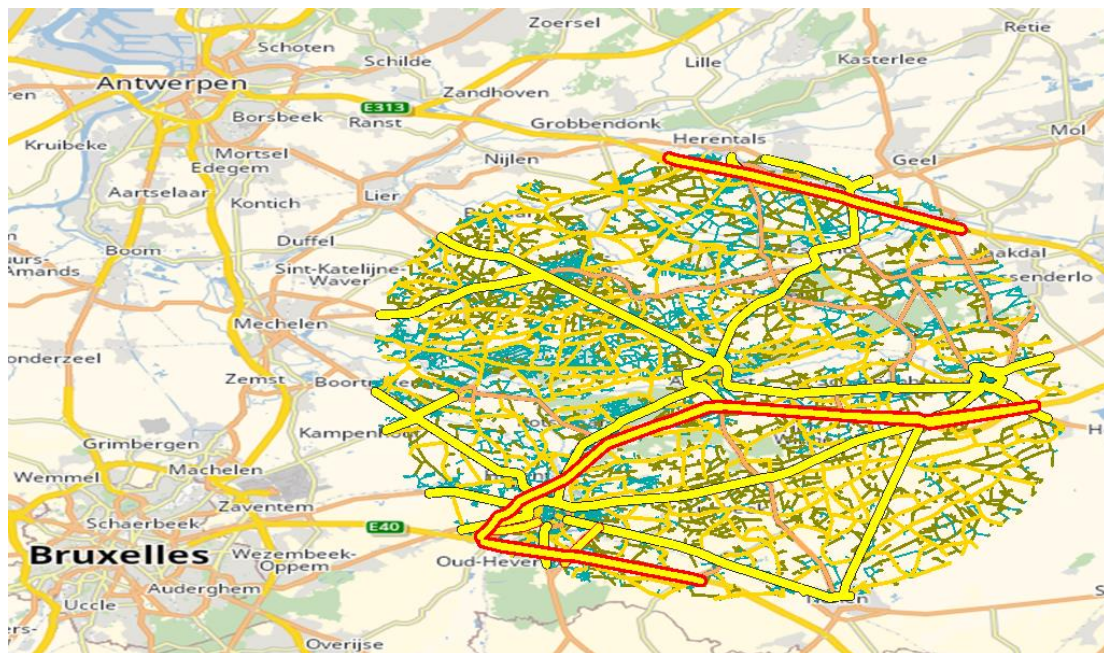


Figure 2: Network shown against the background of OpenStreetMap

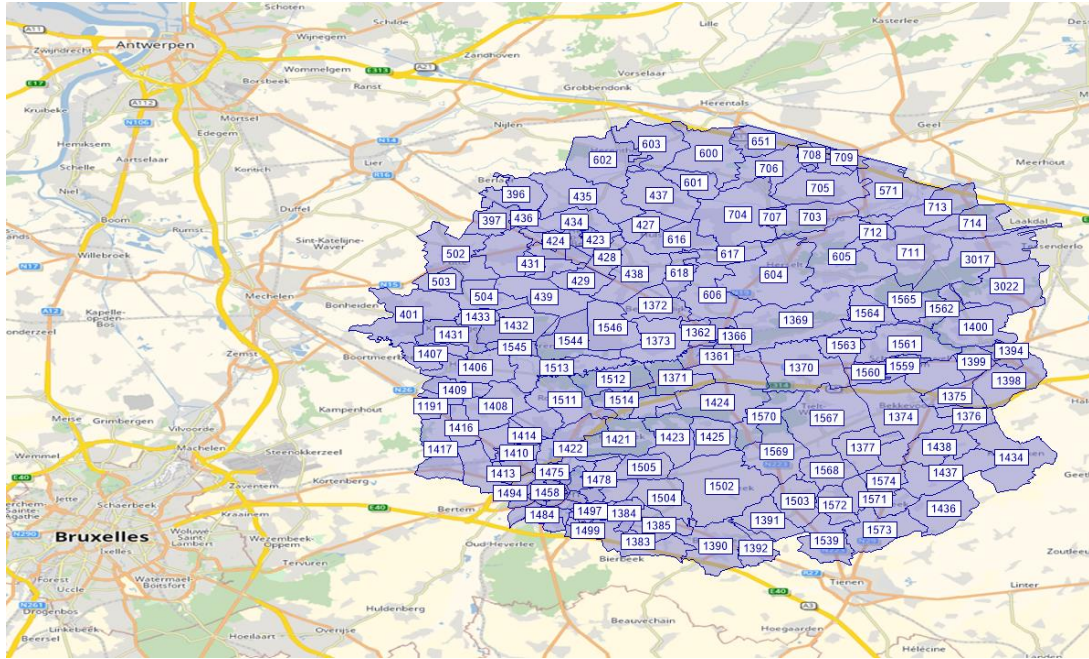


Figure 3: 173 zones used for the study

Table 1: Static Traffic Assignment details [1: (Boeing, 2017), 2: (Vanderhoydonc & Borremans, 2020), 3: (Gentile, 2014)]

Network		
	Centre	Aarschot
	Buffer	17 kms
	No. of nodes	43440
	No. of links	143368
	Source	OpenStreetMap
	Comment	using OSMnx (1)
Demand Data		
	Centre	Aarschot
	Buffer	17 kms
	No. of zones	173
	OD Matrix size	173 X 173
	Source	Belgium-wide data by Flemish Road Authority (2)
	Comment	Additional in-house mining on top of source
Assignment		
	Software	PTV Visum
	Method	LUCE (3)

To include transit traffic through Aarschot, a proportion of the external demand of interest is projected onto the periphery zones. The city and the neighboring municipalities are assumed to have jurisdiction over Territory 6 and Territory 5 respectively (**Figure 4**). The three tolls are added in units of time to travel costs (BPR) of the appropriate **entry** links.

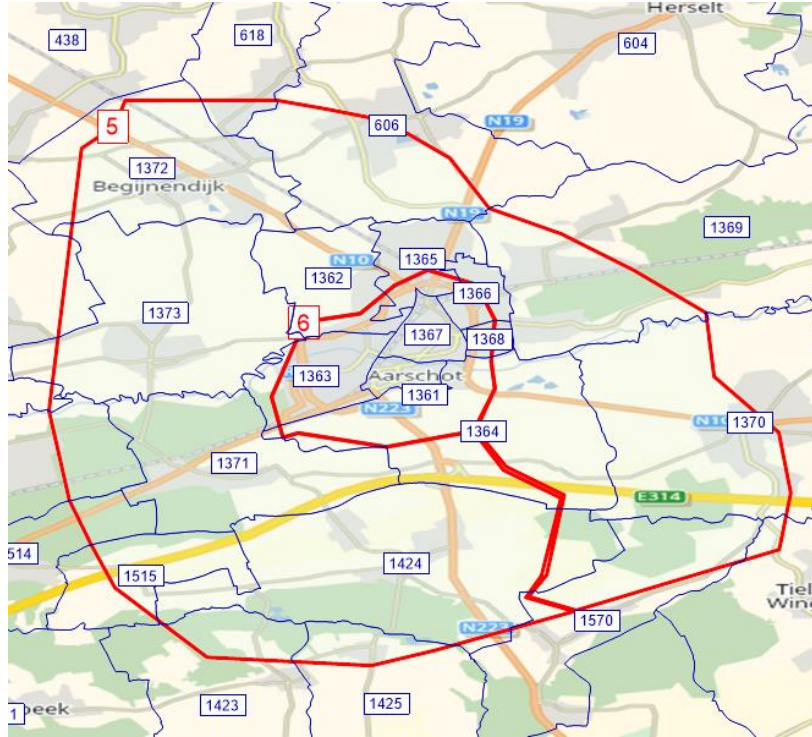


Figure 4: Territory 6 is under city jurisdiction and Territory 5 is under jurisdiction of rural municipalities.

Demand Model:

The demand model is a standard linear inverse demand function of the type mentioned in **Equation 1** and it is used to get the demand (D_{od}) for each of the $173 \times 173 = 29929$ OD pairs as a function of their cost skims (C_{od}).

$$A_{od} - B_{od} * D_{od} = C_{od} \quad (1)$$

For this study, A_{od} and B_{od} have been derived by using the reference OD matrix and reference Cost Skim matrix (obtained after assigning reference OD matrix) and a realistic assumption about the maximum possible demand (at zero cost) for each OD pair.

The horizontal interface between the STA and the demand model shown in **Figure 1** is solved as a fixed-point problem with Method of Successive Averages (MSA) smoothening. In future, more advanced demand models, such as activity-based demand models, may replace this basic demand model.

Metamodel

The metamodel is inspired from transport economics models (Bruno De Borger & Proost, 2021). The schematic network chosen for this case study is shown in **Figure 5** and other details are mentioned in **Table 2**.

Table 2: Details of metamodel network

No. of nodes	9
No. of links	13
Centroid nodes	1, 4, 7, 8 and 9
Left to Right links	All except 9 and 12
Right to left links	9 and 12
Radial links	5 and 6
Ring links	7, 4 and 8
Rural links	3, 9, 10, 11 and 12
Radial toll	link 5
Ring toll	link 7
Rural toll	links 10 and 12

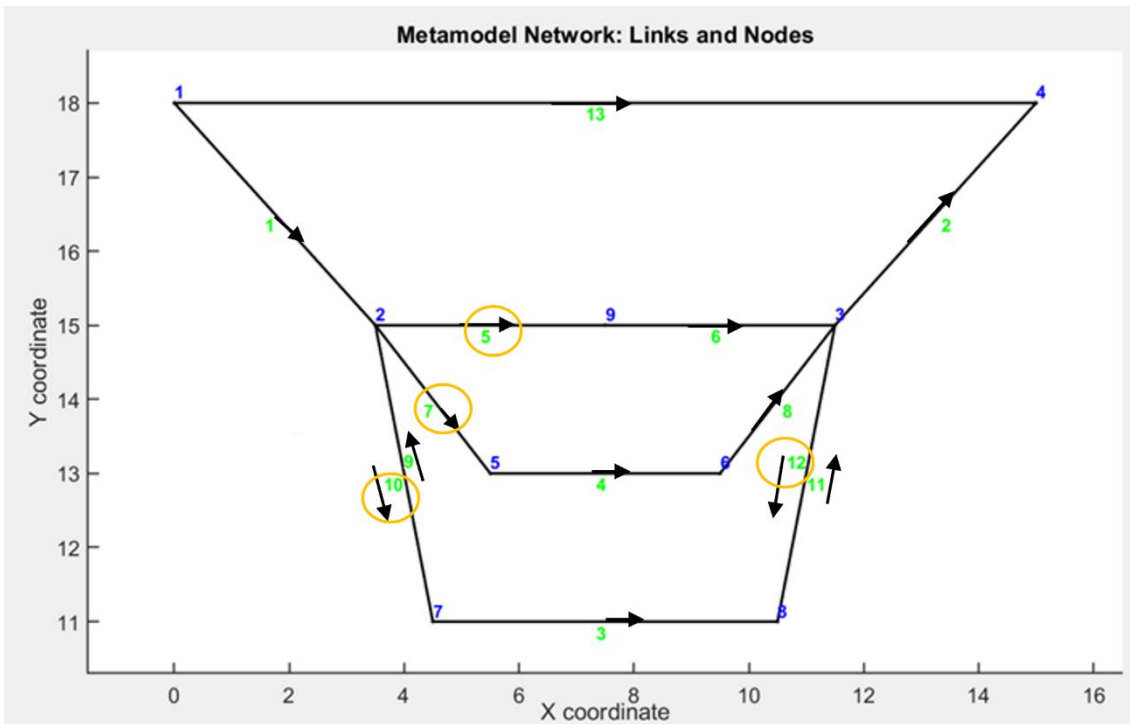


Figure 5: Network for metamodel: Links (green labels), Nodes (blue labels) and Tolled Links (yellow circles)

Links are assumed to have a linear congestion cost function (lc_i) of the type:

$$lc_i = a_i + b_i * f_i \quad (2)$$

Demand is assumed to be comprised of ten schematic OD pairs between the five centroids (see **Table 3**). There are nineteen paths for the ten OD pairs. The paths using radial links, ring links and rural links are mentioned as ‘r’, ‘R’, and ‘M’ respectively. For OD pair 1_4, the path that escapes all three tolls is mentioned as ‘e’ in **Table 3**. The demand in metamodel is also assumed

to be elastic with linear inverse demand functions of the type mentioned in **Equation 1** with A_{od} and B_{od} (for each of the 10 OD pairs) being calibration parameters for the metamodel.

Table 3: Metamodel OD pairs

S.No.	OD pair (from_to)	Available paths	Explanation (from_to)	Code name
1	1_4	M,R,r,e	external_external	OD
2	1_7	M	external_closerural	ON
3	1_8	M,R,r	external_opprural	ONopp
4	1_9	r	external_city	OA
5	7_4	M,R,r	opprural_external	ND
6	8_4	M	closerural_external	NoppD
7	9_4	r	city_external	AD
8	7_8	M,R,r	rural_rural (opp.)	MoMd
9	7_9	r	rural_city	MoA
10	9_8	r	city_rural	AMd

Since the scope of the paper is only to provide a proof of concept, the metamodel is used only for solving joint optimization problem of the city and rural municipalities as opposed to solving Nash-Cournot or Stackelberg competitions. The objective function for this joint optimization (minimization) is a quadratic function of the optimization variables i.e., three tolls (T_{Tli}) and nineteen path flows (X_{pi}). It is a sum of (negative) user welfare, total costs including tolls, external costs and (negative) total revenue from tolls. Objective function and constraints are shown in **Equations 3-11**.

$$Obj = -Wlf + TC + EC - TR \quad (3)$$

where:

$$Wlf = \sum_{od=1}^{10} (A_{od} * D_{od} - 0.5 * B_{od} * D_{od}^2) \quad (4)$$

$$TC = \sum_{pi=1}^{19} (C_{pi} * X_{pi}) \quad (5)$$

$$EC = \sum_{li=1}^{13} (\lambda_{li} c_{li} * f_{li}) \quad (6)$$

$$TR = \sum_{Tli=1}^4 (T_{Tli} * f_{Tli}) \quad (7)$$

subject to:

$$all X_{pi} \geq 0, all T_{Tli} \geq 0 \quad (8)$$

$$Wardrop's equilibrium (route - choice) \quad (9)$$

$$Elastic demand equilibrium \quad (10)$$

$$all T_{Tli} \leq T_{Max} \quad (11)$$

Wardrop's equilibrium condition makes the problem highly non-convex because of which special checks are required for ensuring global optimality. Details on the solution methods for the metamodel are the subject of a forthcoming paper (Malik & Tampère, n.d.).

Interface

Calibration Interface:

At the beginning of each new metamodel optimization routine, the metamodel is (re-)calibrated by using the underlying set-up. This is done by the calibration interface (represented by red lines going up in **Figure 1**). Specifically, it calibrates: 1) elastic demand parameters for the metamodel

i.e., the ten A_{od} and B_{od} mentioned in the previous section, 2) cost parameters of the thirteen metamodel links i.e., a_i and b_i (**Equation 2**) and 3) thirteen external cost parameters i.e., λ_{li} (**Equation 6**).

The zeroth step in this calibration is classifying each of the 29929 physical OD pairs as belonging to either one of the ten metamodel OD pairs. This is performed as a preprocessing step using geometrical logic. **Figure 6** provides an example of the classification for the categories of 1_4, 1_7, 1_8 and 1_9. Red area represents the city territory and pink + violet areas represent rural territory. For an external origin **1**, an OD pair is classified based on the area in which the destination lies i.e., 7, 8, 9 or 4. A similar process is followed for the six remaining OD pair categories.

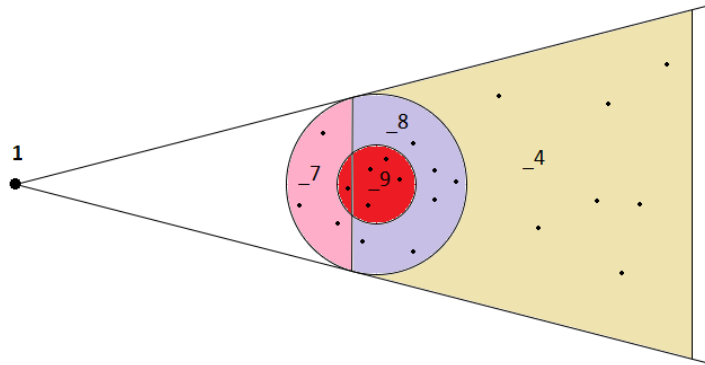


Figure 6: Geometric logic for classification of physical OD pairs as 1_4, 1_7, 1_8 and 1_9 in metamodel.

Then, for each meta-OD category, demands of all belonging physical OD pairs are summed together to give category demand (=metamodel demand for that OD pair) and category cost is given by the average of belonging physical OD cost skims weighted by their corresponding maximum possible demand levels. We do this at two closest elastic demand equilibrium points of the underlying set-up and use them to find locally linearized aggregated inverse demand parameters i.e., the ten A_{od} and B_{od} for the metamodel.

The next step is to calibrate link cost parameters of **Equation 2**. We do this by estimating two flow-cost points. The flows of each of the metamodel links are estimated by using their physical interpretation e.g., flow on link 1 represents all traffic that's entering the two territories shown in **Figure 4**. Then, the corresponding metamodel link costs can be estimated by using the total vehicle hours spent on the real physical network on links associated with a particular metamodel link. By equating the total vehicle hours on the real and metamodel links, we find the metamodel link costs after already having estimated metamodel link flows. Repeating this for a marginally higher demand level gives us another metamodel link flow-cost point. This allows us to find a_i and b_i i.e., locally linearized meta-model link cost parameters.

For calibrating the external cost parameters λ_{li} for **Equation 6**, we need to assume certain external costs in the underlying set-up first. We used a parameter Λ each for the rural, radial and ring physical links. Then, the obvious choice of λ_{li} for: 1) links 5 and 6 is Λ_{radial} , 2) links 7, 4 and 8 is Λ_{Ring} , and 3) links 9, 10, 11 and 12 is Λ_{rural} . λ_{li} for remaining metamodel links is zero.

Toll Interface:

This interface is represented by red lines coming down in **Figure 1**. After the calibration interface has calibrated the metamodel, the metamodel computes optimal tolls for that version of the metamodel. The Toll interface applies an MSA smoothening step on these optimal tolls using the

optimal tolls of the last iteration. The smoothed tolls are then implemented in the underlying set-up as tolls for the next iteration.

Metaheuristics-based optimization for benchmarking

We aim to evaluate optimality and speed of this metamodel-based optimization. This necessitates the benchmarking of this framework against a state-of-the-art metaheuristics-based optimization framework. Traditionally, metaheuristics like Simulated Annealing (SA) and Genetic Algorithms (GA) are used for optimization of such a large-scale network. However, we use SHERPA (“SHERPA,” n.d.) as it has been shown to significantly outperform SA and GA both in terms of efficiency and robustness.

As compared to **Figure 1**, the same underlying set-up is used in this case as well. However, the interface i.e., the calibration, and the toll interface and the metamodel are replaced by a toll optimization outer loop. As mentioned in the introduction, computational loads for both achieving horizontal consistency as well as finding optimal tolls are combined in this case. The objective function for optimization is formed completely analogous to **Equations 3 -7**; however, in this case, it is formulated using 29929 OD pairs and 143,368 links.

3. RESULTS AND DISCUSSION

We obtained results for the case when both the city and the rural municipalities cooperatively optimize the three tolls. Bounds of $0 \leq T \leq 4$ were used for the tolls. Λ_{radial} , Λ_{Ring} and Λ_{rural} were set to 40, 0 and 10 respectively. For SHERPA based optimization, the total number of evaluations was set to 25. The value of objective function for successive designs is shown in Figure 7 and Figure 8. The design with Design Id = 23 proved to be the best design with an objective value of 128.877 vehicle-hours and the corresponding best tolls were $T = [3.92, 0.64, 1.4]$ hours.

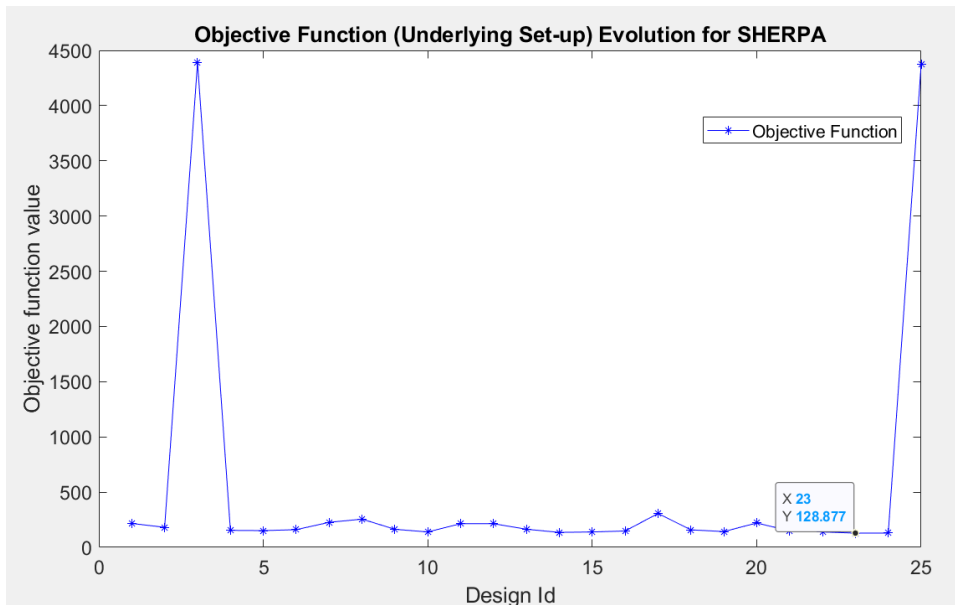


Figure 7: Objective Function (veh.-hr.) for successive designs in SHERPA-based Optimization (Data Tip shows the Best Design)

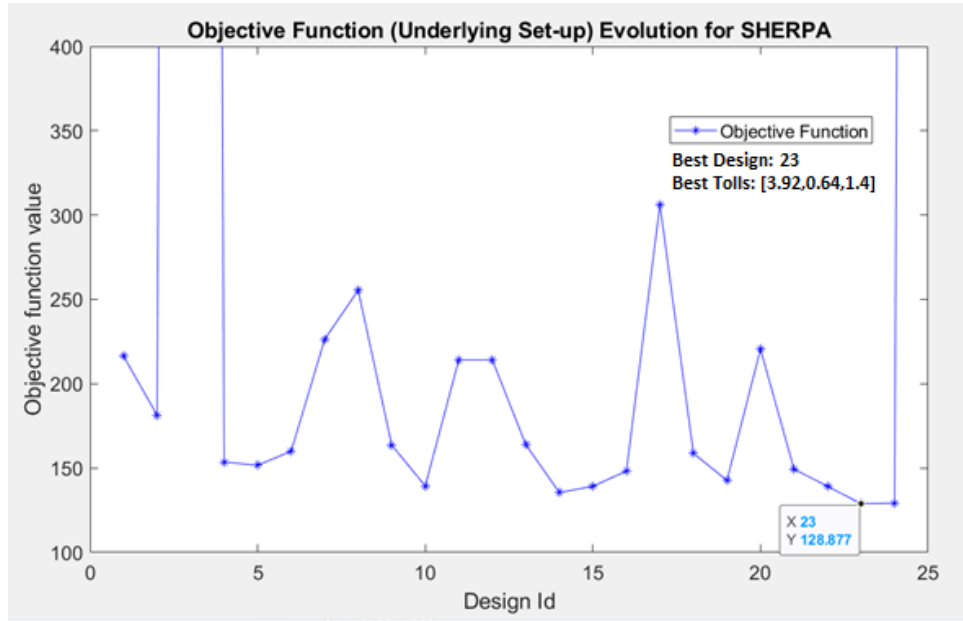


Figure 8: Objective Function (veh.-hr.) for successive designs in SHERPA-based Optimization (ignoring outliers)

For metamodel-based optimization, the number of evaluations was set to 10. The evolution of objective function and tolls values over successive iterations is shown in **Figure 9** and **Figure 10** respectively. The iterations start with an educated guess of T_0 motivated from the values of $[\Lambda_{radial}, \Lambda_{Ring}, \Lambda_{rural}]$. This initial guess proved to be a highly favorable point as the objective value is extremely low; however, the tolls (and consequently the objective function) move quite aggressively away from this point in next MSA-iteration. Regardless, it is remarkable that in the subsequent iterations, the model, almost monotonously, manages to find its way back to values significantly lower than those suggested by the best design of SHERPA-based optimization. It is even more encouraging that it did that within the first 6 iterations while not directly optimizing the actual objective function on which optimality is evaluated but instead an aggregated and linearized version of it.

Results of the two approaches are summarized in **Table 4**. It should be noted that for additional players e.g., in Nash Cournot/ Stackelberg scenarios, computational time for SHERPA/metaheuristics-based approach will increase exponentially but for metamodel-based approach, it will stay practically the same.

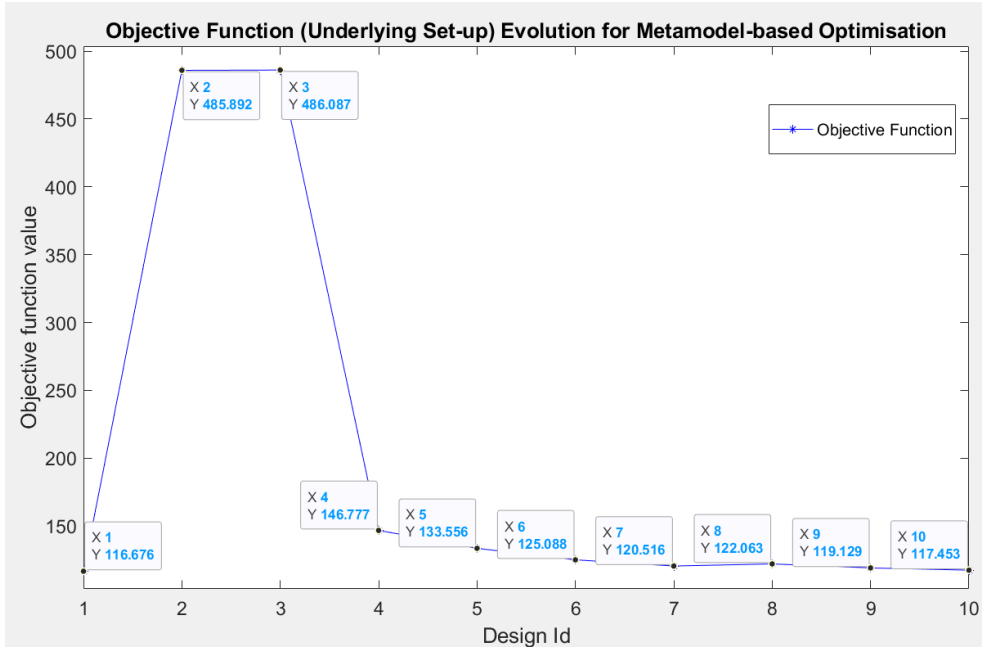


Figure 9: Evolution of Objective Function (veh.-hr.) for Metamodel-based Optimization

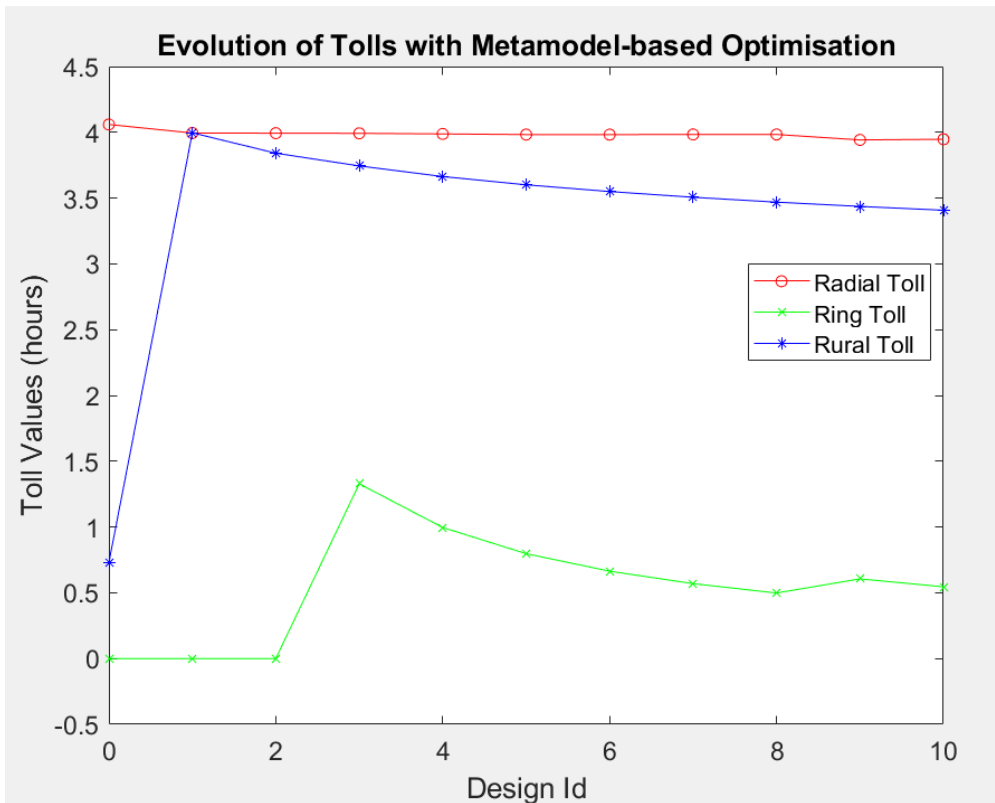


Figure 10: Evolution of Tolls with Metamodel-based Optimization

Table 4: Comparison of two approaches

Optimisation type	SHERPA-based	Metamodel-based
No. of iterations	25	10
Computation time per iteration	13-25 mins	15-26 mins
Best Tolls (hr.)	[3.92, 0.64, 1.4]	[3.94, 0.61, 3.43]
Objective function value of best design	128.77 veh. hr.	117.453 veh. hr
Total computation time (both run parallelly)	8h 6m 5s	3h 39m 58s

4. CONCLUSIONS

This paper discussed a new metamodel-based optimization approach in which a transport economics inspired metamodel is used for optimizing the underlying directly interfaced traditional transport models. A case study about a joint toll optimization problem of a fictional city municipality and its neighboring rural municipalities is developed and presented as a proof of concept of this approach. Preliminary results suggests that the problem can be solved at a much lower computational cost and with appreciable accuracy in terms of optimality of tolls. This study serves as a motivation for interfacing additional models in the underlying set-up as well as solving Nash Cournot and Stackelberg competition scenarios in the metamodel where traditional metaheuristics-based optimization can be prohibitively expensive.

ACKNOWLEDGEMENTS

This research is funded by the FWO (Project Number: 1S44922N).

REFERENCES

- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>
- De Borger, B., Dunkerley, F., & Proost, S. (2007). Strategic investment and pricing decisions in a congested transport corridor. *Journal of Urban Economics*, 62(2), 294–316. <https://doi.org/10.1016/j.jue.2007.01.001>
- De Borger, Bruno, & Proost, S. (2021). Road tolls, diverted traffic and local traffic calming measures: Who should be in charge? *Transportation Research Part B: Methodological*, 147(C), 92–115.
- Ekström, J., Rydergren, C., & Sumalee, A. (2014). Solving a mixed integer linear program approximation of the toll design problem using constraint generation within a branch-and-cut algorithm. *Transportmetrica A: Transport Science*, 10(9), 791–819. <https://doi.org/10.1080/23249935.2013.813988>

- Gentile, G. (2014). Local User Cost Equilibrium: A bush-based algorithm for traffic assignment. *Transportmetrica A: Transport Science*, 10(1), 15–54. <https://doi.org/10.1080/18128602.2012.691911>
- Horni, A., Nagel, K., & Axhausen, K. W. (2016). The Multi-Agent Transport Simulation MATSim. In *Ubiquity Press*. Ubiquity Press. <https://doi.org/10.5334/baw>
- Malik, G., & Tampère, C. M. J. (n.d.). Transport Economics Models for Problems with Several Underlying Routing Options: Solving for Global Optimality and Applications to Strategic Competition Between Different Levels of Governments. *Paper in Preparation*.
- SHERPA. (n.d.). Retrieved August 2, 2022, from <https://www.redcedartech.com/solutions/heads-software/efficient-search/9-solutions/214-sherpa>
- Vanderhoydonc, Y., & Borremans, D. (2020). *Strategische Verkeersmodellen Vlaanderen versie 4.2.1: Overzichtsrapportage* (No. 14258). Vlaamse Overheid, Departement MOW afdeling Beleid.

Assessing the Long-term Impact of E-bikes on Sustainable Mobility: A National-Level Study in the Netherlands

Bingyuan (Amelia) Huang^{*1}, Hans Wust¹, Mathijs de Haas¹

¹ KiM Netherlands Institute for Transport Policy Analysis, Bezuidenhoutseweg 20, 2594 AV, The Hague, The Netherlands

SHORT SUMMARY

Over the past decade, e-bikes have become increasingly popular, sparking interest in their potential replacement for car use and benefit for the environment. However, studies on e-bike substitution effects have limitations, including a lack of assessments of the effects on mobility on the national level, a narrow focus on commuting travelling, and insufficient consideration of future expected e-bike use. This study proposes a new approach that combines an intention-based method with time-series forecasting to estimate e-bike use and investigate its potential for sustainable mobility in the Netherlands. The results show that e-bike ownership strongly reduces the conventional bicycle use and, to a lesser extent, car and public transport use, especially for commuting travelling. This study provides insight into how e-bikes substitute for car use and other modes of transportation, and how the expected growth in e-bike use in coming years may impact national mobility in the Netherlands.

Keywords: Matching method, Unified-Richards growth curve, Transport policy, Cycling behaviour, Substitution, Dutch national mobility survey

1. INTRODUCTION

Pedal-assisted-bikes, also known as e-bikes or electric bikes, are bicycles equipped with a battery-powered motor that assists with pedalling, providing support up to a maximum speed or power. E-bikes have increased in popularity over the past decade. In Europe, e-bike growth per year has an average of 30% between 2010 and 2016. Germany and The Netherlands accounted for over 50% of e-bike sales in the EU in 2016 (CONEBI 2017). In the Netherlands, since 2018, more new e-bikes have been sold each year than conventional bikes (BOVAG, 2023). Moreover, more younger aged people are adopting the e-bike which was originally popular among people over the age of 65 in the Netherlands. As the e-bike allows travelling at greater speeds with less effort compared to a conventional bicycle, it has the potential to replace a substantial part of car trips and bring health benefits. It is thus worth investigating the potential contribution of e-bikes in the shift towards a more sustainable transport system.

Previous studies have shown that the sustainability effects of e-bikes are complex. It mainly depends on whether the e-bike replaces motorized modes (e.g. car trips) (Wolf and Seebauer, 2014). In reality, whether e-bike use will result in a reduction of car use depends on local context. For instance, the substitution of public transport by e-bikes happens in cities with a high-quality transit

system (Fishman and Cherry, 2016). The substitution of car trips can be observed in a car-dominated countries (Wolf and Seebauer, 2014), whereas in European countries with a bicycle orientation, the e-bike seems to substitute the conventional bicycle in addition to the car (Cherry et al., 2016; de Haas et al., 2021; Kroesen, 2017).

A limitation of previous studies is that they are not representative for mobility on the national mobility level nor focused on the future expected e-bike adoption and use. These studies typically focus on individual-level effects with selective survey samples, which may be representative of the national population, but not necessarily of national mobility. While evidence on substitution effects can be collected from such studies, assessing the effect of e-bikes on mobility on the national level is difficult. Moreover, previous studies tend to focus specifically on commuting and do not consider the use of the e-bike for other trip motives. Due to these limitations, the current knowledge on e-bike substitution effects provides an incomplete picture. This limits policymakers in making well-informed decisions on how to use e-bikes to promote sustainable travel behaviour. This study addresses these shortcomings by assessing the effects of e-bike substitution over the long term at a national mobility level in the Netherlands, a leading country in e-bike adoption (Fishman and Cherry, 2016; CONEBI 2017).

We aim to provide a new approach to tackle the question of e-bike adoption and usage at the national level in the long run, and we provide a robust validation of our findings. For this purpose, we employ a combination of an intention-based method and a time-series forecasting method to estimate e-bike use and travel behaviour in the future, providing insights into the substitution of other transport modes in the Netherlands. These insights can be used by policymakers to decide if, and how, the e-bike can be used as a means of promoting more sustainable travel behaviour.

2. METHODOLOGY

To investigate how e-bikes can replace other modes of transportation, we used a combination of an intention-based method and time-series models to estimate e-bike use in the coming years. The intention-based method estimates future e-bike ownership, e-bike use and travel behaviour based on people's intention to buy an e-bike and their intended use of the e-bike by using the Netherlands Mobility Panel (MPN). Further, we validated this estimation by a time series forecasting method based on data from the Dutch national travel survey from 2013 to 2021 (Statistics Netherlands, 2013–2021).

Intention-based method

To predict future e-bike ownership and usage, we used a two-step approach. First, we collected people's intention to purchase an e-bike within the next five years and their intended use of the e-bike using the annually conducted Netherlands Mobility Panel (MPN) from the year 2021 (KiM, 2021). Second, we estimated the future national-level e-bike use and other modes of transport by assuming that those with a buying intention will purchase e-bikes and their usage will mirror that of current e-bike owners with similar demographic profiles. This assumption is based on the finding from the first step that future e-bike owners intend to use the e-bike in a similar manner to current owners.

For the first step, the MPN, an annual household panel that represents the Dutch population, was used to gather data on future e-bike buying intentions. A total of 1046 e-bike owners and 1461

non-owners participated in the questionnaire. For the second step, the Dutch national travel survey (ODiN) was used to predict future e-bike usage and travel behaviour. The annually conducted ODiN involves approximately 40,000 individuals (0.2% of the Dutch population) and is representative of the daily mobility of the Dutch population.

ODiN provides more reliable information on yearly statistics of total Dutch mobility than MPN. But MPN is more practical to collect e-bike buying intention. The two datasets were connected through a matching process as shown in Figure 1 in order to link the future adoption intention gathered from MPN to ODiN. Since ODiN includes more respondents than MPN, the matching process involved linking each MPN respondent with buying intention to multiple ODiN respondents with the same sociodemographic profile and do not yet own an e-bike. This allowed us to identify individuals in ODiN who do not own an e-bike, but do intend to purchase one in the near future.

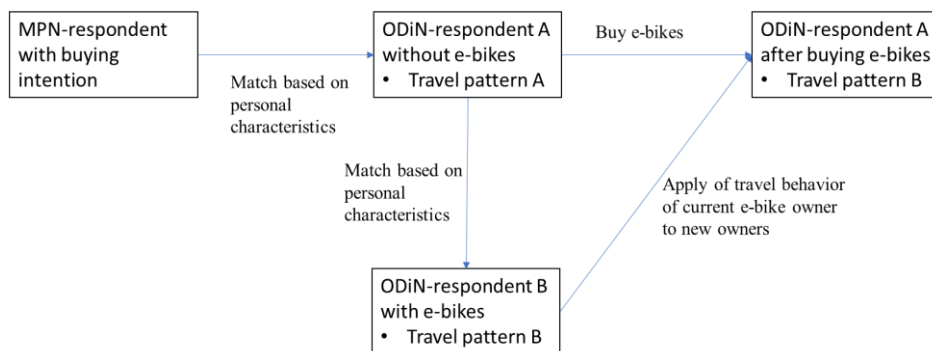


Figure 1 Schematic representation Calculation potential e-bike

To further estimate the e-bike use of future owners in ODiN and their travel behaviour on the national level, we assume that the future e-bike owners will use their e-bikes in a similar manner as current owners with similar demographic profiles. This assumption is backed by the MPN survey, that showed that future e-bike owners expect to use the e-bike in a similar manner as current owners. To do so, we replaced the travel diaries of the future e-bike owners with the travel diaries of their matched e-bike owners in ODiN. The new ODiN data is still representative of the mobility of the Dutch population in the five years following the reference year 2019. Respondents were matched based on personal characteristics available in both MPN and ODiN, such as gender, age, urbanity, education level, car ownership, and commute distance, using the Mahalanobis distance and the R-package MatchIt (Stuart et al., 2010).

The estimate of the e-bike use presented above does not take into account other relevant factors that may affect usage, such as demographic and economic developments. Therefore, we do not have a complete picture of the expected development of the e-bike use. Furthermore, this estimate is based on a number of key assumptions as described above. However, our second method can partially address this issue by validating the results using a time series model.

Time series forecasting

The goal of the time series forecasting is to model the e-bike share and extrapolate the share to 2024, so as to evaluate the e-bike usage results from the above mentioned intention-based method. This method only extrapolates the future e-bike share and provides no information about e-bike substitution and future usage of other transport modes.

The e-bike share was estimated with a multilevel time series models (MTSM). The combination of fixed and random effects of the MTSM allows the sharing of information across all group aggregates (5 travel purposes, 9 age groups and 2 genders). This results in more precise estimates as compared to modelling each group separately. Assuming that the e-bike share follows an S-shaped growth process in general, we applied the Unified-Richards growth curve formulation (1) of (Vrána et al., 2018) in the MTSM because it covers a wide range of S-shaped growth curves.

$$w = A(1 + (d - 1) * \exp(\frac{-k_U(t-T_i)}{d^{d/(1-d)}}))^{1/(d-1)} + b_1 * cvd_1 + b_2 * cvd_2 \quad (1)$$

The Unified Richards growth curve parameters are:

1. saturation level A (the upper asymptote of the share curve).
2. (relative) growth rate k_U at the inflection point of the growth curve.
3. time-location T_i of the inflection point. (t represents time)
4. form parameter d that locates the vertical location of the inflection point.

Additionally, the effect of the COVID-19 pandemic on e-bike share in 2020 and 2021 was modelled by the parameters b_1 and b_2 respectively, with corresponding dummy variables cvd_1 and cvd_2 .

Each parameter was modelled as follows. The fixed effects of the parameters A , k_U and T_i were modelled as a monotonic function of age (Bürkner & Charpentier, 2020), with gender added for parameter A . Additional fixed effects are the interaction of gender and age for parameter d and the interaction of age and purpose for parameters b_1 and b_2 . Random effects varying over all combinations of purpose, age and gender were included for parameters A , k_U , b_1 and b_2 and varying over purpose and gender for parameter d . The random effects of the parameters A and k_U were modelled as correlated.

The multilevel model was fitted to ODin data from 2013 to 2021, using the R-package brms (Bürkner 2017). Brms is an interface to the Bayesian Markov Chain Monte Carlo programming language Stan (Stan Development Team, 2022). Model checking and model comparison were done using goodness of fit values of the waic information criterion and approximate leave-one-out cross validation using the R-package loo (Vehtari et al., 2017).

3. RESULTS AND DISCUSSION

Over the next 5 years, the MPN survey results suggest that 22% of non-owners intend to adopt an e-bike (see Table 2). However, it is likely that not all of these individuals will actually end up purchasing an e-bike. To provide a realistic estimate, we assume that all individuals with an intention to buy within the next 6 months will make a purchase, while 90% of those intending to buy within the next 2 years, and 85% of those intending to buy within the next 5 years will eventually buy an e-bike by 2024.

Table 2 Intention of e-bike adoption among non-owners

e-bike adoption	Share of the non-owners	Share that actually purchases an e-bike
yes, within 6 months	2%	100%
yes, between 6 months – 2 years	8%	90%
yes, between 2-5 years	12%	85%
yes, but after 5 years	17%	-
No	61%	-

If we take into consideration the increase in e-bike ownership and the travel habits of the new owners, the distance covered by e-bike is expected to rise to 69% between 2019 and 2024, from 0.65 km per person per day to 1.1 km per person per day. The distance covered by regular bicycles will then drop by 10%. The total distance covered by e-bike will rise more than the distance covered by regular bicycle will decrease, causing the total distance covered by bicycle to increase by about 8%. This will cause the e-bike's share of the total distance covered by bicycle to increase from 23% to 35%. This e-bike share estimation is in line with our time series forecasting result, as shown in Figure 2.

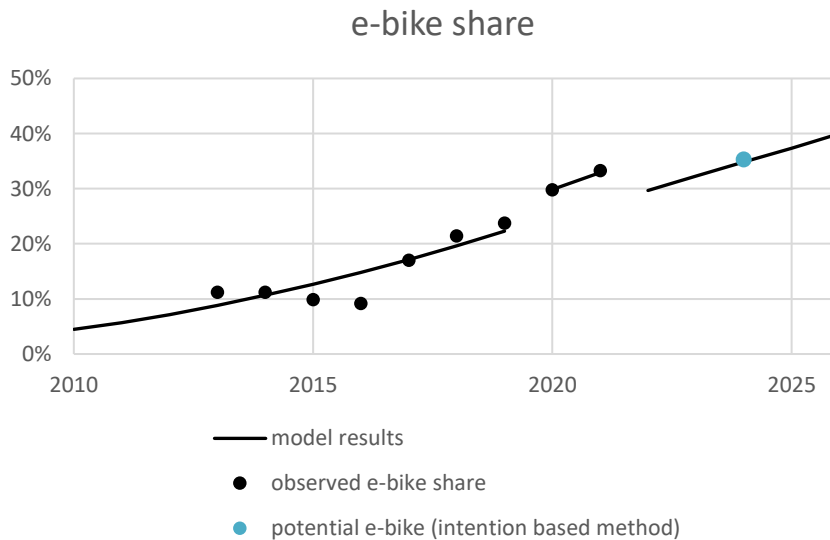


Figure 2 The forecasting of the e-bike's share of the total distance covered by bicycle based on time series forecasting

E-bikes also substitute other means of transport than the regular bicycle. Figure 3 shows the changes in the share of trips for a number of transport means classified by distance. It shows the regular bicycle being substituted by the e-bike for distances under 7 km, while the use of the car (as driver) also slightly declines for longer distances above 7 km. Additionally, short-distance BTM trips and long-distance train trips have both decreased, but due to the small sample sizes of these types of public transportation trips in ODiN, it is difficult to draw strong conclusions from these findings. Moreover, car passenger trips above 25 km show a slight increase, but more other evidence, such as longitudinal analysis, is needed to make definitive conclusions on the effects of the e-bike car passenger use and public transport use.

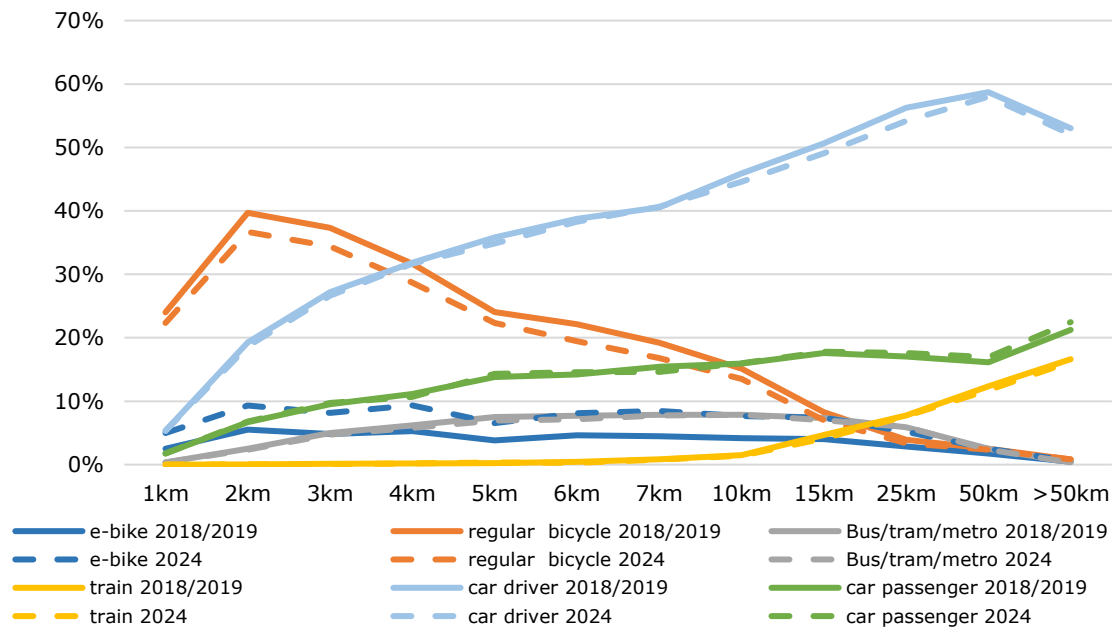


Figure 3 Effect of expected development of e-bike ownership on the modal split classified per distance by intention-based method

The development of e-bike ownership does not lead to an equally large increase in e-bike use for all purposes (see Figure 4). A relatively high number of working people intend to buy an e-bike. Therefore, we expect that the e-bike will have the largest impact on commuter traffic. E-bike use could rise by 122% for this purpose. The e-bike's share of the distance that commuters cover by bicycle would then rise from 23% to 44% and the total distance cycled for this purpose would rise by about 17%. The e-bike currently accounts for a quarter (26%) of all bicycle trips made for leisure or for shopping, this share will increase to 40% and 36% respectively. For both purposes, the total distance covered by bicycle would increase by about 5%. The use of the e-bike for going to school would increase by 110%, which would increase the e-bike's share in the total distance covered by bicycle from 7% to 14%. The total distance covered by e-bike would also increase by about 5% for this purpose.

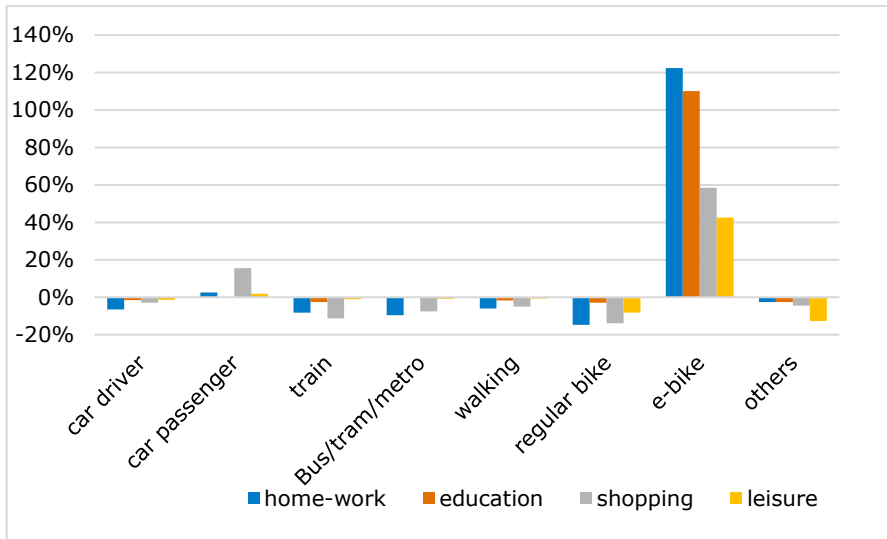


Figure 4 Changes in distances covered per purpose per means of transport by intention-based method

4. CONCLUSIONS

We estimate the role of e-bikes in the five-year period from 2019 to 2024, taking into account the growth in e-bike ownership and travel behaviour of new e-bike owners. To validate our findings, we also used time series forecasting to cross-check the e-bike share estimation.

We find the distance covered by e-bikes is expected to increase significantly between 2019 and 2024, while the distance covered by regular bicycles is expected to decrease, with e-bikes primarily substituting regular bicycles on shorter distances. We expect that the use of e-bikes for commuting will increase significantly and that there may be a reduction in car use for longer distances. Furthermore, there is a possibility that e-bike use will lead to a decrease in public transport use, including bus/trams/metros and train, in the future.

The expected substitution of car use by e-bikes represents a positive contribution to sustainable mobility. This indicates that, to a certain extent, promoting e-bike use results in a shift towards more sustainable travel behaviour. At the same time, promoting e-bike use may also result in a reduction of the normal bicycle and public transport. If policymakers want to promote e-bike use, our previous study (de Haas & Huang, 2022) identified a number of key action points that policymakers could use to develop policies aimed at encouraging use of the e-bike. These include improving facilities and infrastructure such as guarded bicycle parking facilities and broader cycle paths with safer crossing points, increasing the cost of other modes of transport like cars and addressing barriers to commuting such as improving facilities at the work place (e.g., showers, changing areas, and providing secure bicycle parking).

REFERENCES

- BOVAG. (2023). Aandeel e-bikes in fietsverkoop blijft stijgen in 2022. <https://www.bovag.nl/nieuws/aandeel-e-bikes-in-fietsverkoop-blijft-stijgen-in>
- Bürkner, P. C. (2017). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–411.
- Bürkner, P. C., & Charpentier, E. (2020). Modelling monotonic effects of ordinal predictors in Bayesian regression models. *Br. J. Math. Stat. Psychol.* 73(3), 420-451.
- CONEBI: European Bicycle Market 2017 Edition. Industry & Market Profile, Brussels (2017)
- Cherry, C. R., Yang, H., Jones, L. R., & He, M. (2016). Dynamics of electric bike ownership and use in Kunming, China. *Transp. Policy*, 45, 127-135.
- de Haas, M., & Huang, B. (2022). Aanschaf en gebruik van de elektrische fiets.
- de Haas, M., Kroesen, M., Chorus, C., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2021). E-bike user groups and substitution effects: evidence from longitudinal travel data in the Netherlands. *Transportation*, 49(3), 815-840.
- Fishman, E., Cherry, C., 2016. E-bikes in the mainstream: reviewing a decade of research. *Transp. Rev.* 36 (1), 72–91.
- KiM. (2021). The Netherlands Mobility Panel.
- Kroesen, M. (2017). To what extent do e-bikes substitute travel by other modes? Evidence from the Netherlands. *Transp. Res. Part D Transp. Environ.* 53, 377-387.
- Stan Development Team. (2022) Stan Modeling Language: User’s Guide and Reference Manual, Version 2.31. <https://mc-stan.org>.
- Statistics Netherlands (CBS), 2010–2021. Mobility Survey Netherlands and Dutch National Travel survey.
- Stuart, Elizabeth A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science* 25 (1): 1–21.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27, 1413-1432.
- Vrána, J., Remeš, V., Matysioková, B., Tjørve, K. M., & Tjørve, E. (2019). Choosing the right sigmoid growth function using the unified-models approach. *Ibis*, 161(1), 13-26.
- Wolf, A., Seebauer, S.: Technology adoption of electric bicycles: A survey among early adopters. *Transp. Res. Part A Policy Pract.* 69, 196–211 (2014).

Attitudes and Latent Class Choice Models using Machine Learning

Lorena Torres Lahoz*¹ and Francisco Camara Pereira¹, Georges Sfeir¹, Ioanna Arkoudi¹, Mayara Moraes Monteiro¹, Carlos Lima Azevedo¹

¹DTU Management, Technical University of Denmark

SHORT SUMMARY

We present a method of efficiently incorporating attitudinal indicators in the specification of Latent Class Choice Models (LCCM), extensions of Discrete Choice Models (DCMs) that segment populations based on the assumption of preference similarities. We introduce Artificial Neural Networks (ANN) to formulate the latent variables constructs. This formulation overcomes structural equations in its ability to explore the relationship between the attitudinal indicators and the decision choice, given the machine learning (ML) flexibility and power to capture unobserved and complex behavioural features, such as attitudes and beliefs. All of this, while maintaining the consistency of the theoretical assumptions presented in the Generalized Random Utility model and the interpretability of the estimated parameters. We test our proposed framework for estimating a car-sharing service subscription choice with stated preference data. The results show that our proposed approach provides a complete and realistic segmentation, which helps design better policies.

Keywords: Car-sharing, Discrete choice modelling, Machine learning, Psychometric Indicators.

1 INTRODUCTION

This study explores a new method of efficiently incorporating attitudinal indicators in the specification of LCCM by relying on ML techniques while preserving the benefits of the economic and behavioural interpretability of DCMs.

Walker & Ben-Akiva (2002) presented a practical generalized random utility model with extensions for latent variables and classes. They extended the Random Utility Model (RUM) to relax its assumptions and enrich the model's capabilities. They refer to latent classes as unobserved population groups, in which each individual has an associated probability of belonging to each group/class on the assumption of preference similarities. On the other hand, psychometric indicators measure the effect of unobserved attributes on individuals' preferences on topics related to the choice and they are additional information that helps specify and estimate latent classes.

Atasoy & Bierlaire (2011) estimated an LCCM where psychometric indicators are included in the maximum likelihood estimation to improve the model's accuracy. The psychometric indicators were modelled, conditional on the latent class, as parameters jointly estimated with the choice and the class membership model. The model showed that the psychometric indicators allow for richer analysis and generate significantly different class membership estimates. In another approach, Hurtubia et al. (2014) introduced psychometric indicators by computing the probability of giving an agreement level to an attitudinal statement as an ordinal logit, also dependent on the individual class. However, complex interactions between attitudinal variables and the decision-making process should be expected Bahamonde-Birke et al. (2017). We hypothesise that ML could be a good starting point to explore such interactions, given its flexibility and power in capturing unobserved and complex interactions.

In recent years, the use of ML techniques has increased, mainly due to their power to improve prediction accuracy. However, one of the main critiques of ML techniques in contrast to econometric models, is that they tend to generate less interpretable results. Thus, transportation researchers have focused on providing meaningful estimates from ML applications, that can be useful for travel analysis and policy decisions. For example, Arkoudi et al. (2021) proposed an embedding encoding for the socio-characteristic variables that provided a latent representation of these variables in concordance with individuals' choices. Han (2019) included a nonlinear LCCM using a neural network to specify the class membership model. Their model outperformed the traditional ones

in prediction accuracy with the trade-off of losing some interpretability. Sfeir et al. (2021, 2022) presented two model formulations for the construct of latent class choice models using Gaussian process and Mixture models. All these works employed ML in DCMs to allow for more flexibility in the definition of the latent constructs. However, there is still a lack of effective use of these techniques for incorporating attitudinal information into the model formulation.

2 METHODOLOGY

We follow the generalized RUM structure presented by Walker & Ben-Akiva (2002) for interpretability purposes and we include the information on the attitudinal indicators by employing an ANN to formulate with greater flexibility the latent variables. Figure 1 shows the graphical representation of the proposed formulation.

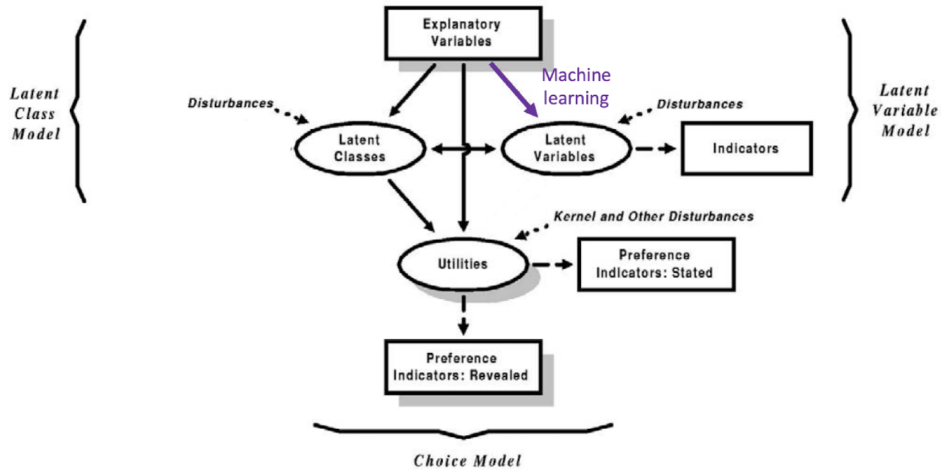


Figure 1: Graphical representation of the model formulation

LCCMs are composed of two sub-models: a *class membership model* and a *class-specific choice model*. The former computes the probability of an individual n belonging to a certain class, while the latter assigns the probability of choosing each alternative, given that individual n belongs to a certain class k .

The utility of the class membership model can be written as:

$$U_{nk} = V_{nk} + v_{nk} \quad (1)$$

where V_{nk} is the representative utility of individual n belonging to class k and v_{nk} is the error term that is assumed to be independent and identically distributed (iid) Extreme Value Type I over individuals and classes. In this case, we define V_{nk} as:

$$V_{nk} = ASC_k + Q_n \gamma_k + r_n \delta_k + \omega_n b_k \quad (2)$$

where ASC_k is the alternative-specific value for class k , Q_n is the vector containing socio-characteristics of individual n , and γ_k the vector of unknown parameters that need to be estimated for each class k . In addition, r_n is a vector of length Z containing the latent variables for individual n and δ_k the corresponding vector of unknown parameters specific to class k . Finally, ω_n is an individual-specific constant with its corresponding coefficient b_k for each class k . It represents the individual variation of all the latent variables caused by the variance of their underlying distributions. It is formulated as a one-layer ANN that gets activated by the ID of each individual in the train set (Id_n is one for individual n and 0 otherwise),

$$\omega_n = \sum_1^N w_{1n}^{(1)} Id_n \quad (3)$$

where $w_{1n}^{(1)}$ are the weights of the layer.

Given the distribution of the error term (v_{nk}), the probability $P(q_{nk}|Q_n, \gamma_k, r_n, b_k)$ can be expressed as:

$$P(q_{nk}|Q_n, \gamma_k, r_n, b_k) = \frac{e^{V_{nk}}}{\sum_{k'=1}^K e^{V_{nk'}}} \quad (4)$$

The novelty of this work is the employment of ANN for the construction of latent variables. We propose a non-linear relationship between the socio-characteristics of the individuals and the latent constructs by employing two densely connected layers:

$$r_{zn} = a_2\left(\sum_{h=0}^H w_{zh}^{(2)} a_1\left(\sum_{m=0}^M w_{hm}^{(1)} Q_{mn}\right)\right) \quad (5)$$

where M is the number of socio-characteristic variables used to predict the answer to the indicators, and H is the number of hidden units in the hidden layer. $w_{hm}^{(1)}$ are the weights of the first layer, and a_1 represents the first activation function defined as a Rectified Linear Unit (ReLU) ($a_1(x) = \max(0, x)$); for the second layer, a linear activation function is applied $a_2(x) = x$, and the weights are represented by $w_{zh}^{(2)}$. By adding an extra input Q_{0n} , which is set to one and extending the sum to go from zero, we avoid writing the intercept term.

The number of latent variables Z , the number of hidden neurons in the hidden layer H , and the number of densely connected layers should be tuned since they are not observed in the data.

The formulation presented is based on the hypothesis that the socio-characteristics of the individuals define the latent variables. Moreover, these latent constructs influence the response to specific attitudinal indicators. We focus on the case where indicators take the form of statements that receive an ordered response, in Likert (1932) scale. Thus, we define the utility of individual n for indicator p , as a measurement of the level of agreement with the statement, and we formulated it as:

$$U_{pn} = V_{pn} + \nu_{pn} = r_n \alpha_p + c_p \omega_n + \nu_{pn} \quad (6)$$

where V_{pn} is the representative utility of individual n to indicator p and ν_{pn} is the error term that is assumed to be iid Extreme Value Type I over individuals and indicators. r_n is a vector of length Z containing the latent variables of individual n , α_p is the vector of corresponding parameters to be estimated. ω_n is the individual-specific parameter estimated together with the latent class model, and c_p is its corresponding coefficient for each indicator p .

Therefore, the probability that individual n answers with a certain level of agreement l to indicator p is expressed as:

$$P(I_{pln} = 1|r_n, \alpha_p, c_p, \omega_n) = P(\tau_{l-1}^p < U_{pn} < \tau_l^p) \quad (7)$$

where we define I_{pln} as 1 if individual n answers with a level of agreement l to indicator p and 0 otherwise. τ_l^p are strictly increasing class-specific thresholds that define an ordinal relation between the utility U_{pn} and the level of agreement to indicator p .

The probability of individual n providing an answer l to indicator p can be computed as an ordinal softmax:

$$\begin{aligned} P(I_{pln} = 1|r_n, \alpha_p, c_p, \omega_n) &= P(\tau_{l-1}^p < U_{pn} < \tau_l^p) = P(\tau_{l-1}^p < V_{pn} + \nu_{pn} < \tau_l^p) = \\ &= Prob(\nu_{vp} < \tau_l^p - V_{pn}) - P(\nu_{vp} < \tau_{l-1}^p - V_{pn}) = \frac{e^{\tau_l^p - V_{pn}}}{1 + e^{\tau_l^p - V_{pn}}} - \frac{e^{\tau_{l-1}^p - V_{pn}}}{1 + e^{\tau_{l-1}^p - V_{pn}}} \end{aligned} \quad (8)$$

where one threshold per indicator is set to zero, as only the difference between them matters.

We estimate all components of the proposed model simultaneously by employing the EM Dempster et al. (1977) algorithm, which combines an expectation step with a maximization one until convergence is reached. The final model architecture is presented in Figure 2.

3 RESULTS AND DISCUSSION

We test the model on a dataset from a 2020 tailor-made online survey in Copenhagen (CPH). Respondents needed to be at least 18 years old and have a valid driver's license. The sample consists of 542 complete answers from which 80% are used for training and 20% for testing. The relevant parts employed in the estimation include:

1. A survey on the respondent's socio-characteristic characteristics

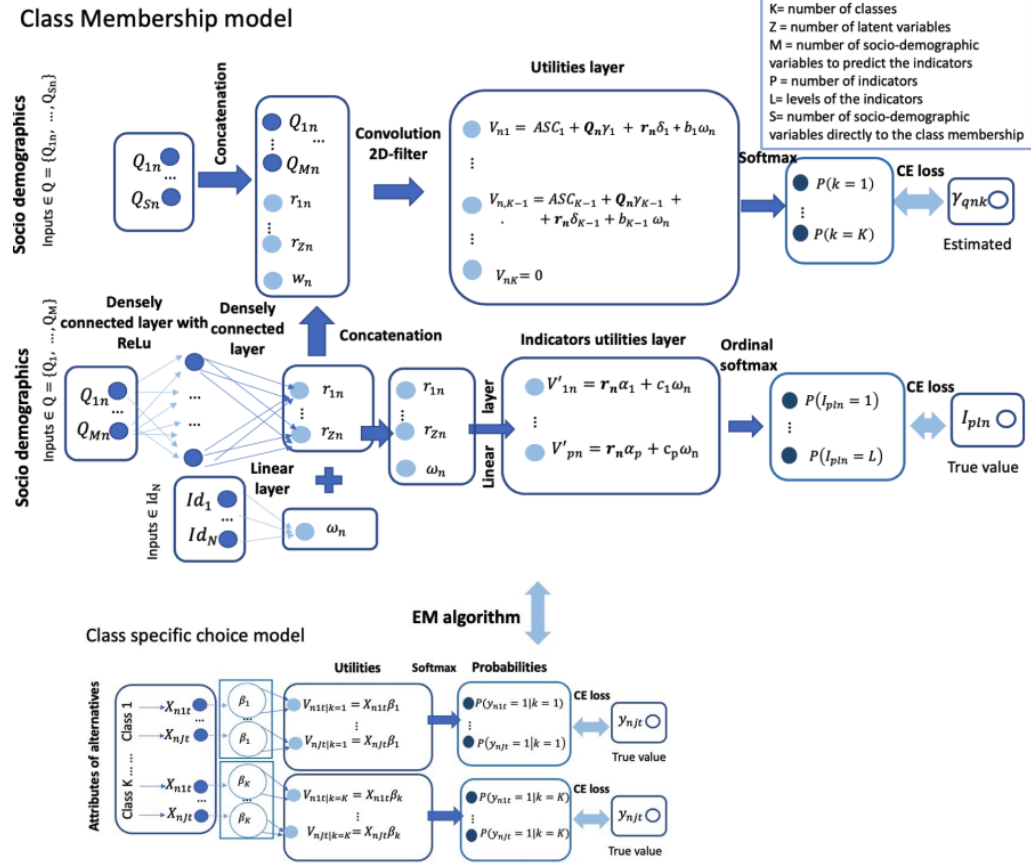


Figure 2: Model architecture

2. A survey addressing questions regarding respondents' attitudes toward private and car-sharing (CS) using a 5-point Likert (1932) scale
3. A Stated Preference (SP) experiment with different options for CS plans

For further details on the data, the reader is referred to Frenkel et al. (2021).

Baseline Results

To benchmark the proposed model, we tried to follow the formulation from Walker & Ben-Akiva (2002) which can be shown as a graphical representation (Figure 2), with the difference that the relation between the indicators' utility and the individuals socio-characteristic characteristics is linear. However, the class membership formulation's simplicity made the model unable to converge, resulting in a non-invertible Hessian matrix. Therefore, to get a comparable magnitude for the likelihood, we estimated the model with a traditional LCCM as a baseline, where the representative part of the class membership utility is just a linear combination of the socio-characteristic variables. Results are presented in Table 1.

Table 1: LCCM results without attitudinal variables

Model	N ^o Classes	N ^o parameters	Null LL	LL	AIC	BIC	R-squared	Test null LL	Test LL
LCCM	2	30	-2047.21	-1599.41	3258.81	3413	0.22	-515.02	-400.52
LCCM	3	43	-2047.21	-1568.14	3234.29	3487	0.23	-515.02	-398.73

We employ the same socio-characteristics for constructing the class membership as in our proposed formulation. More specifically, we include age, binary variables that indicate if the individual has a bike, a car or kids at home and if they are students, retired or CS members. However, we found that having a car at home, being retired or being a student, were not statistically significant under

the LCCM formulation. Instead, our proposed formulation allows us to significantly include this information in the model, improving our characterization of the latent variables. Given the probability of each individual in the sample and its corresponding socio-characteristics, we have represented the classes from the baseline model in Figure 3, by using the Bayes Theorem to compute:

$$P_n(\text{socio-characteristic} | K = k) \quad (9)$$

Figure 3 is compared in the next subsection with the proposed model results.

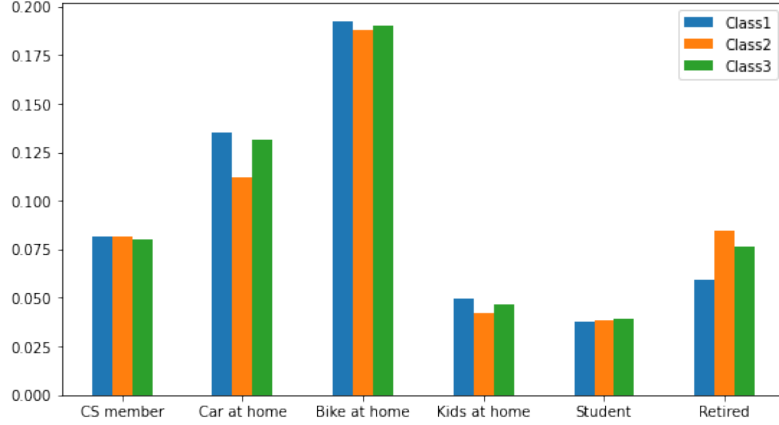


Figure 3: Representation of the class membership of the LCCM model

Proposed Model Results

We have employed the same train/test split as for the baseline model. The EM process has been estimated multiple times with random initializations. We have computed the likelihood variance between the different model estimates to check for stability. The results are summarized in Table 2.

Table 2: Model results

Nº Classes	Nº latent variables	Iterations	Null LL	LL	Variance LL	R-squared	Test null LL	Test LL	Variance Test LL
2	2	15	-2047.21	-1575.32	14.62	0.23	-515.02	-404.72	0.28
3	2	30	-2047.21	-1539.56	22.01	0.25	-515.02	402.72	9.8
3	3	25	-2047.21	-1531.41	26.36	0.25	-515.02	-400.73	9.36

The model with three latent classes and two latent variables is selected as the best model. The one with three classes and three latent classes has a slightly better fit, however, its corresponding latent variables parameter estimates were not statistically significant.

Comparing the results from tables 1 and 2, we observed an increase in the training likelihood for our formulation. We do not provide better results for the test data, but just comparable ones. This could be due to the small size of the test sample and/or to the fact that we don't have access to attitudinal information or w_n values in the test stage, which affect prediction accuracy.

Table 3 shows the estimated parameters of the class-specific choice model with their corresponding standard deviations, where all the data has been used for the estimation. The utility for not choosing any of the CS services is set to zero due to parameters' identification.

Based on the values and signs of the estimated beta parameters, we observe that class 1 and class 2 are more negatively affected by the subscription cost, while class 3 is less influenced by this cost, but more negatively affected by the usage cost. Moreover, individuals with a high probability of belonging to class 3 are the most concerned if the type of engine is combustion. Thus, CS could be seen as an electric alternative for them. Given the beta values for displaying the cost in hours ($\beta_{Usage\ cost\ per\ hour}$) or days ($\beta_{Usage\ cost\ per\ day}$), there is a bias towards displaying the price per minute (baseline), related to the fact that CS users tend to drive for short time periods. Regarding the probability of finding a car, it is a more important feature for classes 2 and 3, which make them more dependent on the availability of the service. Overall, class 2 seems to be less prone to use any CS (including P2P), given all its estimated parameters.

Table 3: Estimate and standard deviation of the parameters of the class-specific choice model

Variable	Class specific choice model		
	Class 1	Class 2	Class 3
ASC_{CS} free-floating	3.50(0.46)	-2.71(1.34)	-1.68(0.89)
ASC_{CS} station-based	3.04(0.48)	-2.64(1.34)	0.07(0.80)
ASC_{CS} peer to peer	4.12(0.52)	0.47(1.45)	1.50(0.83)
$ASC_{roundtrip}$	2.97(0.48)	-3.50(0.40)	0.15(0.80)
$\beta_{One\ time\ subscription\ cost}$	-1.00(0.17)	-1.12(0.51)	-0.35(0.22)
$\beta_{Usage\ cost(OWFF,OWST,RT)}$	0.05(0.04)	-0.16(0.09)	-0.34(0.08)
$\beta_{Usage\ cost(P2P)}$	-1.30(0.39)	-5.57(1.31)	-3.86(0.80)
$\beta_{Usage\ cost\ per\ day}$	-0.35(0.24)	-2.40(0.74)	-1.09(0.39)
$\beta_{Usage\ cost\ per\ hour}$	-0.10(0.21)	-0.99(0.50)	-0.43(0.34)
$\beta_{Only\ combustion\ cars}$	-0.23(0.12)	0.09(0.33)	-0.66(0.20)
$\beta_{Probability\ of\ finding\ a\ shared\ car}$	0.13(0.44)	1.38(1.35)	2.70(0.75)
$\beta_{Walking\ time\ from\ parking\ to\ destination}$	-0.05(0.01)	0.02(0.04)	0.03(0.02)

Table 4: Parameters of the class membership model

Variable	Parameter	St error	P-value
ASC_{class_1}	0.97	0.45	0.031
ASC_{class_2}	-1.40	0.51	0.0057
$\gamma_{kidsathome,class_1}$	0.56	0.29	0.050
$\gamma_{kidsathome,class_2}$	-0.45	0.37	0.22
$\delta_{r_1,class_1}$	0.18	0.13	0.17
$\delta_{r_1,class_2}$	-0.48	0.14	0.0009
$\delta_{r_2,class_1}$	0.12	0.10	0.27
$\delta_{r_2,class_2}$	-0.44	0.11	0.00
b_{class_1}	0.61	0.44	0.17
b_{class_2}	-4.014	0.54	0.00

The parameters of the class membership model are summarised in Table 4. Given the probability of each individual in the sample, we have characterised the classes in Figure 4. Individuals with a higher probability of belonging to class 1 have around 20% probability of being a CS member, a bit above the sample average (17.5%). They also tend to have more kids at home, as well as bikes than other classes. Studies like Uteng et al. (2019) have shown that when there are significant life changes (e.g., birth of a child), people become more inclined to use CS. In opposition, class 2 presents the lowest probability of being a CS member and having kids or bikes at home. Retired people tend to have more predisposition for this class, while students have less. This is aligned with Prieto et al. (2017) which suggested that young people are more prone to use this service. Finally, class 3 has the same probability of being a CS member as class 1, but it also has a lower probability of owning a car, making people more reliant on the service’s availability. Comparison between Figures 3 and 4, shows that the configuration of the classes changes when we include attitudinal information, as it is expected.

By analysing the parameters for the latent variables in Table 4 and looking at their distributions over individuals in Figure 5, we notice that the values of the first latent variable (r_1) are always negative. The more negative value of r_1 , the more probable is to belong to class 2, and therefore, the less inclined people are to use CS services. A negative value of r_2 seems to have the same effect. Thus, individuals with a more negative combination of r_1 and r_2 tend to be less inclined about CS and the other way around. Figure 5 suggests that students are more prone to use the service while retired people are the least predisposed. Moreover, having or not having a car seems to determine the clusters in which the r_s values are structured. Finally, Figures 6 and 5 show that people with a car at home agree more with the statement that the car is a status symbol. For indicator 15, people with a more positive value of r_2 seem to agree more with the statement that

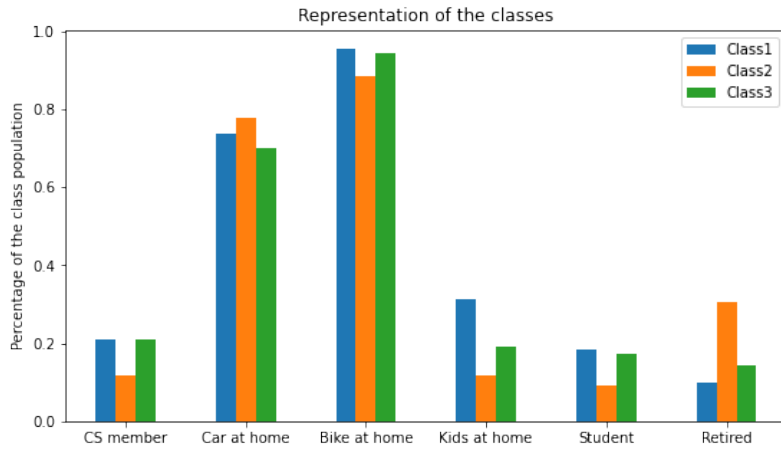


Figure 4: Representation of the class membership of our proposed model

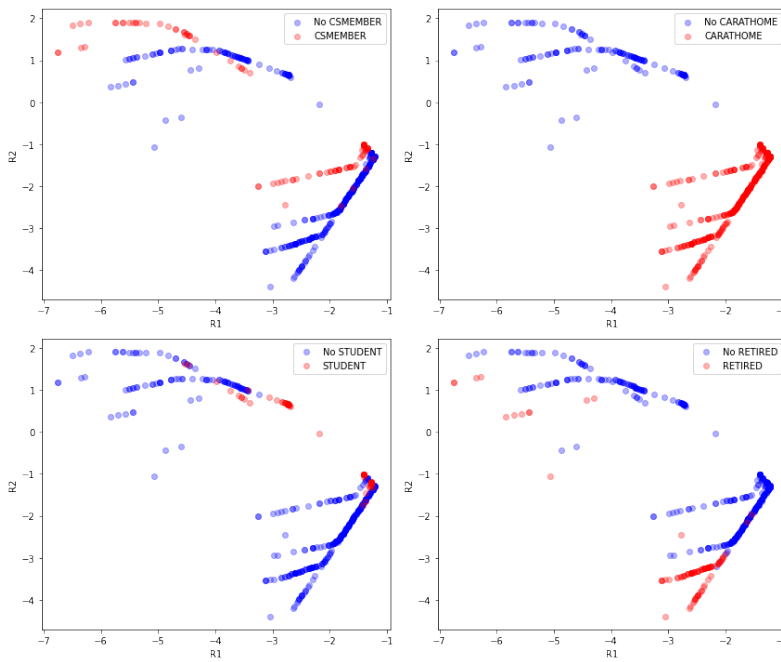


Figure 5: Latent variables representation

they wouldn't need a car if they have CS, as we would expect given the r_2 coefficients of Table 4.

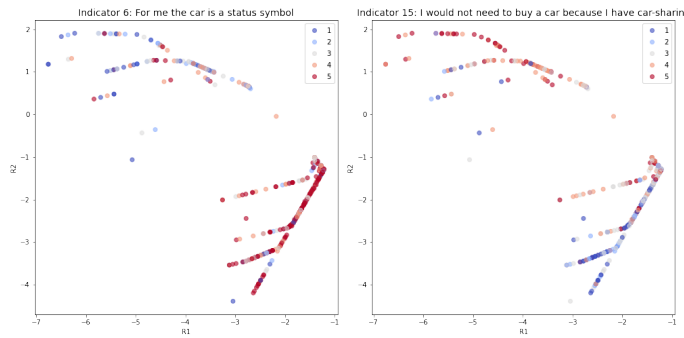


Figure 6: Latent variables representation characterized by the answers to indicators 6 and 15

4 CONCLUSIONS

Our results suggest that the inclusion of attitudinal variables provides a DCM that is more behaviorally realistic. For example, individuals who are more inclined towards the concept of CS tend to be grouped together in clusters with higher parameter estimates of the utility of choosing CS plans. This indicates that beliefs and attitudes play a key role in decision-making, and including this information allows for more accurate estimation and a better understanding of the classes that help design better policies.

Within the limitations, convergence is defined empirically by setting the number of iterations due to small fluctuations in the convergence of the EM algorithm. In addition, given the small sample size, we could not divide the dataset in training, validation, and testing; therefore, the hyperparameters of the ANN were not tuned according to the validation samples. This could be solved by employing a bigger dataset. Moreover, to improve the prediction performance, other types of explainable AI (e.g., SHAP) could be explored.

Although the limitations, we are optimistic that this analysis has opened the door to future research on integrating attitudinal variables in DCMs through ML techniques.

ACKNOWLEDGEMENTS

We thank Share-More project (<https://eitum-sharemore.net.technion.ac.il>) for providing the data used in this study

REFERENCES

- Arkoudi, I., Azevedo, C. L., & Pereira, F. C. (2021). Combining discrete choice models and neural networks through embeddings: Formulation, interpretability and performance. Retrieved from <https://arxiv.org/abs/2109.12042>
- Atasoy, G. A., B., & Bierlaire, M. (2011). Mode choice with attitudinal latent class: a swiss case-study. *Proceedings of the Second International Choice Modeling Conference (ICMC)*.
- Bahamonde-Birke, F. J., Kunert, U., Link, H., & Ortúzar, J. d. D. (2017). About attitudes and perceptions: finding the proper way to consider latent variables in discrete choice models. *Transportation*, *44*, 475–493. doi: <https://doi.org/10.1007/s11116-015-9663-5>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, *39*(1), 1–38. Retrieved 2023-02-08, from <http://www.jstor.org/stable/2984875>
- Frenkel, A., Shiftan, Y., Gal-Tzur, A., Tavory, S., Lerner, O., Antoniou, C., . . . Peretz, D. (2021). Share more: Shared mobility rewards - summary report.
- Han, Y. (2019). *Neural-embedded discrete choice models* (Unpublished doctoral dissertation).
- Hurtubia, R., Nguyen, M. H., Glerum, A., & Bierlaire, M. (2014). Integrating psychometric indicators in latent class choice models. *Transportation Research Part A: Policy and Practice*, *64*, 135-146. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0965856414000755> doi: <https://doi.org/10.1016/j.tra.2014.03.010>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140).
- Prieto, M., Baltas, G., & Stan, V. (2017). Car sharing adoption intention in urban areas: What are the key sociodemographic drivers? *Transportation Research Part A: Policy and Practice*, *101*, 218-227. doi: <https://doi.org/10.1016/j.tra.2017.05.012>
- Sfeir, G., Abou-Zeid, M., Rodrigues, F., Pereira, F., & Kaysi, I. (2021). Latent class choice model with a flexible class membership component: A mixture model approach. *Journal of Choice Modelling*, *41*.
- Sfeir, G., Rodrigues, F., & Abou-Zeid, M. (2022). Gaussian process latent class choice models. *Transportation Research. Part C: Emerging Technologies*, *136*.

- Uteng, T. P., Julsrud, T. E., & George, C. (2019). The role of life events and context in type of car share uptake: Comparing users of peer-to-peer and cooperative programs in oslo, norway. *Transportation Research Part D: Transport and Environment*, *71*, 186-206. doi: <https://doi.org/10.1016/j.trd.2019.01.009>
- Walker, J., & Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical social sciences*, *43*(3), 303-343.

How mobile are persons with mobility restrictions? Analysis of number of days with activities using one-week activity schedules in Germany

Ana Tsui Moreno*¹, Matthias Langer¹, Rolf Moeckel¹

¹ School of Engineering and Design, Technical University of Munich, Germany

SHORT SUMMARY

In Germany, 13% of all residents are disabled and 9.3% are even classified as severely disabled, which includes elderly people with limited mobility as well as physically disabled and mentally disabled people. Persons with mobility restrictions often report on barriers to meet daily needs, which is usually reflected on fewer days to perform out-of-home activities. The objective of this research is to evaluate whether persons with mobility restrictions are less mobile using one-week activity schedules. The results of the models confirm that persons with mobility restrictions are generally less mobile; being statistically significant for work, shop and recreation activities. It was found a significant interaction between occupation status and mobility restriction on the number of mobile days of most activity types, as well as an impact of the number of mobile days for mandatory activities on the number of mobile days for discretionary activities.

Keywords: Activity-generation, Mobile, Activity-based models, Disability, Week travel diary, Household travel survey

1. INTRODUCTION

Transportation is an important component of reaching amenities such as care facilities, social and family contacts, education, or work, and generally contributes to quality of life (Best et al. 2022). Due to physical, psychological, social, or socio-economic factors, individuals with impeded mobility often face difficulties while travelling. In view of the United Nations' Convention on the Rights of Persons with Disabilities transport research must focus in more detail on barriers and solutions. The convention does not only aim "to promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities" but in its article nine focusses in detail on travel accessibility (United Nations 2006). Pursuing these goals can lead to inclusivity and social justice as parts of transport equity (Litman 2022). The objective of this research is to evaluate whether the persons with mobility restrictions are less mobile than persons without mobility restrictions, in terms of number of days that they perform out-of-home activities, by activity type.

The meaning of disabilities

Disabilities are complex, dynamic, multidimensional, and controversial conditions that involve health conditions following to activity limitations and societal participation restrictions (World Health Organization 2011). They are as diverse as the people who suffer from them and can be short or long term, painful or painless, or even be visible or invisible. By understanding disability as an interaction, not as a characteristic of a person, it is recognized that affected people, are differentiated by factors such as gender, socio-economic status, or origin, which bring with them varying social disadvantages (World Health Organization 2011). In Germany, 13% of all residents

are disabled and 9.3% are even classified as severely disabled, which includes elderly people with limited mobility as well as physically disabled and mentally disabled (Statistisches Bundesamt, Wirtschaft und Statistik 2012).

Speaking of the diversity of disabilities, Frye (2019) lists physical, vision, hearing, and cognitive impairments, as well as mental health issues leading to limitations and restrictions in transportation. It is important to understand that both the type of impairment and surrounding environmental factors influence the consequences for affected individuals.

Daily mobility for mobility-impaired individuals

Studies of travel behavior on cognitively impaired persons (Rosenkvist et al. 2009) and those with mobility-impaired persons (Best et al. 2022), studies in different global regions (Frye 2019), in urban or rural areas, or even in differently developed neighbourhoods may differ substantially.

Many countries have laws that guarantee daily accessibility for people with disabilities (Bekiaris et al. 2018). Nevertheless, impaired individuals often report on barriers to meet daily needs. Measures to improve accessibility include information and driver training, pedestrian walkways and street crossings, public transit stops and station infrastructure, public transit vehicles, and appropriate private transportation (World Bank 2013). Best et al. (2022) summarize those as Availability, Accessibility, Accommodation, Affordability, Acceptability, and Awareness. Opposing this are several obstacles impeding the daily commute.

In order to live a qualitative life, people with reduced mobility adapt to the mentioned circumstances. A number of studies have looked at their resulting travel behavior. Schmöcker et al. (2008) focus on shopping trips of elderly and disabled people, stressing the heterogeneity of these individuals and their behavior. As such, Rosenkvist et al. (2009) conduct interviews with cognitively impaired people who they believe are rarely studied. Using data from the UK National Travel Survey, Lucas et al. (2016) found that disabled people travel fewer and shorter distances on average. This can be attributed primarily to the lower number of leisure trips (Jansuwan et al. 2013), which could be a result of lacking accessible social activities (Lucas et al. 2016). Schmöcker et al. (2008) add that the trips also differ in their complexity and that trips are primarily made with a single destination.

Park et al. (2022) conducted a systematic literature review, analysing 115 peer-reviewed papers on travel behavior for persons with reduced mobility. They found that people with disabilities take up to 30% fewer trips than people without disabilities. Likewise, a lower amount of non-work trips, increased use of public transportation, cabs, and ridesharing, and in turn, decreased walking distances and car trips were identified. The review highlighted that “environmental, social, and system barriers make specific modes unavailable to travelers with disabilities, increase travel time, and eventually decrease their trip frequency” (Park et al. 2022). They concluded that the sum of negative travel experiences “can ultimately lower social inclusion and the quality of life” (Park et al. 2022).

2. METHODOLOGY

Data source

To the authors' knowledge, travel behavior in Germany has not previously been studied in relation to persons with mobility restrictions in particular. The 2017 National Travel Survey included one chapter on “Health-related limitations – influence on mobility in an aging society”, which primarily focused on the elderly (Nobis and Kuhnimhof 2018). It is mentioned that 13 percent of the population is affected by health limitations, half of them suffer from mobility restrictions, which

is in line with official statistics. They also report that more than 1.5 million people in Germany do not own a car solely for health reasons, affecting travel behavior.

For this study, another important data source for understanding the mobility behavior of the German population, the German Mobility Panel, was used. This large-scale, nationwide survey by the German Federal Ministry of Transport and Digital Infrastructure is conducted every two years and collects information on travel behavior, costs, satisfaction and individual participants. To do this, participants fill out a travel diary on seven consecutive days and provide information on their choice of transportation, reason for travel, travel time and distance. In addition, a household survey is answered about the place of residence and public transport connections, as well as personal questions about age, gender, employment and also mobility restrictions (Vallée et al. 2022). Mobility restrictions were self-reported, without distinguishing by type or degree. Panel data from 2010 to 2019 was analyzed, with a total of 18,700 individual records.

Model estimation

The main dependent variable is the number of days on which an individual perform a given activity (mobile days model). Person was selected as unit of analysis, in line with activity-based models (Hilgert et al., 2018).

The dependent variables are the number of days with a given activity, by activity type and the number of tours per day, by activity type. Therefore, their values could only be non-negative integers (e.g. 0, 1, 2, 3 etc.) and the responses are ordered. Previous approaches using linear regression (Vickerman 1974; Cervero and Kockelman 1997; Seo et al. 2013), multinomial logit models (Hilgert et al. 2018) or nested logit models (Yun and O'Kelly 1997) fail to account for the nature of the dependent variable. Count regression models or ordered logit models could be used instead. Typical distributions for count variables are Poisson or Negative Binomial distributions. The Poisson distribution requires the mean of the count process to be equal to its variance (Washington et al. 2020). If the requirement does not hold, Negative Binomial distributions could be used instead. An excessive number of zeros in the data could mean that it reflects both a normal count and a zero-count process. Models that can handle both states are denominated zero-inflated (Washington et al. 2020).

Preliminary analysis of the data showed overdispersion and a preponderance of zeros for some activity types, such as accompany or education, so zero-inflated negative binomial regression models were initially selected. However, these models could not capture a higher concentration of responses around 5, which reflects the 5-day commute pattern of the majority of full-time employees. To overcome this issue, we selected a two step model: a binomial logit model to model the zero-count state and a ordered logit model to model the count process state.

For each dependent variable, one model per activity type was estimated: work, education, accompany, shop, recreation and other discretionary activities. The models were executed in the R statistical software tool using the `pscl` package (Zeileis et al. 2008) and the `MASS` package (Ripley et al. 2023). Akaike Information Criteria (AIC), correlation between fitted values and simulation values, p-value of the variables and number of parameters were then used to determine the best model by activity type. Furthermore, we compared the observed and estimated number of individuals with zero mobile days to seven mobile days.

In line with most of activity generation models, independent variables included household size, gender, age, occupation status, economic status, car ownership or area type. Individual mobility restrictions were included, as well as their interaction with occupation status. For example, an employed person with mobility restrictions will perform more work activities than an unemployed person with mobility restrictions, but he/she may perform less work activities than an employed person without mobility restrictions.

On one hand, the number of mobile days for mandatory activities could impact the number of mobile days for discretionary activities. Likewise, not all discretionary activities may have the

same priority. In this sense, we established a hierarchy among activities: work, education, accompany, shop, recreation and other. The models were estimated in sequence following the hierarchy, and the number of mobile days of higher rank activities were added as explanatory variables to the model.

Independent variables were checked for correlation. Variables with correlation higher than 50% were not considered in the same set of independent variables. Alternative models were estimated with each set of independent variables, and the final model was selected based on goodness of fit.

3. NUMBER OF MOBILE DAYS

Preliminary analysis

The preliminary analysis on the number of mobile days by activity type is summarized in Figure 1. Each one of the subfigures indicates the distribution of mobile days by activity type, where blue highlights the share of individuals that did not perform that activity across the whole week, and the darker greens highlight individuals that are highly mobile. Each row and column represent the distribution by occupation status and mobility restriction (MR), respectively.

The top left subfigure we can observe that most of full-time workers without MR do go to work 5 days per week (52 %), compared to 19% that go to work 4 days per week. It is observed that, on an average week, 7% of full time workers do not go to work, either due to vacation or sickness. On the other hand, 98% of full-time worker do not travel for education any day of the week, and 64% do not travel for accompany or other discretionary purposes. There is a higher distribution among how many days do they perform shopping and recreation activities. It is observed that more than half only go for shopping or recreation 2 or fewer times per week, and that only 10% do shop 5 or more time per week. Most individuals distribute such activities either one, two or three days per week (17 - 26 %).

Not surprisingly, part-time employed individuals without MR commute less days than full-time employed, with only 24% of them being mobile 5 days per week (compared to 52%), and between 18 and 22% commuting to work 3 or 4 days per week. Regarding their discretionary activities, part-time employed allocate more days to accompany acts but they show similar distribution of days allocated to shop, recreation and other as full-time employed.

Students present a similar activity pattern as full-time employed to commute for education. However, their distribution of other discretionary activities is different, with fewer days for shopping and accompany. Last but not least, unemployed individuals hardly commute to work or education, and have similar distribution of accompany acts as full-time employed. They allocate more days to shop (only 4% do not shop in the entire week, and more than half shop at least 3 days per week),

The comparison between persons without and with MR shows that, generally, persons with MR do travel fewer days that persons without MR. A notable exception are shop days for students with MR, however the sample size was limited and may lead to non-representative results.

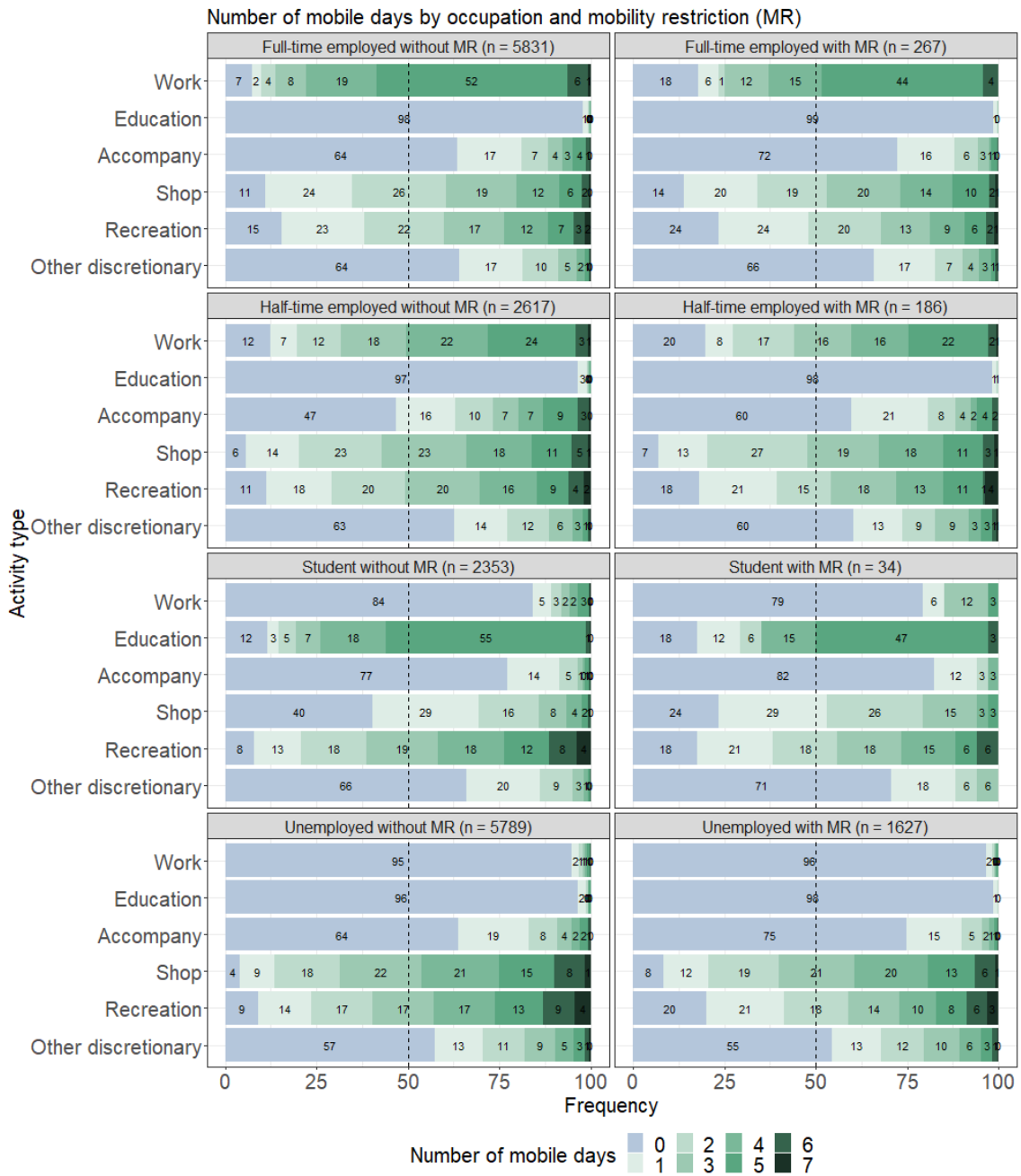


Figure 1: Number of mobile days per activity type by occupation and mobility status

Model estimation

The next step is model estimation. As the preliminary analysis showed patterns by occupation status and MR, their interaction term was added for model estimation. Model estimates for the zero-state and count-state are summarized in Tables 1 and 2, respectively. Observed vs. predicted frequencies are shown in Figure 2. As seen in the Figure, the models reproduce the aggregate

distribution of number of mobile days by activity type. The distributions show different patterns: from a preponderance of zeros and five mobile days for mandatory activities (work, education), to preponderance of zeros with low mean mobile days (accompany, other discretionary) to skewed distribution with mode equal to two mobile days (shop, recreation).

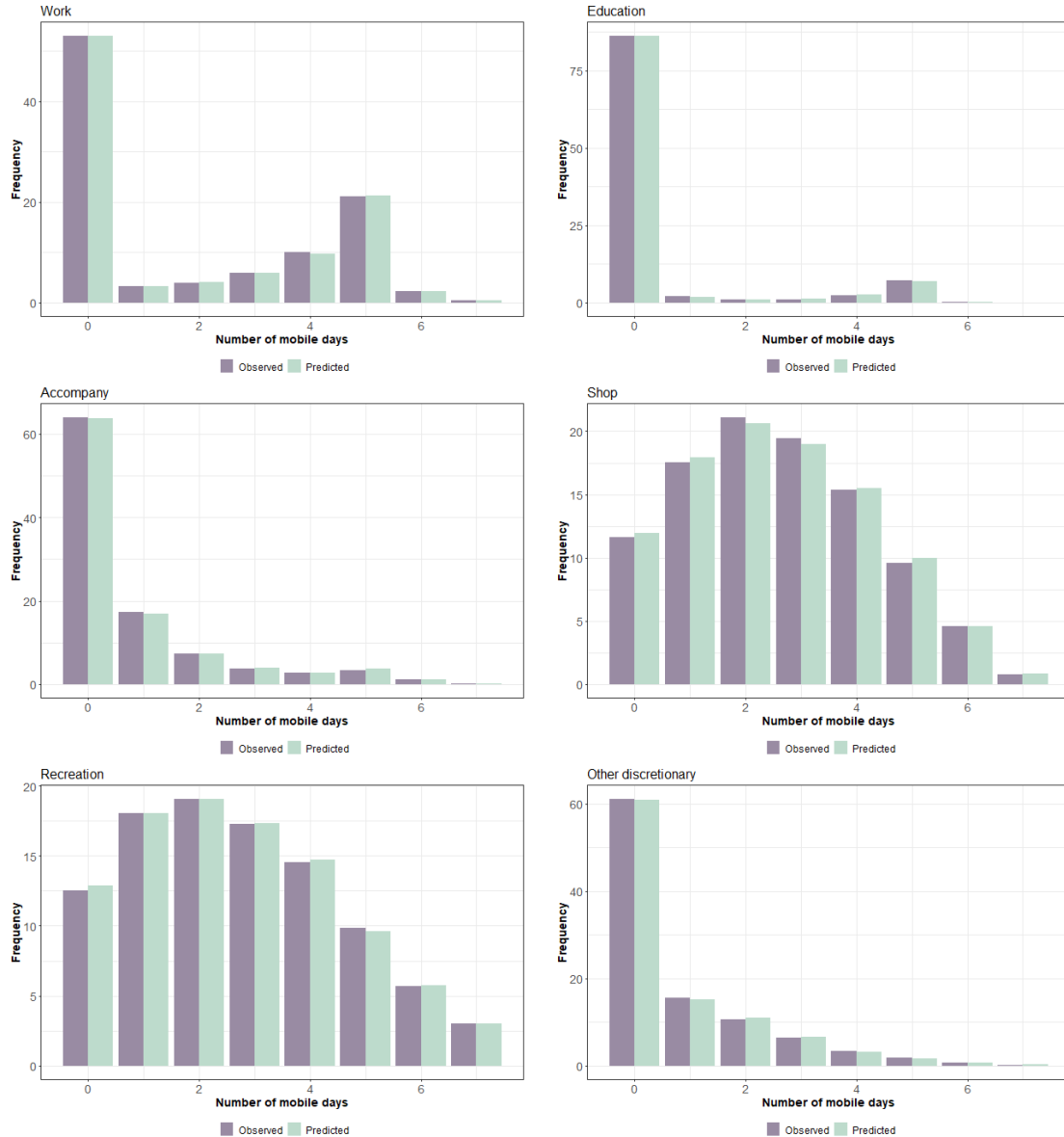


Figure 2: Number of mobile days per activity type. Observed vs. Predicted

Table 1 shows the estimates of the zero-state model. This binomial model estimated the likelihood of an individual to perform or not some activity along the week (non-mobile vs. mobile). Positive coefficients indicate higher likelihood of performing the activity compared to the baseline.

Regarding occupation status, the individuals with higher likelihood of not being mobile across the week for commute are students and unemployed; and among unemployed persons, persons with MR. Part-time employed also presented higher likelihood of not being mobile, compared to employed persons. Not surprisingly, students and part-time workers had higher likelihood of conducting at least one education activity along the week. Part-time employees were the most likely

to conduct any accompany, shop or recreation activity. We found a distinction based on mobility restriction: unemployed without MR are more likely to be mobile for shop and work activities, while unemployed with MR are less likely to be mobile for the same activities.

Being more mobile for mandatory activities increased the likelihood of not being mobile for discretionary activities, especially for accompany, shop and other. On the other hand, being mobile for accompany increased the likelihood of being mobile for shop, recreation and other; showing a complementary effect. All other variables show intuitive results, as larger households with more children have higher likelihood of doing accompany activities; and individuals aged 60 or older are less mobile than middle-age individuals.

Table 1: Zero-state model estimation

Variable	Work	Education	Accompany	Shop	Recreation	Other
(Intercept)	1.716	-3.706	-2.322	2.753	0.861	-1.091
Highly agglomerated areas		-0.219		-0.193	-0.427	2.713
Urbanized areas				-0.522	-0.465	2.691
Lower density urban/higher density rural				-0.383	-0.517	2.686
Household economic status: very low			-0.184	-0.217	-0.658	
Household economic status: low					-0.224	
Household economic status: very high						
Household size 2	-0.281		0.141	-0.707		
Household size 3	-0.185		0.476	-1.1		
Household size 4	-0.185		0.476	-1.1		
Household size 5 or more	-0.185		0.476	-1.1		
Children per household: 1			1.135			
Children per household: 2	-0.415		1.576			-0.253
Children per household: 3 or more	-0.853		2.035			-0.293
Adults per household: 1					0.495	
Adults per household: 2					0.502	-0.124
Adults per household: 3					0.424	-0.227
Adults per household: 4					0.46	-0.227
Between 10 und 17 years old			-1.67	-1.424	0.773	
Between 18 und 25 years old		-1.165	-0.363	-0.786	0.773	0.207
Between 36 und 50 years old		-1.169	-0.169		-0.251	0.104
Between 51 und 60 years old	-0.837	-1.401	-0.132			0.088
Over 61 years old	-0.837	-1.37	-0.132			
Occupation: Student	-3.705	5.81		-0.455	0.255	-0.286
Occupation: Part-time employed	-0.503	0.526	0.479	0.317	0.29	
Occupation: Unemployed		1.215				
Mobility Restriction: yes						
Employed with mobility restriction	-0.908			-0.447	-0.442	
Unemployed without mobility restriction	-4.916			0.49	0.509	
Unemployed with mobility restriction	-5.34			-0.286	-0.168	
Student (18 - 60 years old)						
Unemployed (18 - 60 years old)						
Part-time employed (18 - 60 years old)						
Gender: Female				0.397		0.068
Driver license holder	1.24		0.969	0.341		0.379
Bicycle ownership		0.416		0.214	0.512	-0.483
Cars per household: 1			0.801			
Cars per household: 2			0.732			0.11
Cars per household: 3 or more			0.771			0.283
Number of mobile days for work			-0.092	-0.111		-0.152
Number of mobile days for education						
Number of mobile days for accompany				0.29	0.072	0.068
Number of mobile days for shop					0.165	-0.064

Table 2: Count-state model estimation

Variable	Work	Education	Accompany	Shop	Recreation	Other
Highly agglomerated areas			0.186	-0.284	-0.428	0.432
Urbanized areas			-0.211	-0.532	-0.455	0.329
Lower density urban/higher density rural				-0.484	-0.39	0.554
Household economic status: very low				-0.108	-0.518	
Household economic status: low					-0.231	
Household economic status: very high						0.233
Household size 2			0.382	-0.256	-0.143	
Household size 3	-0.237	0.499	0.692	-0.47	-0.212	
Household size 4	-0.366	0.585	0.692	-0.464	-0.112	
Household size 5 or more	-0.692	0.443	0.692	-0.394	-0.112	
Children per household: 1	0.243		0.888			
Children per household: 2	0.566		1.234			
Children per household: 3 or more	0.935		1.643			-0.645
Adults per household: 1	-0.763					
Adults per household: 2	-0.88	-0.302				
Adults per household: 3	-0.567	-0.547				
Adults per household: 4	-0.418	-0.547				
Between 10 und 17 years old	-2.532	3.849	-0.757	-1.372	0.419	-0.526
Between 18 und 25 years old			-0.757	-0.356	0.519	-0.411
Between 36 und 50 years old			-0.602			
Between 51 und 60 years old	-0.201	-2.325	-0.816			0.147
Over 61 years old	-2.092	-1.896	-0.816	-0.265		0.178
Occupation: Student			-0.265	-0.243	0.344	-0.321
Occupation: Part-time employed			0.507	0.551	0.347	0.136
Occupation: Unemployed			0.417		0.516	
Mobility Restriction	-0.189					
Employed with mobility restriction					-0.291	
Unemployed without mobility restriction	-1.39			0.767		
Unemployed with mobility restriction	-1.864			0.561		
Occupation: Unemployed			0.417		0.516	
Student (18 - 60 years old)	-2.253	2.49				
Unemployed (18 - 60 years old)	-0.912	1.413				
Part-time employed (18 - 60 years old)	-1.028					
Gender: Female	-0.116		0.205	0.071		
Driver license holder			0.956	0.199		0.242
Bicycle ownership	-0.225		-0.172	0.181	0.453	
Cars per household: 1						
Cars per household: 2						
Cars per household: 3 or more						
Number of mobile days for work			-0.047	-0.133	-0.057	-0.172
Number of mobile days for education						
Number of mobile days for accompany				0.103		0.068
Number of mobile days for shop					0.069	0.058
Number of mobile days for recreation						-0.037
1 2	-4.463	0.168	1.235	-1.591	-1.098	-0.07
2 3	-3.451	1.197	2.259	-0.325	0.003	1.176
3 4	-2.582	1.926	2.951	0.69	0.839	2.188
4 5	-1.574	3.06	3.599	1.728	1.695	3.119
5 6	1.371	7.981	5.263	2.924	2.588	4.324
6 7	3.331	10.609	7.686	4.954	3.734	5.789

Table 2 shows the estimates of the count-state model. The ordered logit model provides the probability of an individual to have between one and seven mobile days. Similarly, occupation status and mobility restriction play a role on how many days are mobile. Half-time employed allocate more days for all discretionary activities. Specifically, students tend to travel fewer days for shop, accompany and other, but they do travel more days for recreation activities. As expected, unemployed persons do shop more days per week than full-time employed and commute fewer days per week. Having MR does accentuate the differences even more, with fewer commute days per week and fewer shop days per week than persons without MR.

As for the likelihood of being mobile, an increased number of mobile days for mandatory activities decreased the number of mobile days for discretionary activities. Therefore, increasing the number of mobile days for mandatory activities detracts time from discretionary activities. Regarding number of mobile days for shop, it could be observed that higher mobility for accompany activities do increase the number of mobile days for shop, as well as being female, have driver license or own a bicycle. On the other hand, persons younger than 25 years old or over 70 do travel fewer days for shop.

4. CONCLUSIONS

This study presented an analysis on how mobile individuals are, being defined as the number of days per week that they perform activities out-of-home. The analysis has been focused on individuals with mobility restrictions, as they are usually reported as being less mobile. Two main methodological contributions included the analysis of a week-long travel diary as well as the statistical modeling of number of mobile days using a combination of a binomial logit model for the zero-state and an ordered logit model to model the count process state.

The results of this paper provide a better understanding of the individuals who are not mobile across a complete week for a certain purpose. Occupation status and mobility restrictions were key factors to determine how mobile individuals are, as well as the number of mobile days for higher hierarchy activities. Part-time employed individuals allocate more days to accompany acts, as well as shop, recreation and other, compared to full-time employed; and have lower likelihood of not being mobile. Unemployed individuals with mobility restriction do have higher likelihood of not being mobile for work, shop and recreation, and, if they are mobile, they tend to perform the activities in fewer days. On the other hand, unemployed individuals without mobility restriction have higher likelihood of being mobile and also conduct activities in more days per week than employed persons.

This research only scratches the surface on travel behavior of individuals with mobility restrictions. Future research will include the analysis of weekly variability, number of tours per day, distance per tour or mode choice. Furthermore, the data did not allow for distinctions among types of mobility restriction. It would be interesting to evaluate whether the type of disability (visual, cognitive, etc.) does play a role on their activity generation pattern. This may shed more light to uncover the motivations of performing fewer out-of-home activities and whether being less mobile does affect the overall individual well-being. In a broader future research, a survey would be carried out to identify the needs of individuals with mobility restrictions, by type of restriction, and whether their needs are met with their travel behavior.

REFERENCES

- Bekiaris, E.; Loukea, M.; Spanidis, P.; Ewing, S.; Denninghaus, M.; Ambrose, I. et al. (2018): Research for TRAN Committee: Transport and tourism for persons with disabilities and persons with reduced mobility. Brussels: European Parliament, Policy Department for Structural and Cohesion Policies.
- Best, Krista L.; Noreau, Luc; Gagnon, Marie-Andrée; Barthod, Corentin; Hitzig, Sander L.; Routhier, François (2022): Housing, Transportation and Quality of Life among People with Mobility Limitations: A Critical Review of Relationships and Issues Related to Access to Home-

- and Community-Based Services. In *Disabilities* 2 (2), pp. 204–218. DOI: 10.3390/disabilities2020015.
- Cervero, Robert; Kockelman, Kara (1997): Travel demand and the 3Ds: Density, diversity, and design. In *Transportation Research Part D: Transport and Environment* 2 (3), pp. 199–219. DOI: 10.1016/S1361-9209(97)00009-6.
- Frye, Ann (2019): Disability Inclusive Public Transport. In *High Volume Transport Applied Research*.
- Hilgert, Tim; Kagerbauer, Martin; Vortisch, Peter (2018): Considering joint trips and activities in week activity schedules. In : TRB 97th Annual Meeting Compendium of Papers. Ed.: Transportation Research Board, Art.-Nr.: 18-01378.
- Jansuwan, Sarawut; Christensen, Keith M.; Chen, Anthony (2013): Assessing the Transportation Needs of Low-Mobility Individuals: Case Study of a Small Urban Community in Utah. In *Journal of Urban Planning and Development* 139 (2), pp. 104–114. DOI: 10.1061/(asce)up.1943-5444.0000142.
- Litman, Todd (2022): Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transportation Planning. In *Victoria Transport Policy Institute* 8 (2), pp. 50–65.
- Lucas, Karen; Bates, John; Moore, José; Carrasco, Juan Antonio (2016): Modelling the relationship between travel behaviours and social disadvantage. In *Transportation Research Part A: Policy and Practice* 85 (October 2021), pp. 157–173. DOI: 10.1016/j.tra.2016.01.008.
- Nobis, Claudia; Kuhnimhof, Tobias (2018): Mobilität in Deutschland - MiD Ergebnisbericht. Studie von Infas, DLR, IVT und infas 360 im Auftrag des Bundesministers für Verkehr und digitale Infrastruktur (FE-Nr. 70.904/15). Bonn, Berlin. Available online at www.mobilitaet-in-deutschland.de.
- Park, Keunhyun; Esfahani, Hossein Nasr; Novack, Valerie Long; Sheen, Jeff; Hadayeghi, Hooman; Song, Ziqi; Christensen, Keith (2022): Impacts of disability on daily travel behaviour: A systematic review. In *Transport Reviews*, pp. 1–26. DOI: 10.1080/01441647.2022.2060371.
- Ripley, Brian; Venables, Bill; Bates, Douglas M.; Hornik, Kurt; Gebhardt, Albrecht; Firth, David (2023): The MASS package. Available online at <https://cran.r-project.org/web/packages/MASS/MASS.pdf>.
- Rosenkvist, Jenny; Risser, Ralf; Iwarsson, Susanne; Wendel, Kerstin; Ståhl, Agneta (2009): The Challenge of Using Public Transport: Descriptions by People with Cognitive Functional Limitations. In *Journal of Transport and Land Use* 2 (1), pp. 65–80. DOI: 10.5198/jtlu.v2i1.97.
- Schmöcker, Jan Dirk; Quddus, Mohammed A.; Noland, Robert B.; Bell, Michael G.H. (2008): Mode choice of older and disabled people: a case study of shopping trips in London. In *Journal of Transport Geography* 16 (4), pp. 257–267. DOI: 10.1016/j.jtrangeo.2007.07.002.
- Seo, Sang-Eon; Ohmori, Nobuaki; Harata, Noboru (2013): Effects of household structure and accessibility on travel. In *Transportation* 40 (4), pp. 847–865. DOI: 10.1007/s11116-013-9468-3.
- Statistisches Bundesamt, Wirtschaft und Statistik (2012): Lebenslagen der behinderten Menschen. Available online at https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2012/03/lebenslagen-behinderte-032012.pdf?__blob=publicationFile.
- United Nations (2006): Convention on the Rights of Persons with Disabilities.

Vallée, Jan; Ecke, Lisa; Chlond, Bastian; Vortisch, Peter (2022): Deutsches Mobilitätspanel (MOP) – Wissenschaftliche Begleitung und Auswertungen Bericht 2021/2022: Alltagsmobilität und Fahrleistung.

Vickerman, R. W. (1974): Accessibility, Attraction, and Potential: A Review of Some Concepts and Their Use in Determining Mobility. In *Environ Plan A* 6 (6), pp. 675–691. DOI: 10.1068/a060675.

Washington, Simon; Karlaftis, Matthew G.; Mannering, Fred L.; Anastasopoulos, Panagiotis Ch (2020): Statistical and econometric methods for transportation data analysis. Third edition / Simon Washington, Matthew G. Karlaftis, Fred Mannering, Panagiotis Anastasopoulos. Boca Raton: Chapman & Hall/CRC (Chapman & Hall/CRC interdisciplinary statistics series).

World Bank (2013): Improving accessibility to transport for people with limited mobility (PLM): A practical guidance note. In 2012, p. 91.

World Health Organization (2011): World report on disability. WHO Press, World Health Organization. Geneva.

Yun, Dae-Sic; O'Kelly, M. E. (1997): Modeling the day-of-the-week shopping activity and travel patterns. In *Socio-Economic Planning Sciences* 31 (4), pp. 307–319. DOI: 10.1016/S0038-0121(97)00010-4.

Zeileis, Achim; Kleiber, Christian; Jackman, Simon (2008): Regression Models for Count Data in R. In *J. Stat. Soft.* 27 (8). DOI: 10.18637/jss.v027.i08.

Prediction of Passengers Demand for Customized Bus Systems

Saumya Bhatnagar^{*1}, Rongge Guo¹, Jihui Ma², and Mauro Vallati¹

¹School of Computing and Engineering, University of Huddersfield, United Kingdom

²School of Traffic and Transportation, Beijing Jiaotong University, China

SHORT SUMMARY

The customized bus system is an innovative demand-responsive public transit service with the potential to significantly alleviate congestion and environmental footprint. To fully exploit the flexibility of this approach, it is pivotal to forecast the demand for the service, in order to optimize the use of vehicles and resources. In this paper, with the aim for supporting the use of customized bus systems, we formalize the predictive task and assess the performance of a range of machine learning techniques. We introduce a two-step predictive task aiming at (i) identifying the presence of demand and, if there is actual demand, (ii) estimating the number of passengers to be served. The experimental analysis, based on realistic data from the Beijing area, shed some light into the performance of different classes of approaches.

Keywords: Customized Bus System, Machine Learning, Artificial Intelligence.

1 INTRODUCTION

The customized Bus (CB) system is an innovative and extremely flexible public transit (PT) service with the benefits of congestion alleviation and environmental friendliness (Shu & Li, 2022), that is emerging as an alternative to conventional buses and private cars. CB is a class of demand-oriented transit that holds the promise to provide "door-to-door" service to passengers with similar travel requirements in both space and time (Liu & Ceder, 2015).

The vast majority of existing work on CB systems assumes that travel demands are collected in advance from dedicated online booking platforms. The CB system can then produce plans including routes and schedules to satisfy the time- and space-restricted requests accordingly (Huang et al., 2020). In literature, a class of approaches allows to generate optimal solutions for these static demands, but at the cost of reduced flexibility for the service, as dynamic travel requests are ignored and not handled during operation. To tackle this pivotal issue, that reduces the usability of the service, a growing number of approaches are exploring the idea of a two-stage optimization procedure for on-demand CB service, as shown in Fig. 1. The first stage is similar to the traditional CB service design and generates the initial plans (including routes, timetables and schedules) for static requests. Each request contains a paired origin-destination (OD) and preferred departure time. The second stage responds to dynamic requests received by adjusting the initial plans during operation (Wang et al., 2020). However, even this class of approaches is not fully exploiting the flexibility of the CB system. It may be impractical to satisfy all the dynamic requests, that also provide strict constraints on preferred time windows for pickup and drop off, starting from the generated static plan.

During the operation, the dynamic stage is triggered once new requests pop up. The re-planning process activities are implemented to insert each emerging request into the current network. If existing CB routes are able to fully satisfy the spatial-temporal constraints of the request, the service is updated and followed by drivers via real-time communication; otherwise, the request is ignored by the system and defined as unserved. The main reason behind this is that the short response time makes it impossible to rapidly input new requests into the current network while fulfilling travel restrictions of both already assigned and new requests. Therefore, the prediction of the dynamic requests is essential for planning to prevent the unserved requests and minimize uncertainty in operation of the dynamic stage, which can improve the operational efficiency of CBs.

To fully reap the benefits of CB systems, it is crucial to have in advance an estimate of the dynamic requests, to generate plans that are robust with regards to dynamic changes and adjustments. Having approaches that can accurately estimate future passenger demands would help ensuring

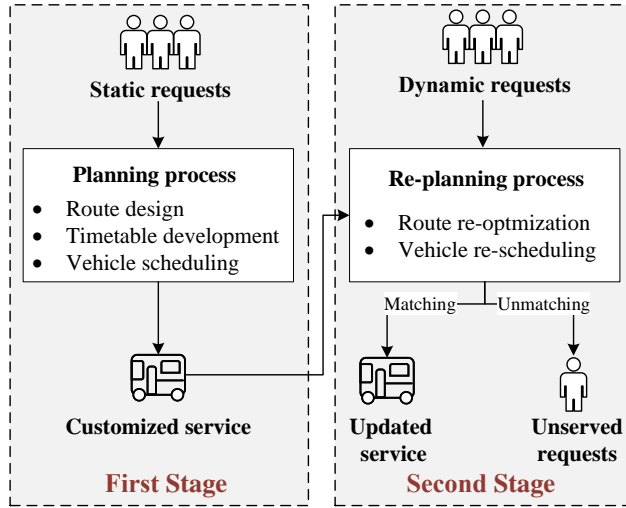


Figure 1: The planning process of the on-demand customized bus service.

Table 1: Data structure of collected SCD information.

Field Name	Description
CardID	Card identification number
LineID	Bus line number
BusID	Vehicle identification number
UpTradeStation	Boarding station number
UpTradeTime	Boarding time
DownTradeStation	Alighting station number
DownTradeTime	Alighting time

a higher level of optimization in CB systems, as well as better service and higher environmental benefits. To this end, in this paper we formalize the predicting task, we describe best practice, and investigate the use of a range of machine learning algorithms. Our analysis is performed considering real-world smartcard data collected by the Automatic Fare Collection system in Beijing over 2 months, and the results demonstrate that machine learning is a promising approach to be used to support CB systems in improving their effectiveness. Further, the proposed analysis provides useful insights into how to preprocess data and to represent complex variables, that can help in fostering the use of machine learning techniques in similar applications.

2 METHODOLOGY

A quantitative case-study approach has been adopted to define the predictive task and determine the machine learning technique that best predicts passenger demands for a CB system.

Data Source

In this study we use real-world data extracted from the SCD collected by the Automatic Fare Collection (AFC) system in Beijing, China, which can provide a complete overview of the trips. The SCD contains more than 12 million transactions per day. Each transaction records details of the get on and get off location, and time for the trip information (see Table 1).

For this experimental analysis, we consider passenger spatial-temporal dynamics collected from December 1, 2018 to January 31, 2019. The considered period of time is before the start of the COVID-19 pandemic, so no restrictions were in place at the time. We focus only on working days, as travel requests of CB are expected to be mainly from commuters. For the same reason, we select records from typical residential and working areas during the morning peak hours (7:30-9:00). This is the time when the CB system is expected to be most under stress from dynamic travel demands. Starting from the raw SCD data, we follow the approach proposed by Guo et

al. (2019) to generate corresponding demands for a CB system. In particular, the implemented approach relies on three major stages, namely trip chain generation for non-transfer and transfer trips, station identification, and OD matrix generation. After that, travel demands for each OD and corresponding boarding and alighting timestamps for passenger groups are collected, and used as the input for the prediction system. The distribution of extracted travel demands is given in Fig. 2.

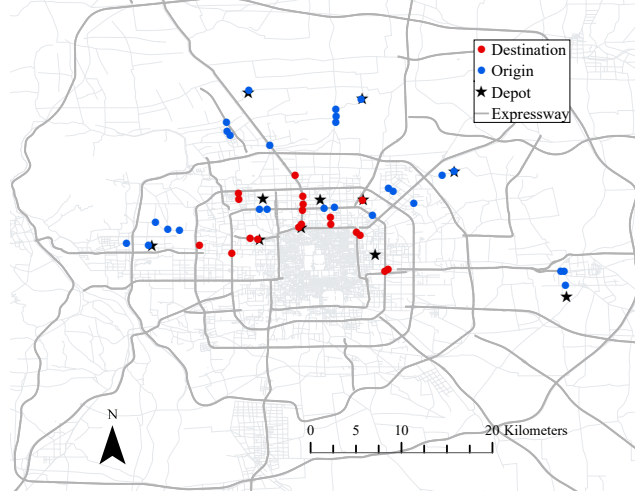


Figure 2: Distribution of extracted travel demands.

The data set includes 5 features, all alphanumeric, referred to as original features in Table 2.

Data Preprocessing and Transformation

To support the predictive task, the limited initial set of available features has been extended as shown in Table 2. From the original date, we extracted two additional features referring to the specific day of the week (Monday, Tuesday, etc.) and the indication of whether the day is a Monday, Friday, or any other day of the week. This is because this 3 classes of days can show different travel patterns. From the original time stamp, 5 time categories have been identified, namely 0700-0730, 0730-0800, 0800-0815, 0815-0830, 0830-0845 and converted to ordinal categorical values. The categorization of starting demand time decreases the complexity of the predicting element, while preserving the usefulness of the information.

The 2 features reporting the IDs of the origin and destination bus stops have been combined into a single variable by concatenating the corresponding values. In order to do that, all the possible combinations have been generated. This lead to some empty demands, that augmented the size of the data set, with a ratio of datapoints showing demand vs without demand to be in the imbalanced ratio of 8/23. Finally, the original numeric demand feature has been divided into 2 different features, namely the size of the passenger group for the demand, and a Boolean feature indicating the presence of the demand for a specific set of features (Origin-Destination pair, Date and Time). In the considered data set, the size of demands ranges between 0 (where there is no demand) and 30.

Table 2: Extended Features

Original features	Extracted features	Type
Date	Date	datetime format
	Week Day	Categorical
	DayType	Categorical
Time Stamp	Time period	Categorical
Origin Destination	Origin-Destination	Categorical
Demand	Size	Integer
	Presence	Boolean

We performed a correlation analysis of the extended variables, that confirmed that there is no

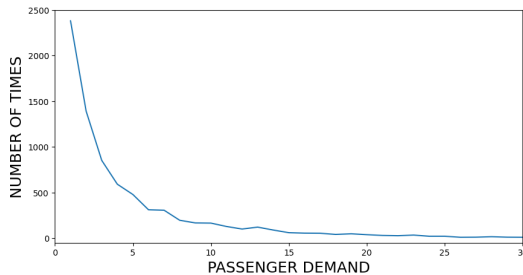


Figure 3: Demand size distribution.

correlation among them, hence the set is suitable to be used for training purposes.

Fig. 3 shows the distribution of the passenger groups sizes in the considered data set. Notably, smaller groups are very common, while larger groups – here the largest considered is of 30 passengers – are much less usual.

Predictive Task

Given the complexity of the predictive task at hand, we decided to define it as two different predicting tasks. First, a classification step that allows to identify the presence of demand (last row, Table 2), then there is a subsequent regression step that aims at predicting the size of the corresponding passengers group. This separation gives also us the opportunity to assess the abilities of a range of well-known machine learning approaches on two different yet crucial tasks.

The following approaches have been considered for this analysis: K-Nearest Neighbors (KNN)(Wu et al., 2008), Decision Tree, Random Forest (Breiman, 2001), AdaBoost (Schapire, 2013), LightGBM (Ogunleye & Wang, 2019), and Multi Layer Perceptron (MLP) (Haykin, 1994). The approaches have been selected based on the range of implemented techniques, and on their performance on well-known benchmarks.

As a baseline algorithm, to better contextualize the achieved results, we use a traditional Linear Regression approach, declined in its Logistic Regression form for the classification task.

Evaluation Metrics

For the binary classification problem of predicting the presence of demand, we relied on the well-known notions of true positive (negative), false positive (negative). We then consider as metrics for assessing the abilities of systems, *accuracy*, *F1 score*, and *Area Under the Curve* (AUC).

When it comes to the metrics to assess the capabilities of the regression models, those selected are *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), and *Root Mean Squared Error* (RMSE). MAE measures the average of the absolute residuals (the difference between the actual value and the predicted value) in the dataset. MSE measures the variance of the residuals and the RMSE measures the standard deviation of those residuals. For those metrics, the lower the value, the better the performance of the regression model.

Training Settings

For training, testing, and validating purposes we used sklearn (Pedregosa et al., 2011). sklearn is a robust and well-known Python library for performing ML tasks. It natively supports hyperparameter optimization (Feurer & Hutter, 2019). In the following, we shortly summarize the parameters considered for the optimization of the learning algorithms.

- KNN: The value of K, i.e. the number of nearest neighbors to be considered.
- AdaBoost: we optimized 3 aspects: the base estimators, the number of estimators in the ensemble, and the learning rate. The base estimators were given three choices between Support Vector Machine, Logistic Regression, and Decision Trees.
- LightGBM: the 4 parameters to be optimized include the boosting type (GBDT or DART), the number of maximum leaves in each learner, the boosting learning rate, and the number of boosted trees.
- MLP: we optimized 5 parameters, namely the set of hidden layer sizes, activation functions, solvers, alphas, and learning rates.

Table 3: Classification Results Summary

Classifiers	Accuracy	F1	AUC
Logistic Regression	73.81	2.42	58.75
KNN	82.65	49.15	80.52
MLPClassifier	79.49	45.95	79.88
Decision Tree	78.14	39.50	75.98
AdaBoost	75.88	15.89	76.00
Random Forest	72.59	71.81	82.67
LightGBM - DART	87.17	77.85	94.65

Ensemble Models		RF		DART	
		Predicted			
		Positive	Negative	Positive	Negative
TRUE	Positive	2206	962	1425	193
	Negative	770	2382	618	4084

Figure 4: Confusion matrices of Random Forest and LightGBM - DART.

The learning rate ranges for all the aforementioned hyper-parameter optimizations, have been kept as floating values ranging between 0.05 to 1.

Finally, we also optimized the best Logistic Regression algorithm, selecting among those implemented by the sklearn library.

To reduce noise, random seeds have been fixed.

For the classification task, the hyper-parameters optimization aims at improving two metrics, namely F1 and ROC-AUC. Ties are broken in favor of AUC. For regression, the focus is on minimizing the Mean Absolute Error (MAE).

3 RESULTS

In this section we present the results of our extensive experimental analysis. Some preliminary tests, not shown in this paper, performed by considering the single straightforward regression task of predicting at the same time both presence and size of demand lead to models not usable in practice, due to the extreme error. Therefore, in this section we consider the two steps predictions discussed in Methodology. First, we focus on the classification step that aims at predicting the likeliness of travel demand between two stops, and second we consider the regression task of predicting how many passengers will be part of the travel group.

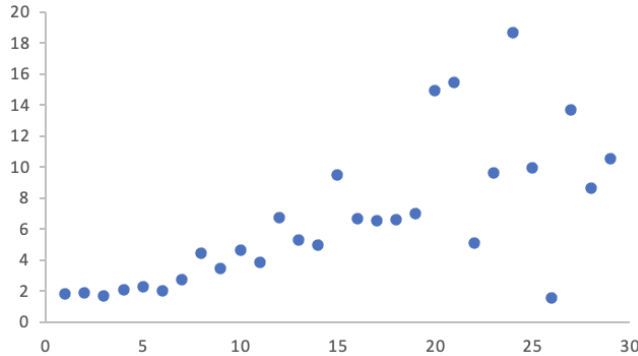
Table 3 shows the performance of the considered algorithms on the binary classification task. The results indicate that, unsurprisingly, Random Forest is the algorithm that shows the worst accuracy; it has been outperformed even by Logistic Regression and Decision Tree. This is due to the limited number of available features, even in the extended set, that is a known element that can reduce the performance of this class of approaches. On the other hand, Random Forest should be the second best choices when evaluating on F1 score and AUC, demonstrating that the unbalanced nature of the data is not strongly negatively affecting the generated models. The optimized LightGBM using DART is delivering outstanding performance according to all metrics; however, we observed that with almost all other hyper-parameter configurations of LightGBM showed significantly worse results (not shown in the table). Finally, it is worth noticing the very low F1 score and AUC of the Logistic Regression approach, suggesting that the approach does not cope well with the unbalanced data set.

To shed some light into the relative performance of Random Forest and LightGBM, Figure 4 shows the corresponding confusion matrices. Random Forest (RF) is better in predicting positive cases, but shows low performance when predicting cases for which no demand is expected.

We now turn our attention to the regression task. Results are presented in Table 4. As a first remark, we do not include KNN results because, due to the way KNN models are generated, it quickly runs out of memory for large data sets – hence providing predictions of very low quality.

Table 4: Regression Results Summary

Regressors	MAE	MSE	RMSE
Linear Regression	1.87	13.49	3.67
MLPRegressor	1.92	13.69	3.70
Decision Tree	1.10	9.30	3.04
AdaBoost	3.62	23.54	4.85
Random Forest	0.82	4.05	2.01
LightGBM - DART	0.88	3.62	1.90

**Figure 5:** Mean absolute error (y-axis) for each actual demand value (x-axis) when LightGBM is used for the regression task.

With regards to the other considered algorithms, LightGBM is again providing extremely good performance, but in terms of MAE Random Forest is, surprisingly, the approach that shows best results. AdaBoost is instead the worst one, that produces predictions that are not usable in practice for optimizing the routes of CB systems.

Figure 5 shows how the MAE varies according to the actual demand to be predicted, when the best predictor is used. As expected, for ranges of demand that are extremely well represented in the data set, the error is very limited. When demands get larger and less common, the predictions show a higher degree of variability. Pragmatically, this means that the predictions will be more precise for common demands – and the graph shown already provides some information about the reliability of predictions made for different demand sizes.

Finally, a general result can be derived by comparing Random Forest and AdaBoost results on both classification and regression tasks that using Bagging ensembles over Boosting ensembles lead to better results in regression, but the contrary is true in terms of classification. While this behavior may not generalize on different data sets, it can provide an interesting take-home message for this class of applications.

4 CONCLUSION

With the aim of fostering the use and efficiency of customized bus systems, in this paper we tackled the task of forecasting travel demands to be served. Our analysis provided insights into how data can be processed to provide high quality input for machine learning algorithms, and we showed how a challenging task can be divided into more amenable tasks that are easier to be dealt with. On this regards, we demonstrated the abilities of a range of predictive techniques on both classification and regression tasks, using realistic data from the Beijing area. Our results indicate that machine learning provides suitable approaches to tackle the challenges of CB systems due to the limited ability to forecast ongoing demand.

Future work will focus on extending the analysis to different metropolitan areas, and to increase the considered features to include aspects such as weather, traffic, etc.

ACKNOWLEDGEMENTS

This work is supported by the UKRI Future Leaders Fellowship [grant number MR/T041196/1].

REFERENCES

- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Feurer, M., & Hutter, F. (2019). Hyperparameter optimization. In F. Hutter, L. Kotthoff, & J. Vanschoren (Eds.), *Automated machine learning: Methods, systems, challenges* (pp. 3–33). Springer International Publishing.
- Guo, R., Guan, W., Huang, A., & Zhang, W. (2019). Exploring potential travel demand of customized bus using smartcard data. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)* (p. 2645–2650). doi: 10.1109/ITSC.2019.8916843
- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.
- Huang, D., Gu, Y., Wang, S., Liu, Z., & Zhang, W. (2020). A two-phase optimization model for the demand-responsive customized bus network design. *Transportation Research Part C: Emerging Technologies*, 111, 1–21.
- Liu, T., & Ceder, A. A. (2015). Analysis of a new public-transport-service concept: Customized bus in china. *Transport Policy*, 39, 63–76.
- Ogunleye, A., & Wang, Q.-G. (2019). Xgboost model for chronic kidney disease diagnosis. *IEEE/ACM transactions on computational biology and bioinformatics*, 17(6), 2131–2140.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Schapire, R. E. (2013). Explaining adaboost. In *Empirical inference* (pp. 37–52). Springer.
- Shu, W., & Li, Y. (2022). A novel demand-responsive customized bus based on improved ant colony optimization and clustering algorithms. *IEEE Transactions on Intelligent Transportation Systems*.
- Wang, C., Ma, C., & Xu, X. D. (2020). Multi-objective optimization of real-time customized bus routes based on two-stage method. *Physica A: Statistical Mechanics and its Applications*, 537, 122774.
- Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... others (2008). Top 10 algorithms in data mining. *Knowledge and information systems*, 14(1), 1–37.

City-wide bottleneck and deficiency analysis on a road network generated from the Open Street Map road network using Floating Car Data (FCD)

Josephine Grau^{*1}, Lea Fuchs¹, Torben Lelke², and Peter Vortisch³

¹M.Sc., Institute for Transport Studies, Karlsruhe Institute of Technology, Germany

²M.Sc., Institute of Transportation and Urban Engineering, Technische Universität Braunschweig, Germany

³Prof. Dr.-Ing., Institute for Transport Studies, Karlsruhe Institute of Technology, Germany

SHORT SUMMARY

The German Highway Capacity Manual (HBS) and the German Guideline for Integrated Network Planning (RIN) rely on travel time distributions to assess the Level-of-Service (LOS) of roads and road networks. Usually, these values are generated by traffic measurements or with the help of traffic flow simulations. In recent years Floating Car Data (FCD) has become an essential data source for analyzing traffic quality because of its easy accessibility and growing coverage. This paper proposes a method to perform a city-wide analysis on the Open Street Map (OSM) road network using raw FCD. Therefore, OSM road segments of a city are aggregated to longer network sections on which travel times are estimated. Performance indicators can be calculated using these travel time distributions. Examples are shown for the cities of Karlsruhe and Hannover.

Keywords: floating car data, incident analysis, open data, key performance indicators, big data analytics

1 INTRODUCTION

The German Highway Capacity Manual (HBS) is the primary tool used to evaluate the quality of traffic flow. However, the calculation methods of the HBS are not generally applicable, e.g., for roads with a speed limit of 30 km/h or if cyclists are guided in mixed traffic (HBS (2015)). Since the required conditions are often not met in reality, especially in inner-city areas, it is necessary to use different means for the evaluation. This paper will present such an alternative by evaluating FCD.

FCD offers the advantage of mapping historically long periods of data obtained from vehicles moving in traffic. It can add to classical stationary detectors by providing information about driven trajectories. Other than stationary detectors, only a subset of the total traffic volume is represented. However, preceding research shows that even low-frequency FCD (around four data transmissions per minute) has adequate statistical power to calculate travel time distributions (Schäfer et al. (2003), Jenelius & Koutsopoulos (2013)). Furthermore, FCD can be used to analyze traffic flow regarding incident recognition. Altintasi et al. (2017) showed that FCD could recognize traffic patterns defined by LOS and detect congestions and bottleneck locations.

For an intuitive analysis of a city-wide road network, travel time indices are helpful. The TomTom traffic index, for example, uses FCD to generate several indicators, e.g., an average travel time per 10 km and offers thus the possibility to compare different cities (*TomTom Traffic Index* (2023)). Further studies developed and evaluated several indices for motorways (see, for example, Peter et al. (2021), and Radde et al. (2016)) and for urban road networks (Ulm et al. (2015)).

Similar to the study conducted by Axer & Friedrich (2014), we use historical FCD, transformed to trajectories, to calculate average travel times on all road network links resulting in traffic quality indices using the traveled distance and the elapsed time between two data transmissions. In Axer & Friedrich (2014), FCD was mapped to the OSM road network via the TMC (traffic message channel) system, used as an intermediary georeferencing tool. We use a more straightforward approach and map the FCD directly to the OSM road network. This offers an analysis method using FCD for a whole city road network. The determination of decision values (travel time distributions and

percentiles) within the framework of the German guidelines of the HBS and the RIN is possible. The aim is to provide practitioners with a time and cost-efficient method to evaluate the quality of every single link in an urban network and anchor this method in the mentioned guidelines.

2 METHODOLOGY

In order to perform a city-wide analysis, the following method is proposed. First, a suitable road network needs to be constructed. Our work is based on the OSM road network, as it is publicly available. Roads in the OSM network consist of so-called OSM ways, each of which is a list of nodes representing a small road segment of some meters. Inspired by the functional road classes defined in the German Guideline for Integrated Network Planning (RIN (2008)), these OSM ways are grouped into longer road segments representing, for example, main roads, which we refer to as network sections in the following. The FCD can be used to estimate a travel time or speed distribution for each network section. Finally, performance indicators are calculated to evaluate the traffic conditions on each network section. Network sections with low performance measures can be analyzed in greater detail in further steps.

The OSM import is done via the Python OSM API packages *osmapi* (2023) and *overpass* (2019). As they are decisive for travel planning, only higher road classes are considered. Therefore, e.g., residential roads are not taken into account. The road classes are defined by the OSM attribute "highway". Besides that, OSM offers many other attributes for the OSM ways. For this application, we use the id of the OSM way ("id"), the allowed maximal speed on the OSM way ("maxspeed"), whether the OSM way is a oneway road ("oneway"), and the coordinates of the nodes ("geometry.coordinates"). The information is then augmented by node ids, start and end nodes of the OSM ways, and a doubling of non-oneway OSM ways to separate the two driving directions.

After this preparation, the automatic identification of network sections is made. The general idea is to add an OSM way to a network section as long as it belongs to the same road category and does not cross an OSM way of the same or higher-order road category. In the end, each direction of an OSM way is part of precisely one network section. The algorithm works as follows: at the beginning, there are no network sections, and no links of the network are assigned to a network section. Each algorithm step randomly selects a link that must still be assigned to a network section. This link establishes a new network section from which the section is extended in and against the direction of travel until a node is reached where routes of the same or higher category cross. This determines and delimits the network section, and the procedure continues with the next unassigned link. Directions of travel are considered separately, so OSM ways representing two directions of travel must be duplicated and used as two separate links. After all links are assigned, the procedure is finished. For the choice of the next link to be attached to a network section, the start and end nodes and angles are considered, i.e., OSM ways are not attached to a network section if a steep curve occurs at the current node. This reflects typical situations in the network structure and avoids, e.g., U-turns. Particular attention should be paid to avoid round trips; thus, it has to be checked whether the following link is already part of the considered network section.

This method provides plausible and intuitive results in the OSM road networks of the considered cities. The algorithm depends on a correct assignment of coordinates and attributes in OSM, which is mostly the case. The following figures provide some insights into the algorithm's results.

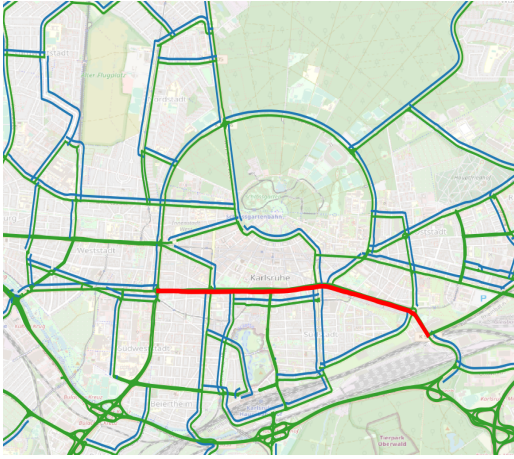


Figure 1: Recognition of the network section Kriegsstraße in the city center of Karlsruhe

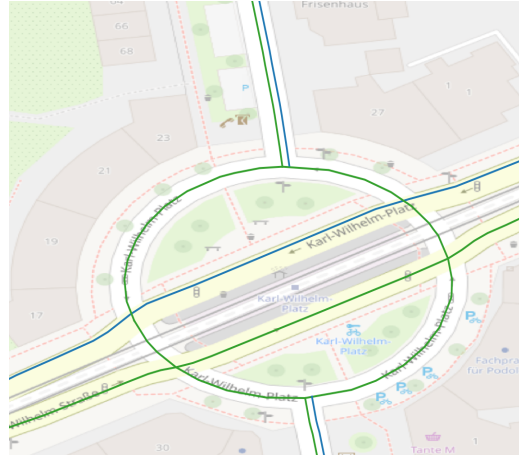


Figure 2: Recognition of a complex roundabout with separation of driving directions

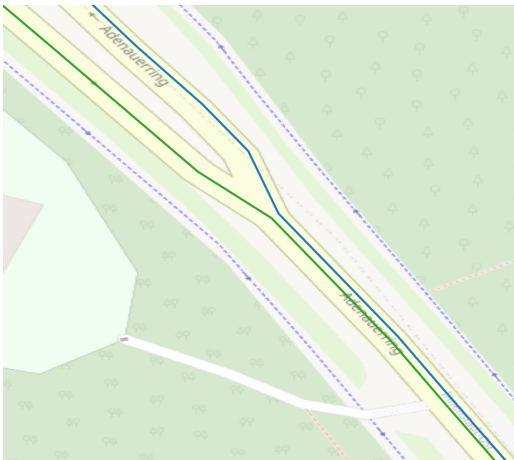


Figure 3: Continuation of the driving directions on roads with structural separation of the driving directions



Figure 4: Faulty detection of two driving directions due to a wrongly assigned oneway attribute in OSM

The following step uses FCD to calculate travel times on the network sections. To obtain a reliable evaluation, the FCD must be processed. First, we use only the weekdays Tuesday, Wednesday, and Thursday to compare similar demand situations that represent typical working days. The data quality is ensured by excluding implausible high speed values and cutting trajectories at stops longer than 60 seconds.

For each network section, each direction's start and end nodes are used to define a catch radius to filter floating cars traversing the entire section. This imitates the principle of an ANPR measurement. Travel times were calculated from the difference between the time stamps. A travel speed distribution can be determined using the haversine function to calculate road lengths. The percentiles of these distributions are used to calculate incident indices. For this step, raw FCD compared to aggregated FCD information is essential. Calculating travel time (or travel speed) distributions of a network section from shorter road segments is impossible. Mathematically, the random variables representing travel times on different road segments are not independently distributed. Therefore, a convolution of travel time distributions to obtain overall distributions is incorrect.

Next, several indicators are used to evaluate traffic flow on the network sections. The reliability index, as proposed in Peter et al. (2021), is calculated as the quotient of the 90% and 50% percentile of travel time denoted by t_{90} and t_{50} accordingly:

$$Reliability\ Index = \frac{t_{90}}{t_{50}}. \quad (1)$$

If the travel times vary little, there are hardly any disturbances (t_{90} and t_{50} are low) or permanent disturbances (t_{90} and t_{50} are high) and the reliability index is close to one. The higher the reliability index, the more the travel times fluctuate on the network section. In particular, the reliability index does not indicate whether there are disruptions but whether the same traffic condition occurs reliably. The reliability index has the same significance as the buffer time index (BTI) used in Radde et al. (2016) normalized with the 50% percentile. It can be interpreted as additional time to the median travel time to be considered in planning to arrive on time. The index is calculated by the difference between the 90% and 50% percentiles and a normalization for comparability:

$$BTI = \frac{t_{90} - t_{50}}{t_{50}} = \frac{t_{90}}{t_{50}} - 1 = Reliability\ Index - 1. \quad (2)$$

Another index proposed by Peter et al. (2021) is the travel time index which evaluates the ratio between the 50% percentile of the travel time and a target travel time t_{target} . We use the travel time when driving at the maximum permitted speed to approximate t_{target} . This is a plausible assumption for a city road network of higher-order road categories.

$$Travel\ time\ index = \frac{t_{50}}{t_{target}} \quad (3)$$

The above indices refer to travel time percentiles, which are robust quantities. However, an aggregation of travel time distributions over all daytimes only gives a rough overview and does not account for peak hours. On the other hand, indices with higher temporal or spacial resolutions can only be used if a sufficiently large amount of data is available. This would apply when evaluating the indices above for daytime hours. Another index of that category is the Cumulated Differences Index (CDI) from Radde et al. (2016):

$$CDI = (n - 1)^{-1} \sum_{t=1}^{n-1} |(v_{t+1} - v_t)|, \quad (4)$$

where n denotes a considered number of days and v_t denotes the mean speed at the considered hour on day t . This index reflects user expectations based on the fact that travel speeds do not vary from day to day. A high CDI indicates the low reliability of the route segment. The index provides good results in the situation considered by Radde et al. (2016) on highways with high traffic volumes. In contrast, it may be considered unsuitable for urban roads with few vehicles per hour at low FCD penetration rates. Indicators that require many data sets due to high temporal or spatial resolution may be of limited suitability for FCD evaluation.

3 RESULTS AND DISCUSSION

The method explained above is used for the cities of Karlsruhe and Hannover in Germany. For Karlsruhe, we used FCD from September 2021, whereas for Hannover, data were available for the entire year 2021. As mentioned above, only Tuesdays, Wednesdays, and Thursdays were considered. The transmission interval of the FCD is mostly around five seconds.

The following pictures show the results for the reliability index (formula 1) and the travel time index (formula 3) for the city of Karlsruhe. It can be seen that the two indices make different statements but lead to plausible results. They allow the identification of network sections with poor traffic conditions, which can be subjected to further, more detailed analyses. For example,

one can take the Kriegsstraße, the east-west axis in the center of Karlsruhe, which performs poorly in both indices as expected. It was the subject of extensive reconstruction measures until 2022.

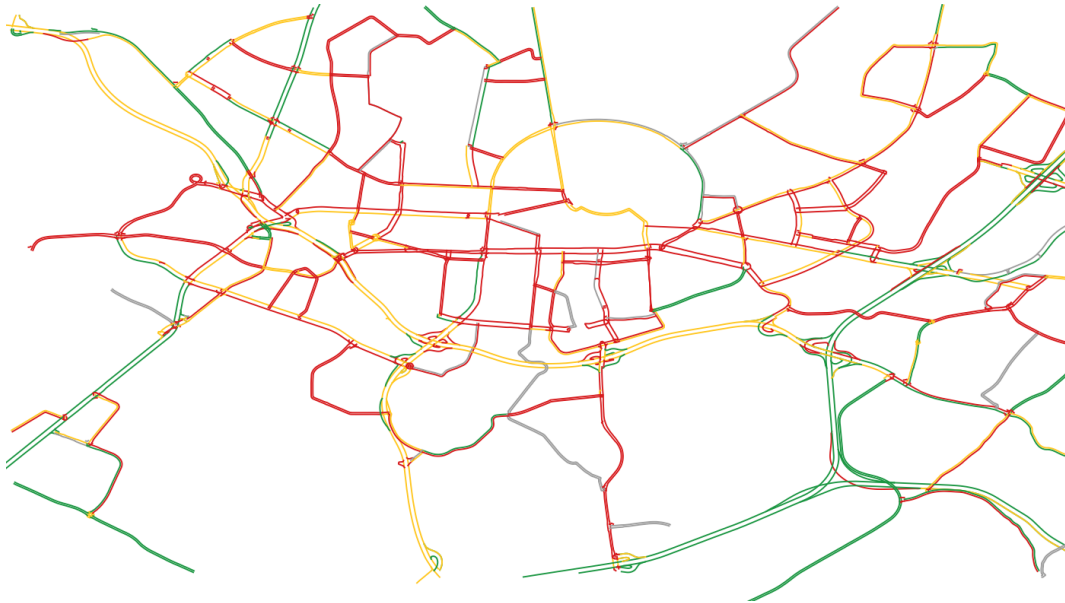


Figure 5: Reliability index Karlsruhe (green: value range 1 - 1.5, yellow: 1.5 - 2, red: > 2)

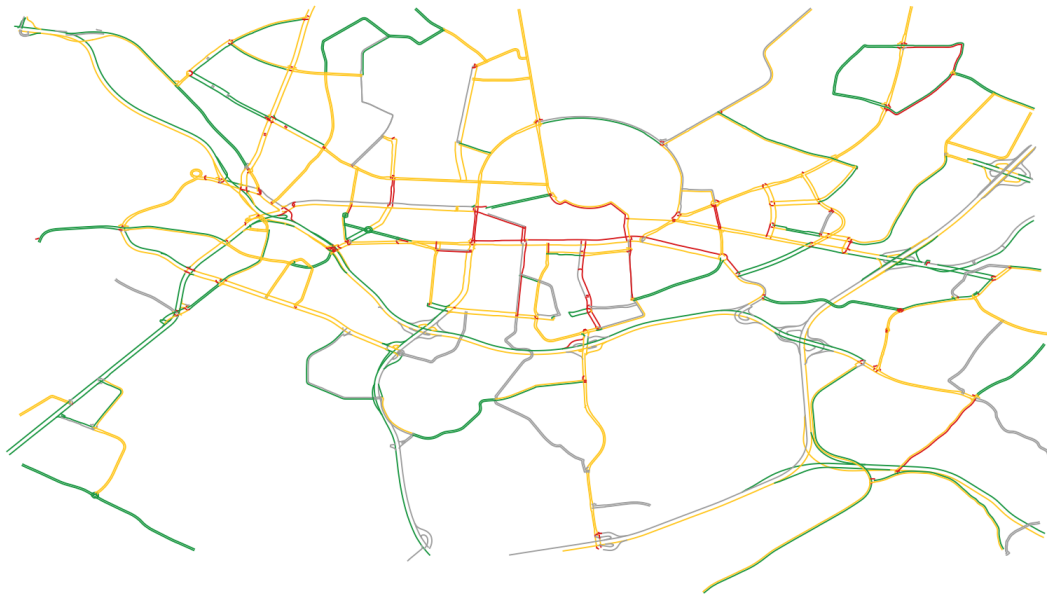


Figure 6: Travel time index Karlsruhe (green: value range 0 - 1, yellow: 1 - 2, red: > 2)

As we had enough data for Hannover, the CDI (formula 4) could be evaluated. The figures 7 and 8 show the results for 7 a.m. and 11 a.m. The results do not show a considerable variation between different evaluated hours.

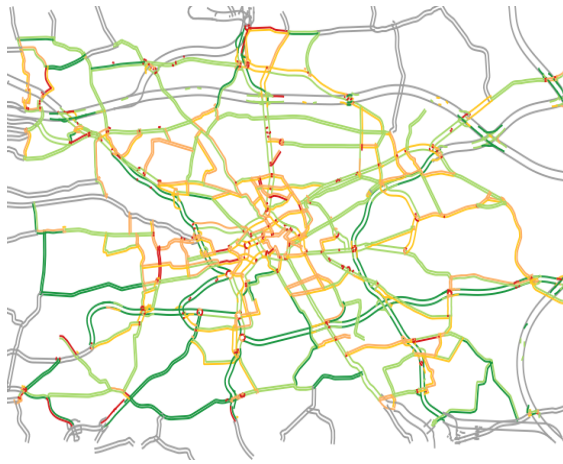


Figure 7: CDI Hannover 2021, 7 a.m.

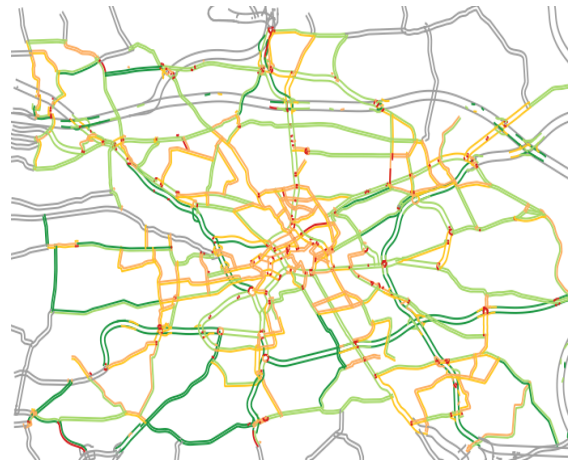


Figure 8: CDI Hannover 2021, 11 a.m.

These indices were chosen as examples from the literature. When interpreting the indices, it is important to ensure that they reflect the different calculation methods. The performance index must be chosen according to the intended application.

The indices rely on FCD's high penetration rates, transmission frequencies, and coverage of all road segments. For the reliability index and the travel time index, an average of 640 vehicles per network section were evaluated in the example of Karlsruhe. For the CDI, around 128 days with an average of 5.6 vehicles per hour were available as a data basis for each network section. Thus, the travel time distribution's representativeness for calculating the CDI is significantly less robust than the distribution used for the other indices. The evaluation also depends on the chosen color scale, which is different for every index and might be shifted to one or another direction based on the knowledge of the local situations or the intention of the statement. A particular artifact of the simple travel time calculation with catch radii is the misevaluation of short segments, roundabouts, or motorway links, where start and end nodes lie too close together. This problem could be tackled using more refined techniques to process the FCD, notably a map-matching algorithm.

4 CONCLUSIONS

The presented method offers an efficient and valuable procedure to gain a quick overview of a city's traffic conditions. It uses raw FCD and evaluates performance indices based on travel time distributions on network sections. The road network is constructed using open-source data, namely OSM ways. The presented method can provide an extension to the existing methods of the German Highway Capacity Manual (HBS) or the German Guideline for Integrated Network Planning (RIN) to evaluate traffic quality.

The method could further be used to identify problematic road segments, which can be analyzed in more detail to find the reasons for disruptions and bottlenecks. A similar analysis is also conceivable for other transport modes, e.g., using floating bike data.

ACKNOWLEDGEMENTS

The research for this paper was conducted as part of the project "Bottleneck Analysis for Urban Road Networks" funded by the German Federal Ministry for Digital and Transport, represented by the Federal Highway Research Institute (BASt). We want to thank our project partners for their excellent cooperation and for providing FCD for the city of Hannover, as well as our colleagues Marvin Baumann and Claude Weyland for their advice and ideas on this work.

REFERENCES

- Altintasi, O., Tuydes-Yaman, H., & Tuncay, K. (2017). Detection of urban traffic patterns from floating car data (FCD). , *22*, 382–391. (Publisher: Elsevier)
- Axer, S., & Friedrich, B. (2014). Level of service estimation based on low-frequency floating car data. , *3*, 1051–1058. (Publisher: Elsevier)
- HBS. (2015). *Handbuch für die Bemessung von Straßenverkehrsanlagen: Teil S: Stadtstraßen*. FGSV - Forschungsgesellschaft für Straßen- und Verkehrswesen.
- Jenelius, E., & Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. , *53*, 64–81. (Publisher: Elsevier)
- osmapi*. (2023). Retrieved 2023-02-28, from <https://pypi.org/project/osmapi/>
- overpass*. (2019). Retrieved 2023-02-28, from <https://pypi.org/project/overpass/>
- Peter, L., Janko, J., Schick, N., Waßmuth, V., Friedrich, M., & Bawidamann, J. (2021). Entwicklung eines aktuellen, echtzeit-verfügbaren Key Performance Indicator (KPI) Systems für das deutsche Autobahnnetz.
- Radde, M., Leerkamp, B., & Klemmer, T. (2016). Messung der Zuverlässigkeit von Strassennetzen mit Floating Car Data. , *60*(11).
- RIN. (2008). *Richtlinien für die integrierte Netzgestaltung*. FGSV - Forschungsgesellschaft für Straßen- und Verkehrswesen.
- Schäfer, R.-P., Gühnemann, A., & Thiessenhusen, K.-U. (2003). Neue Ansätze im Verkehrsmonitoring durch Floating Car Daten.
- Tomtom traffic index*. (2023). Retrieved 2023-02-28, from <https://www.tomtom.com/traffic-index/about/>
- Ulm, M., Heilmann, B., Asamer, J., Graser, A., & Ponweiser, W. (2015). *Identifying congestion patterns in urban road networks using floating car data*.

Exploring the impact of the social network geography on the individual's activity space using structural equation models

Benjamin Gramsch Calvo^{*1} and Kay W. Axhausen¹

¹Institute for Transport Planning and Systems, ETH Zurich, Switzerland

^{*}Corresponding author: bgramsch@ethz.ch

SHORT SUMMARY

Most leisure travel has social motivations, one of them is to be in contact with people from their social network, this means that the decision does not only depends on individual preferences and restrictions but also on the other person (or persons) involved in the activity. This means that the places an individual visits for leisure are strongly correlated with the geographic location of their social network. This hypothesis is tested with a structural equation model that includes social needs and mobility demand as latent variables. The model shows a strong correlation between these two variables, showing a positive impact on the geographic distribution of social networks and the number of contacts with the area of leisure activity space, and the number of frequently visited locations. This model shows the social network's importance in individual mobility decisions and patterns.

Keywords: Social Networks, Activity Space, Leisure Travel.

1 INTRODUCTION

In travel demand, leisure travel plays an important and often overlooked role, which can negatively impact the functioning of the overall transport system. Also, leisure travel is primarily social travel as a small share of leisure is solitary (Axhausen, 2005) leading to an increasing interest in understanding leisure travel demand and the influence of social needs and activities on travel decisions. Studying social network geographies can help understand how the geographic distribution of social networks impacts daily mobility patterns, opening new perspectives to transport modeling.

Social and leisure travel is more complex than work-related travel as it has many variables that influence it while being flexible in time and space (Ruiz et al., 2016). Therefore, the individual's social network geography is an essential determinant of the social travel decision process (Axhausen, 2006). For this reason, there has been a growing amount of literature on the impact of social networks on individual mobility behavior, as between 10 and 30% of all human travel can be explained by social relations and spatial characteristics of social networks (Cho et al., 2011; Axhausen, 2008). For this reason, transport modelers have started to include social networks in their models (see Axhausen (2005)). One of the essential findings is the negative correlation between the probability of being friends and the geographical distance between two people (Liben-Nowell et al., 2005; Kowald, 2013), as the opportunity cost of the meeting is lower (Arentze et al., 2013). Thus, individuals' social networks are essential to urban mobility and access to the opportunities the city generates (Guidon et al., 2017).

One of the main differences between social travel and other types of travel is the motivation to maintain individual social connections. This motivation directly impacts the process of choosing a leisure activity and destination, as it involves not only personal preferences for the characteristics of the location but the preference and geographical location of the alters. Therefore, people with more extensive social networks tend to have higher heterogeneity in the type of locations visited and to perform more socially motivated travel (Baburajan, 2019). Moore et al. (2013) has studied the link between "longer-term" (social networks) and "shorter-term" (social activities) decisions, showing that the social ties and network density of an individual explain the activity duration and distance.

This paper looks to contribute to understanding how social networks impact individual mobility patterns by analyzing the impact of the geography and structure of social networks on the number of regular leisure locations and the size of the activity space of the ego. This question is relevant

as the alters' home locations can be considered as *anchor points* of mobility in the city, acting as pivotal places in the ego's activity space and the selection of leisure activities. To analyze this hypothesis, we use a structural equation model that includes three latent variables, *social needs*, *relationship strength*, and *mobility demand*. This model shows the relation between the unobservable variables described above by analyzing the covariances of observable variables that are associated with the unobservable variables.

The paper continues as follows: section 2 describes the methodology used in more detail. Section 3 describes the results of the models estimated. Finally, section 4 finishes with some conclusions of the paper.

2 METHODOLOGY

The data was collected through a survey conducted in Zurich, Switzerland, that included an egocentric social network and questions on regularly visited leisure venues or locations through a *place generator* and *place interpreter* Gramsch Calvo & Axhausen (n.d.). This information creates the link between the geographic distribution of social networks and the geographic distribution of leisure activities in the city. A structural regression model with three endogenous variables is used to analyze this relationship, following the work by Washington et al. (2003). To estimate the model, we define the latent variables as follows:

- Social needs (exogenous): This variable represents the individual's sociability. It explains the number of contacts in the individual's social network, the size of the social network geography, and how often the individual meets with its alters.
- Relationship strength (endogenous): This variable shows the individual's proximity to their social network. It is measured by the average trust they have in their social network and the average capacity to ask for favors to them.
- Mobility demand (endogenous): This variable shows how much the individual moves and uses the city. It explains the number of leisure locations frequently visited and the size of the leisure activity space.

All these latent variables are constructed with observed variables from the social network structure and distribution of leisure activities. The model solves four equations simultaneously:

$$X_i = \Lambda_x \xi + \delta_i \quad (1)$$

$$Y_{ji} = \Lambda_{y_j} \eta_j + \varepsilon_{ij} \quad (2)$$

$$\eta_1 = \Gamma_1 \xi + \zeta_1 \quad (3)$$

$$\eta_2 = B \eta_1 + \Gamma_2 \xi + \zeta_2 \quad (4)$$

Equation 1 corresponds to the estimation of the dependent variables by the latent dependent variable, and equation 2 is the estimation of the observed dependent variables by the latent dependent variables. Λ_i is the coefficient of X and Y_i or matrix of loadings corresponding to the latent dependent variables, η is the vector of latent dependent variables, and ξ is the vector of latent dependent variables. Finally, equations 3 and 4 are the structural equations, where B and Γ_i are the weights predicted between the independent and dependent variables.

The model used in this paper uses the social network geography as the independent latent variable, which explains the observable variables *number of contacts* and *area of geographical distribution*. The two dependent latent variables are relationship quality and mobility patterns. The first variable explains the observable variables *social capital*, *meeting frequency*, and *trust level*. The second explains *number of places regularly visited* and *area of activity space*. Table 1 explains the type of variables used and their description.

To better explain the relationship between variables in the model, figure 1 shows a path diagram with the variables of interest and their relationship, adapted from the diagrams presented in Lin (2021). The circles represent latent variables, and squares are observable variables; vectors represent the direction of the correlation, single-pointed arrows represent the direction of the effect, and double-pointed arrows represent residuals of the covariance.

Table 1: Description of variables considered in the model

Name of Variable	Mean	SD	Type	Description
# of contacts	11.94	6.13	Ordinal	Number of contacts specified in the <i>name generator</i> .
Area of social network	6.30	3.37	Continuous	Logarithm of the area in km^2 of the convex hull generated by the geographical distribution of the social networks' house locations.
Trust level	82.62	19.21	Continuous	Percentage of the ego's alters with whom they discuss important problems or can ask for help.
Social capital	16.61	16.61	Continuous	Average social capital levels on the ego's social network, measured by the number of contacts that the ego would ask for help in different situations.
Meeting frequency	18.03	10.31	Continuous	Total number of time the ego meets with their alters in an average week.
Area of activity space	3.07	2.29	Continuous	Logarithm of the area in km^2 of the convex hull generated by the leisure locations visited by the ego.
# places visited	7.21	4.03	Ordinal	Number of places mentioned by the ego in the <i>place generator</i>

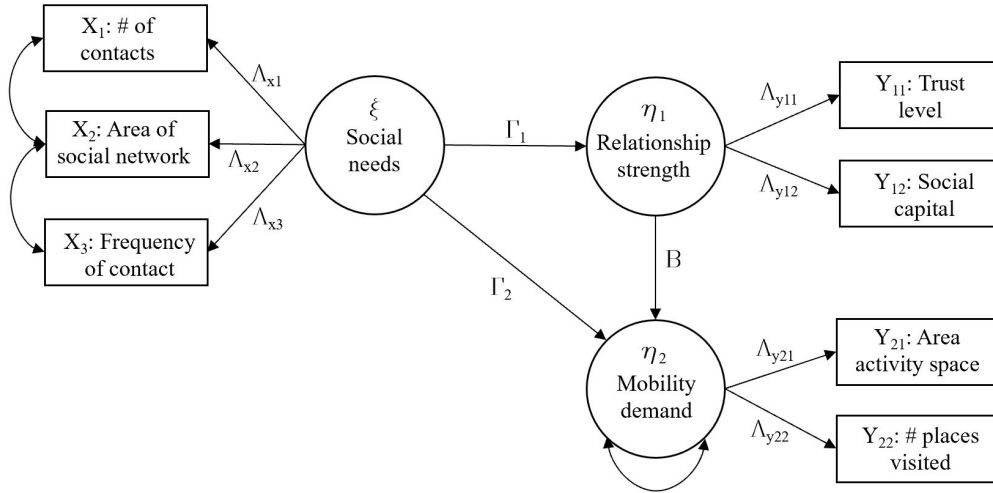


Figure 1: Specification of variables of the model and their effects

3 RESULTS AND DISCUSSION

To measure the impact of the social needs of an individual on their mobility demand, we have estimated two models following the framework explained in section 2. The first one includes all the variables except the latent variable *relationship strength* and its endogenous variables *trust level* and *social capital*. The second model includes all the variables of the model. Table 2 presents the estimate of the loadings. The latent variable *social needs* influences the three measured variables compared to the number of contacts. The social network area is explained approximately two times more by social needs, while the frequency of contact is explained by a factor of approximately three and a half. Analyzing the *relationship strength*, we can see that both *social capital* and *average trust* is explained similarly by *mobility demand*. *Mobility demand* also has a positive impact on *leisure activity space*, and *number of places visited*, with the former variable explained twice as much as the latter. Both models show similar loading factors with a difference of 0.315 in the impact of *social needs* on *social network area* and 0.228 on *frequency of contact*. The covariances of the model have the expected direction. There is a negative correlation between the *social network area* and the *frequency of contact*, as the farther away the ego's social connections live, the costlier it is to visit them. Therefore there is a lower number of face-to-face meetings. Conversely, the more contacts the ego has, the more frequently they meet with people.

Regarding the model's goodness-of-fit, models 1 and 2 have similar indices. The Comparative Fit Index and the Tucker-Lewis Index are one or higher, and the Root Mean Square Error is close to zero. The only noticeable difference is in the Chi-square, in which model two has a robust value of 8.939 against 1.079.

Table 2: Loadings of the measurement model

A. Estimates of loadings		Model 1			Model 2				
		Measurement variables	Loadings	Standardized estimates	<i>p-value</i>	Loadings	Standardized estimates	<i>p-value</i>	
Social needs	Number of contacts	1.000	0.826	1.000	0.960	0.960			
	Social network area	2.734	2.259	2.025	1.944	1.944	<0.001		
	Frequency of contact	4.286	3.541	3.451	3.313	3.313	<0.001		
Relationship strength	Social capital	-	-	1.000	9.502	9.502			
	Average trust	-	-	0.978	9.293	9.293	<0.001		
Mobility demand	Leisure activity space	1.000	1.609	1.000	1.614	1.614			
	Number of places visited	0.498	0.801	0.495	0.798	0.798	<0.001		
B. Covariances of measurement error									
Measurement variables		Covariances		<i>p-value</i>		Covariances		<i>p-value</i>	
Social network area	Frequency of contact	-7.415	0.009	-5.856	0.002				
Number of contacts	Frequency of contact	2.015	0.100	1.761	0.085				
C. Goodness-of-fit									
Name of test		Standard		Robust		Standard		Robust	
Chi-square	Test statistic	0.269	1.079	5.155	8.939				
	Degrees of freedom	2	2	9	9				
	<i>p-value</i> (chi-square)	0.974	0.583	0.821	0.443				
Goodness of fit index	Comparative Fit Index (CFI)	1.000	1.000	1.000	1.000				
	Tucker-Lewis Index (TLI)	1.009	1.006	1.008	1.000				
Root Mean Square Error Approximation	RMSEA	0.000	0.000	0.000	0.000				
	<i>p-value</i> RMSEA	0.980	0.909	0.999	0.991				

Table 3: Estimates of the structural model

A. Regression weights							
		Model 1			Model 2		
Endogenous variables	Explanatory variables	Estimate	Standardized estimates	<i>p-value</i>	Estimate	Standardized estimates	<i>p-value</i>
Mobility demand	Social needs	0.589	0.302	<0.001	0.526	0.313	<0.001
	Relationship strength	-	-	-	0.012	0.068	0.362
Relationship strength	Social needs	-	-	-	-4.131	9.502	<0.001

Table 3 shows the results of the regression of the latent variables. The difference between the estimations of model 1 and model 2 is non-significant. The latter model standardized estimate shows an increase of 0.09. There is a significant impact of *social needs* in *mobility demand*, but *relationship strength* shows a small non-significant impact. Individuals with more alters tend to visit more places for leisure and to have a higher leisure activity space. However, this mobility demand does not depend on the strength of the relationship with those alters. Finally, the more social needs a person has, the lower the ego’s average *relationship strength*; this is explained because individuals with more contacts tend to have proportionally more peripheral alters, reducing the average social capital and trust level of the network.

4 CONCLUSIONS

Leisure travel is highly influenced by the individual’s friends, family, and acquaintances. As most leisure travel has a social motivation associated, the places a person regularly visits are directly affected by where the individual’s social network lives. This paper tries to understand that correlation by comparing the geographic size of the social network with the leisure activity space. The results show a correlation between these two variables. The latent variable *social needs* explains the number of contacts in the ego’s social network and their distribution in space. At the same time, it also significantly impacts *mobility demand*. Also, we can see that *relationship strength* does not impact mobility demand.

The results of this paper contribute to the literature on social networks and mobility, as it links the number of friends and their geographic distribution with the number of leisure locations a person visits and their leisure activity space. These two variables of mobility demand can help explain many other travel decisions, such as mode choice and mobility tool ownership. These variables will be included in future work related to social networks and leisure activity spaces.

REFERENCES

- Arentze, T. A., Kowald, M., & Axhausen, K. W. (2013). An agent-based random-utility-maximization model to generate social networks with transitivity in geographic space. *Social Networks*, 35(3), 451–459. doi: 10.1016/j.socnet.2013.05.002
- Axhausen, K. W. (2005). Social networks and travel: Some hypotheses.. doi: 10.3929/ETHZ-A-004663201
- Axhausen, K. W. (2006). Social factors in future travel: A qualitative assessment. *Research Collection, ETH Zurich*, 29.
- Axhausen, K. W. (2008). Social networks, mobility biographies, and travel: Survey challenges. *Environment and Planning B: Planning and Design*, 35(6), 981–996. doi: 10.1068/b3316t
- Baburajan, V. (2019). An investigation into the influence of land-use, social networks and information and communication technologies on destination choice for social activities. *Transactions of the Association of European Schools of Planning*, 14.
- Cho, E., Myers, S. A., & Leskovec, J. (2011). Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11* (p. 1082). San Diego, California, USA: ACM Press. doi: 10.1145/2020408.2020579

- Gramsch Calvo, B., & Axhausen, K. W. (n.d.). Place generator & place interpreter: A new survey method to understand regular destination choice. *Research Collection, ETH Zurich*. doi: 10.3929/ETHZ-B-000561743
- Guidon, S., Wicki, M., Axhausen, K. W., & Bernauer, T. (2017). Investigating the relationship between individuals' social networks, mobility behaviour, and social capital: Survey method and first results. *Research Collection, ETH Zurich*.
- Kowald, M. (2013). Focussing on leisure travel: The link between spatial mobility, leisure acquaintances and social interactions. *PhD Thesis, ETH Zurich*.
- Liben-Nowell, D., Novak, J., Kumar, R., & Raghavan, P. (2005, June). Geographic routing in social networks. , *102*(33), 11623 – 11628. doi: <https://doi.org/10.1073/pnas.0503018102>
- Lin, J. (2021). *Introduction to structural equation modeling (SEM) in R with lavaan*. Los Angeles, California,.
- Moore, J., Carrasco, J.-A., & Tudela, A. (2013). Exploring the links between personal networks, time use, and the spatial distribution of social contacts. *Transportation*, *40*(4), 773–788. doi: 10.1007/s11116-013-9467-4
- Ruiz, T., Mars, L., Arroyo, R., & Serna, A. (2016). Social networks, big data and transport planning. *Transportation Research Procedia*, *7*.
- Washington, S., Karlaftis, M. G., & Mannering, F. L. (2003). *Statistical and econometric methods for transportation data analysis*. Boca Raton: Chapman & Hall/CRC.

Choice-driven Service Network Design and Pricing in Intermodal Transport

Adrien Nicolet¹ and Bilge Atasoy*²

¹PhD Candidate, Dept. of Maritime and Transport Technology, Delft University of Technology,
The Netherlands

²Associate Professor, Dept. of Maritime and Transport Technology, Delft University of
Technology, The Netherlands

SHORT SUMMARY

In intermodal transport, Service Network Design (SND) problems cover most tactical decisions of a carrier. Nevertheless, among the literature on SND, very few works include pricing decisions and the preferences of the shippers. In this study, we contribute to the existing body of knowledge by proposing a choice-driven and cycle-based formulation of the Service Network Design and Pricing (SNDP) problem which considers different aspects of the mode choice decisions of shippers. This formulation aims at finding the itineraries, frequencies and prices of the services that will maximize the profit of an intermodal carrier. Moreover, the mode choice preferences of shippers are modeled as a utility maximization accounting not only for the logistics costs, but also the frequency of the offered services and the accessibility of the transport mode. This bi-level formulation can be reformulated into a single level linear problem. The proposed model is compared to two other models (one cycled-based and one path-based) where shippers are assumed to be purely cost minimizers. While the latter generate higher profits, they also result in unrealistic mode shares, with road transport being negligible. On the other hand, the proposed formulation leads to mode shares that are considerably closer to reality. In addition, higher revenues can be generated with a cycle-based formulation compared to a path-based as it allows for more consolidation opportunities for the carrier.

Keywords: Choice-driven Optimization, Intermodal Transport, Mode Choice, Pricing, Service Network Design

1 INTRODUCTION

In intermodal freight transport, planning at the tactical level is of key importance to make the best use of existing infrastructure and available assets and to ensure reliable transport plans. In particular, Service Network Design (SND) problems cover, among other things, the decisions of transport operators about the itineraries to be served, the offered frequencies and how demand should be assigned to these services. The majority of existing studies on SND are formulated as a cost minimization of the transport operator and, therefore, do not include the revenues of fulfilling the transport orders, as highlighted by Elbert et al. (2020).

For the works actually applying a profit maximization, they mostly assume fixed tariffs that are included as parameters into the model, as in Bilegan et al. (2022). Only a handful of works include pricing as a decision of the problem. Some are using game theory paradigms to solve the SND, see for example Qiu et al. (2021), while others come up with a Mixed Integer Problem formulation, as for Martin et al. (2021).

In their work, Tawfik & Limbourg (2019) propose a bi-level SND and pricing model, with the upper level representing the profit maximization of an intermodal transport operator and the lower level being the costs minimization of shippers. The shippers can choose between the services proposed by the operator or a competition alternative. The latter is represented as direct trucking and a fixed cost is assumed for it. The authors then reformulate this bi-level model into a single-level problem and apply linearization procedures to come up with a Mixed Integer Linear Problem. Yet the representation of shippers here is limited as they look for the minimum cost and other attributes are not being evaluated.

The body of literature on choice-driven optimization has advanced in other domains such as assortment optimization, e.g., Davis et al. (2014), on-demand mobility solutions, e.g., Atasoy et al. (2015) and Sharif Azadeh et al. (2022) etc. Nevertheless, due the complexity of the decision-making

process, it has not been yet sufficiently addressed in intermodal transport. The inputs from the demand side are typically considered as exogenous to the optimization problem or introduced with simplistic assumptions. In this paper, we represent the preferences of shippers more realistically rather than cost minimization and integrate this behavioral response within the service network design problem.

2 METHODOLOGY

In our study, we use the work by Tawfik & Limbourg (2019) as benchmark and expand it by introducing several new elements. Firstly, the path-based formulation is replaced by a cycle-based formulation. The latter is deemed more accurate to represent realistic decision-making. Indeed, most intermodal transport services go back and forth on an itinerary with a defined schedule. The cycle-based representation also enables a more elaborate representation of services as multiple intermediary stops can be added in both directions. In addition, it simplifies the asset management of the operators. In a path-based formulation, they may need to re-balance the vehicles at the end of the planning horizon; whereas a cycle-based representation ensures that each vehicle ends up at its starting point. It is noteworthy that we keep an arc-based pricing representation, as shippers will not be charged for a journey whose distance is longer than between the origin and the destination of their cargo.

Secondly, shippers' mode choice behavior is represented as the maximization of a utility function including not only logistics costs, but also non-monetary attributes. It is indeed known that there exist other influential factors, such as time and reliability, see for example Li et al. (2020). In our work, besides the price charged by the carrier, the utility function also consists of the offered frequency and the accessibility to the transport mode. The estimation of the model coefficients can be found in Nicolet et al. (2022). Compared to the benchmark formulation, where the frequency appeared only in the upper level, it now also appears in the lower level problem. In particular, it has a positive influence on the shippers' utility. Indeed, with the rise of just-in-time logistics and the observed trend of companies to reduce their inventories, it is desirable for shippers to have frequent transport services. We therefore develop a choice-driven and cycle-based service network design model (CD-SNDP) which is formulated next.

Mathematical formulation for the proposed CD-SNDP

The transport network is represented as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where \mathcal{N} is the set of terminals and $\mathcal{A} = \{(i, j) : i, j \in \mathcal{N}, i \neq j\}$ the set of links between these terminals.

Upper level

The operator's fleet is heterogeneous and the different vehicle types are denoted by set \mathcal{K} . The number of available vehicles for type k is V_k and the corresponding capacity is Q_k .

Set \mathcal{S} includes all the transport services that can be run by the operator. Unlike the benchmark, where each service corresponds to a single arc of \mathcal{A} , a service is composed of a sequence of arcs. Each arc in this sequence is called a leg and the whole sequence of legs for a given service s is noted \mathcal{L}_s . The cycle-based formulation of the problem implies that the sequence starts and ends at the same node.

The maximum number of cycles of service s that can be performed by vehicle type k is named W_{sk} : it typically consists of the maximum operating time divided by the cycle time (sum of travel time and time at terminals). Each service s has a fixed cost c_{sk}^{FIX} of operating it with vehicle type k and a variable cost $c_{ij sk}^{\text{VAR}}$ per container transported between terminals i and j . Moreover, we introduce the parameter $\delta_{ij l_s}$, which equals one if a container traveling from i to j uses the service leg l_s and zero otherwise.

The transport operator has three decision variables in the upper level problem:

- v_{sk} is the number of vehicles of type k that the operator allocates to each service s ;
- f_{sk} is the frequency of service s per vessel type k ;
- p_{ij} is the price per container charged to shippers wanting to transport goods from i to j .

Lower level

The shippers are represented as a whole: therefore, their demand is aggregated. The container transport demand between terminals i and j is denoted by D_{ij} . Shippers decide to assign demand to the transport operator or their competitors by the maximization of their utility. The utility function of using the services proposed by the transport operator between i and j is noted U_{ij} and is dependent on p_{ij} and f_{sk} , whereas the utility of using a competing alternative h is written U_{ij}^h . Finally, the decision variables of the lower level consist in the number of containers that are assigned to the operator's services ($x_{ij sk}$) and to every competing alternative (z_{ij}^h). All the aforementioned sets, parameters and decision variables are listed in Table 1.

Table 1: Notation

Sets:	
\mathcal{N}	Terminals (indices: i, j)
\mathcal{A}	Arcs (i, j)
\mathcal{K}	Vehicle types (index: k)
\mathcal{S}	Potential services (index: s)
\mathcal{L}_s	Legs of service $s \in \mathcal{S}$ (index: l_s)
\mathcal{H}	Competing alternatives (index: h)
Parameters:	
V_k	Number of vehicles of type k in the operator's fleet
Q_k	Capacity of vehicle type k [TEUs]
W_{sk}	Maximum number of cycles of service s that can be performed by vehicle type k
c_{sk}^{FIX}	Fixed cost of operating service s with vehicle type k [€]
$c_{ij sk}^{\text{VAR}}$	Variable cost of transport between i and j with service s and vehicle type k [€/TEU]
$\delta_{ij l_s}$	Dummy param. equal to 1 if container traveling from i to j uses service leg l_s , 0 otherwise
D_{ij}	Aggregated transport demand of shippers between i and j [TEUs]
U_{ij}	Utility of using the operator's services between i and j
U_{ij}^h	Utility of using competing alternative h between i and j
Variables:	
v_{sk}	Number of vehicles of type k assigned to service s by the operator
f_{sk}	Frequency of service s operated with vehicle type k
p_{ij}	Price charged by the operator to shippers wanting to transport goods from i to j [€/TEU]
$x_{ij sk}$	Cargo volume using service s operated with vehicle type k between i and j [TEUs]
z_{ij}^h	Cargo volume using competing alternative h between i and j [TEUs]

The proposed CD-SNDP is expressed as a bi-level Mixed Integer Problem as follows:

$$\max_{v, f, p, x, z} \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} p_{ij} x_{ij sk} - \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} c_{sk}^{\text{FIX}} f_{sk} - \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} c_{ij sk}^{\text{VAR}} x_{ij sk} \quad (1)$$

$$\text{s.t.} \quad \sum_{s \in \mathcal{S}} v_{sk} \leq V_k \quad \forall k \in \mathcal{K} \quad (2)$$

$$f_{sk} \leq W_{sk} v_{sk} \quad \forall s \in \mathcal{S}, \forall k \in \mathcal{K} \quad (3)$$

$$\sum_{(i,j) \in \mathcal{A}} \delta_{ij l_s} x_{ij sk} \leq Q_k f_{sk} \quad \forall l_s \in \mathcal{L}_s, \forall s \in \mathcal{S}, \forall k \in \mathcal{K} \quad (4)$$

$$x_{ij sk} \leq \sum_{l_s \in \mathcal{L}_s} \delta_{ij l_s} D_{ij} \quad \forall (i,j) \in \mathcal{A}, \forall s \in \mathcal{S}, \forall k \in \mathcal{K} \quad (5)$$

$$p_{ij} \geq 0 \quad \forall (i,j) \in \mathcal{A} \quad (6)$$

$$v_{sk} \in \mathbb{N} \quad \forall s \in \mathcal{S}, \forall k \in \mathcal{K} \quad (7)$$

$$f_{sk} \in \mathbb{N} \quad \forall s \in \mathcal{S}, \forall k \in \mathcal{K} \quad (8)$$

where x and z solve:

$$\max_{x, z} \sum_{(i,j) \in \mathcal{A}} \left(\sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} U_{ij} x_{ij sk} + \sum_{h \in \mathcal{H}} U_{ij}^h z_{ij}^h \right) \quad (9)$$

$$\text{s.t. } \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} x_{ijsk} + \sum_{h \in \mathcal{H}} z_{ij}^h = D_{ij} \quad \forall (i, j) \in \mathcal{A} \quad (10)$$

$$x_{ijsk} \geq 0 \quad \forall (i, j) \in \mathcal{A}, \forall s \in \mathcal{S}, \forall k \in \mathcal{K} \quad (11)$$

$$z_{ij}^h \geq 0 \quad \forall (i, j) \in \mathcal{A}, \forall h \in \mathcal{H} \quad (12)$$

At the upper level, the objective function of the transport operator (1) is to maximize their profit. It is computed as the revenues from the transported containers minus the fixed and variable costs of the offered services. Constraint (2) is the fleet size constraint for each vehicle type. Constraint (3) ensures that the service's frequency is inferior to the maximum number of cycles that can be performed by the assigned vehicles. Constraint (4) assures that the total number of containers transported on each leg of every service does not exceed the available capacity of the service, whereas constraint (5) ensures that no container can be assigned to a service that does not go through the origin or destination terminal of the container. The domains of the operator's decision variables are defined by constraints (6)-(8).

Regarding the lower level, shippers seek to maximize their utility (9) by assigning their containers either to the operator's services or to the competition. Moreover, constraint (10) enforces the total transport demand to be met. Finally, constraints (11)-(12) define the domain of the decision variables of the shippers.

The presented model is transformed to a single level problem inspired by Tawfik & Limbourg (2019). In doing that, our formulation has multiple nonlinearities to deal with in order to reach a mixed integer linear program. These are due to the pricing decision as well as the utility function involving price and frequency which are both decision variables of the model. We make use of the strong duality theorem, the transformation of frequencies into binaries and the big M method in order to deal with the embedded nonlinearities.

3 RESULTS AND DISCUSSION

To assess the performance of our method, it is applied to Inland Waterway Transport (IWT) of containers on a small network of three nodes (Rotterdam, Duisburg and Bonn) and compared with the results of the benchmark. We consider an IWT operator competing with two other modes (Road and Rail). The operator's fleet is homogeneous and composed of 30 vessels with a maximal capacity of 250 Twenty-Foot Equivalent Units (TEUs) and maximal operation time is assumed to be 120 hours per week. The sailing times and the time spent in ports as well as the costs for IWT and the two competing modes are estimated using the model of Shobayo et al. (2021); whereas the transport demand inputs come from the NOVIMOVE project, see Majoor et al. (2021).

Three different models are compared with each other:

- The Benchmark from Tawfik & Limbourg (2019) in which fleet constraints are added;
- The cycle-based version of the benchmark, SNDP, with Cost minimization of shippers, which is equivalent to the benchmark but with cycles allowed;
- The proposed Choice-Driven SNDP (CD-SNDP), with shippers' utility functions replacing costs.

Table 2 displays the main results of the three models applied to the three-node network. The cycle-based formulation offers more flexibility for the transport operator, who can propose services with intermediary stops and take advantage of consolidation opportunities, instead of offering only direct connections between two terminals. This translates into a 22% increase of the revenue with the SNDP compared to the benchmark. In the SNDP, almost all proposed services are between Rotterdam and Bonn with an intermediary stop in Duisburg; whereas this option is not available in the benchmark. As a result, the operator can serve higher shares of the demand between Rotterdam and Duisburg, where volumes are much greater than between Rotterdam and Bonn. Nevertheless, both the benchmark and the SNDP rely on the assumption that shippers are only interested in minimizing the costs when choosing the transport mode. This results in unrealistic modal shares, where only a tiny fraction of the total demand is assigned to Road. When the cost minimization is replaced by the utility maximization to get the CD-SNDP, the modal shares become much closer to the reality. Indeed, the data collected in Majoor et al. (2021) shows that the shares of container transport between the three considered terminals are 43% for IWT, 48% for Road and 8% for Rail. The integration of utility in the model implies that not all flows can be served by the operator. In particular, no demand is served between Duisburg and Bonn because

Table 2: Results on a 3-nodes network, with served demand without parentheses for import flows \rightarrow and in parentheses for export (\leftarrow).

		Benchmark	SNDP	CD - SNDP
Total revenue [mio €]		3.43	4.17	2.03
Weekly service frequencies	RTM \leftrightarrow DUI	15	0	0
	RTM \leftrightarrow BON	6	4	0
	DUI \leftrightarrow BON	27	0	0
	RTM \leftrightarrow DUI \leftrightarrow BON	-	26	30
Served demand by IWT	RTM \leftrightarrow DUI	58% (45%)	100% (77%)	100% (89%)
	RTM \leftrightarrow BON	79% (100%)	53% (67%)	53% (0%)
	DUI \leftrightarrow BON	100% (100%)	97% (100%)	0% (0%)
Modal shares	IWT	75%	89%	48%
	Road	0%	1%	50%
	Rail	25%	10%	3%

IWT is not competitive enough compared to Road. Undoubtedly, the revenue is lower in the CD-SNDP compared to the two other formulations. Yet it embeds a more realistic response of the demand to the decisions of the IWT operator. Moreover, it still emphasizes the benefits of our cycle-based formulation as the optimal solution includes no direct links at all, but only cycle services that stop at every terminal.

4 CONCLUSIONS

The results support that including a more detailed modeling of the mode choice behavior of shippers allows our service network design model to more accurately represent the situation. Moreover, the benefits of a cycle-based formulation in terms of flexibility and profitability are highlighted. The analysis needs to be carried out for bigger transportation networks in order to see the potential of the proposed methodology. Furthermore, our immediate future research relates to the inclusion of shippers' heterogeneity in the mode choice behavior for further improving the representation of the real situation. When the decision-making process of both the transport operators and shippers are represented more realistically, the transport systems as such can be improved further towards sustainability goals as the resources can be allocated to the right entities at the right time and place.

ACKNOWLEDGEMENTS

This research is supported by the project “Novel inland waterway transport concepts for moving freight effectively (NOVIMOVE)”. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 858508.

REFERENCES

- Atasoy, B., Ikeda, T., Song, X., & Ben-Akiva, M. E. (2015). The concept and impact analysis of a flexible mobility on demand system. *Transportation Research Part C: Emerging Technologies*, 56, 373-392.
- Bilegan, I. C., Crainic, T. G., & Wang, Y. (2022). Scheduled service network design with revenue management considerations and an intermodal barge transportation illustration. *European Journal of Operational Research*, 300(1), 164-177.
- Davis, J. M., Gallego, G., & Topaloglu, H. (2014). Assortment optimization under variants of the nested logit model. *Operations Research*, 62(2), 250-273.

- Elbert, R., Müller, J. P., & Rentschler, J. (2020). Tactical network planning and design in multimodal transportation—a systematic literature review. *Research in Transportation Business & Management*, 35, 100462.
- Li, Q., Rezaei, J., Tavasszy, L., Wiegmans, B., Guo, J., Tang, Y., & Peng, Q. (2020). Customers' preferences for freight service attributes of china railway express. *Transportation Research Part A: Policy and Practice*, 142, 225–236.
- Majoor, I., Ramos, C., Burgees, A., van der Geest, W., Hendriks, I., van Hassel, E., . . . Hofman, P. (2021, 09). *D.2.2: Novimove transport model architecture and data collection* (Tech. Rep.). NOVIMOVE Technical Report.
- Martin, F., Hemmelmayr, V. C., & Wakolbinger, T. (2021). Integrated express shipment service network design with customer choice and endogenous delivery time restrictions. *European Journal of Operational Research*, 294(2), 590–603.
- Nicolet, A., Negenborn, R. R., & Atasoy, B. (2022). A logit mixture model estimating the heterogeneous mode choice preferences of shippers based on aggregate data. *IEEE Open Journal of Intelligent Transportation Systems*, 3, 650–661.
- Qiu, X., Xu, S. X., & Xu, G. (2021). Pricing and scheduling of barge hinterland transportation service for inbound containers. *Transportation Research Part E: Logistics and Transportation Review*, 156, 102515.
- Sharif Azadeh, S., Atasoy, B., Ben-Akiva, M. E., Bierlaire, M., & Maknoon, M. (2022). Choice-driven dial-a-ride problem for demand responsive mobility service. *Transportation Research Part B: Methodological*, 161, 128-149.
- Shobayo, P., Nicolet, A., van Hassel, E., Atasoy, B., & Vanelslander, T. (2021). Conceptual development of the logistics chain flow of container transport within the rhine-alpine corridor. In *European transport conference (etc)* (pp. 1–17).
- Tawfik, C., & Limbourg, S. (2019). A bilevel model for network design and pricing based on a level-of-service assessment. *Transportation Science*, 53(6), 1609–1626.

Estimation of demand models for long-distance cross-border travel

Ida Kristoffersson*¹, Chengxi Liu¹

¹ Dr., VTI Swedish National Road and Transport Research Institute, Box 55685,
10215 Stockholm, Sweden

SHORT SUMMARY

Although long-distance cross-border travel contributes significantly to global emissions from the transport sector, transport models for this type of travel are scarce. In this study, a disaggregated travel demand forecasting model is estimated using Swedish national travel survey data 2011-2016 along with detailed supply data from European road, train, and ferry networks and a Worldwide air network, aiming at forecasting Swede's long-distance travel abroad. Mode choice, destination choice and trip generation are modelled by traditional Nested Logit models and Multinomial Logit models. Results show that values of time of long-distance cross-border travel derived from the model estimation are in general higher than values of time of long-distance domestic travel. Furthermore, elasticity estimates of level-of-service attributes for train suggest that infrastructure investments in high-speed rail network may have a profound effect on demand for long-distance cross-border travel, especially for business trips.

Keywords: Discrete choice modelling; Transportation network modelling; Long-distance cross-border travel; Mode choice; Destination choice; Trip generation

1. INTRODUCTION

Long-distance cross-border travel differs from regional and national travel in many respects, such as what determines traveller trip generation, mode, and destination choice. Due to the long distances of these trips, they usually contribute significantly to a country's total passenger-kilometres travelled, even though the number of long-distance cross-border trips is in general lower than the number of regional and national trips. Passenger-kilometres travelled by mode is important, especially since it is related to CO₂ emissions from transport, for which ambitious reduction targets have been set both at the EU and national levels. Travel demand forecast models are an important part of large-scale modelling to provide accurate input to cost-benefit analyses of large infrastructure investments or policy measures. The major advantage of these forecast models is that planned but not implemented investments and policies can be tested in the models and effects analysed.

One of the few existing demand models of long-distance cross-border travel is Trans-tools, which is a transport model for both passenger and freight transport in 42 European countries. The demand model for passenger transport is described in Rich and Mabit (2012). The networks (car, train, and air) and their level of service attributes are described in Rich et al. (2009). A model called Trust (TRT Trasporti e Territorio, 2018) was developed as a follow-up to the Trans-tools model, however in Trust there is no demand model, instead demand is treated as fixed origin-destination (OD) matrices. Pieters et al. (2012) describe an effort to develop sub-models for border crossing traffic in the Dutch national model. Somewhat more common are so called direct-demand models, especially concerning tourist travel. These models typically calculate the total number of tourists travelling to/from a destination zone as a function of e.g., GDP and population. Due to the aggregate nature of these models, it is not possible to calculate e.g., cross-elasticities

between modes. Examples of direct demand models include Divisekera (2010) for Australia, Santana-Jiménez and Hernández (2011) for Canary Islands and Li et al. (2017) for China. There are also some direct-demand models that focus on a certain mode, especially air travel, and predict e.g., number of air trips to certain airports (Gelhausen et al., 2018; Kim & Shin, 2016; Suh & Ryerson, 2019).

The lack of disaggregated travel demand models for long-distance cross-border travel can be a problem in practice when certain investments or policy measures might have a substantial impact on cross-border travel demand and cross-elasticities are of interest. One such example is high-speed train that connects large cities across countries. Witlox et al. (2022) determine a number of existing bottlenecks for European rail, such as the train travel time not being fast enough and too many interchanges. An analysis of the ability of policy measures and investments to remove these bottlenecks would benefit from travel demand models for long-distance cross-border travel.

2. METHODOLOGY

The travel demand models are formulated using classic discrete choice theory and logit formulations (McFadden, 1974). There are two sub-models per trip purpose (private/business): one nested logit model for mode and destination choice and one multinomial logit model for trip generation.

For the mode and destination choice model, the utility equation for an alternative (mode i and destination j) are formulated as:

$$U_{i,j} = ASC_i + \gamma_i I + \beta_i L_{i,j} + \delta_i D_j + \phi \log(A_j) + \varphi_i + \varepsilon_{i,j} \quad (1)$$

In Equation (1), ASC_i is the alternative specific constant for mode i . I is the vector of individual socio-economic attributes. $L_{i,j}$ refers to the vector of level-of-service attributes for mode i to destination j . D_j is the vector of destination variables per capita, e.g., GDP and number of hotel beds per resident. A_j is a destination attraction variable (size variable) that represents the attractiveness in terms of size and quantity of each destination zone, for which a non-linear log formulation is used, see (Daly, 1982), and φ_i refers to the error term at the mode level. Thus, the alternatives with the same mode i will share the same error term φ_i and therefore those alternatives are not independent of each other. $\varepsilon_{i,j}$ refers to the error term that is unique and independent for each alternative. The mode and destination choice model with the utility function described in Equation (1) then is a nested Logit model where mode is on the upper level. The choice of model structure with mode above destination or the other way around is an empirical question which is determined by the data. The model structure that in estimation yields a logsum parameter which is within the range of 0 and 1 is the preferred structure.

For the trip generation model, the utility function for an alternative k is formulated as follows, where k belongs to {no long-distance cross-border trip; daytrip; 1-5 nights, and 6+ nights} for private travel, and {no trip and trip} for business travel.

$$U_k = ASC_k + \theta_k I + \mu_k T + \varphi_k \logsum_{modeDestModel} + \theta_k \quad (2)$$

In Equation (2), I is again a vector of socio-economic variables, T is a vector of time period variables such as Christmas, and $\logsum_{modeDestModel}$ is the logsum variable calculated from the estimated mode and destination choice model. θ_k , μ_k and φ_k are associated parameter vectors. The trip generation model is then a Multinomial Logit model.

3. RESULTS AND DISCUSSION

The travel demand data consists observations of long-distance (one-way distance 100 km or longer) cross-border trips from the Swedish national travel survey for the years 2011-2016 (Trafikanalys, 2017). The respondents in the Swedish national travel survey were asked which trips longer than 100 km they have made during the last month and which trips longer than 300 km they have made during the last three months. There has been one more national travel survey conducted after this, in 2019, but in the 2019 survey only trips from the measurement day were asked for, which resulted in very few long-distance international trips. Therefore, the 2019 survey could not be used in this study. After the data cleaning process, the trip data consists of 3561 (83%) private trips and 717 (17%) business trips. Out of the private trips, 324 (9%) are daytrips, 1348 (38%) are trips with 1-5 nights away, and 1889 (53%) are trips with 6 or more nights away. The modal shares for private trips differ a lot depending on number of nights away, which is a motivation for testing model segmentation across this variable. Car trips dominate for private daytrips, car and air trips are of about equal size for private trips 1-5 nights away, and air is the dominating mode for private trips 6+ nights away and for business trips.

One of the major tasks of this work was to develop digital European-wide/worldwide networks for major travel modes so that level-of-service data can be generated from these networks. Level-of-service data is generated at zone level using the transport modelling software TransCad (<https://www.caliper.com/tcovu.htm>). The zonal system in the long-distance model component of the Swedish national travel demand model is used for zones within Sweden, while NUTS zone system is used to represent Europe. Outside Europe, nations are represented as zones. In total, four networks are developed for car/bus, train, air, and ferry. Networks for car/bus, train, and ferry are European-wide while network for air is worldwide. As an example, the network for train is shown in Figure 1.

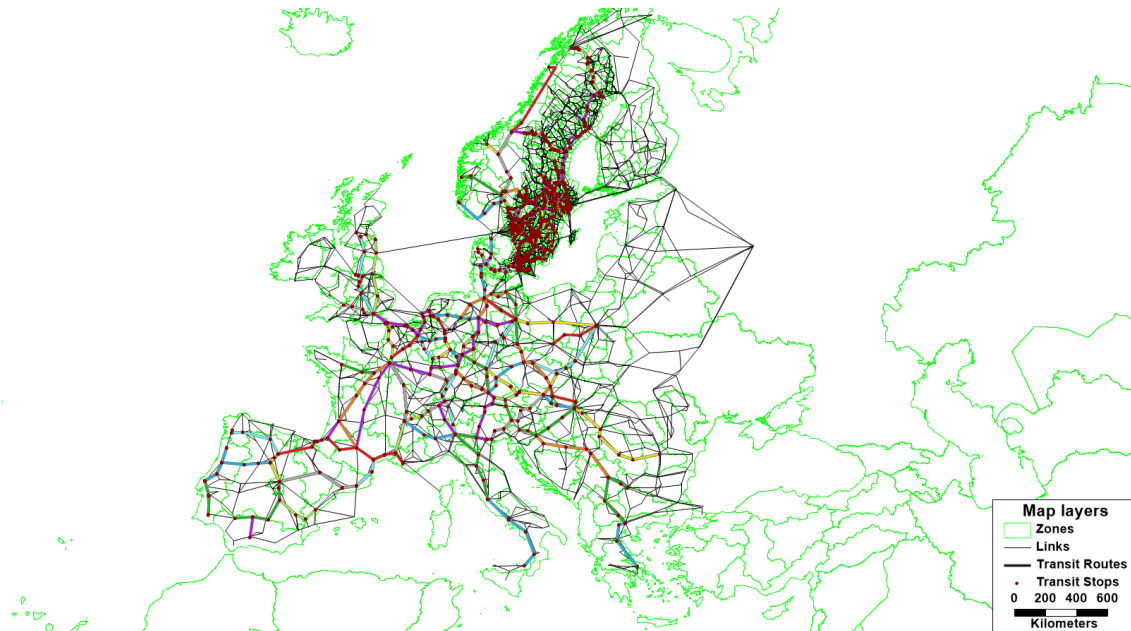


Figure 1: The European train network developed in TransCad.

Table 1 presents the estimation results for the mode and destination choice model for business trips. Note that there are four mode-destination choice models estimated: private daytrip, private 1-5 nights, private 6+ nights, and business, but there is only space to show results of one of these

in this short paper. The table shows the final model specification for business trips. The initial model specifications were set to include all variables that are relevant and then insignificant variables have been removed gradually. A large number of model specifications were tested before selecting the final version. The t-values in the table show the statistical significance of the parameters in the model. A t-value (absolute value) larger than 1.96 means that there is a 95% probability that the parameter is different from zero, i.e., it has an effect in the model. A few parameters with a lower significance level are kept in the model (shown in red in Table 1). These are either alternative specific constants that would be used as calibration constant in an implementation of the model or important level of service variables.

Table 1: Estimated parameter values of the mode and destination choice model for long-distance cross-border business trips

Parameter name	Explanation	Mode	Parameter value	t-value
ϕ	Log-size	all	0.717	23.70
$\beta_{BedPerArea.noAir}$	Hotel beds per area	car, bus, train	0.059	3.65
$\beta_{GDP.noAir}$	GDP per capita	car, bus, train	3.810	8.59
$\beta_{GDP.air}$	GDP per capita	air	3.196	18.65
$\beta_{TT.car}$	Travel time	car	-0.0080	-11.97
$\beta_{TT.PT}$	In-vehicle time	bus, train, air	-0.0039	6.61
$\beta_{AC.train}$	Access/egress time	train	-0.0783	-3.76
$\beta_{AC.air}$	Access/egress time	air	-0.0093	-6.75
β_{Cost}	Travel cost	all	-0.0038	-4.67
$\beta_{LogCostLowMedInc}$	Log(Travel cost) for low/medium income segment	all	-0.6302	-1.70
$\beta_{CarHH.car}$	Availability of car in household	car	0.372	3.26
ASC_{bus}	Alternative specific constant	bus	-1.767	-5.74
ASC_{train}	Alternative specific constant	train	0.341	0.98
ASC_{air}	Alternative specific constant	air	-0.007	-0.02
$Logsum_{destination}$	Accessibility to destination	all	0.786	1.71
Number of observations			717	
Number of parameters			15	
Log-likelihood			-3454.2	
Log-likelihood all parameters=0			-4969.4	
McFadden rho			0.305	

Parameters of destination attraction variables are positive, showing that the quantity in terms of number of hotel beds/population/employment has a positive effect in attracting travellers to given destination zones. When it comes to level-of-service variables, all travel time and travel cost

parameters are negative as expected. The disutility of travel time for car is in general higher than that for public transport which is expected since travel time on PT can be used for work activities. Furthermore, those with individual income lower than 30 TEUR have a higher cost sensitivity. Looking into the effects of socio-economic variables, number of cars in household is, as expected, a strong factor for choosing car. The logsum parameter is within the range of 0 and 1, indicating that the nested-Logit structure with mode at the upper level is valid. Value of time (VOT) estimates are derived from the estimated parameters of in-vehicle time and travel cost. Results are then compared to the VOT derived from the existing domestic long-distance model. In the business trip segment, VOT for car for long-distance cross-border trips is higher than VOT derived from the domestic long-distance trip model, while a reversed trend is found for public transport modes.

Table 2 shows the estimation results of the trip generation model for business trips. The available alternatives are taking no trip or conducting a business trip. Note that trip generation model estimation results for private trips (conducting no trip, private daytrip, private trip 1-5 nights, or private trip 6+ nights) exist but had to be left out due to space limitation.

Table 2: Estimated parameter values of the trip generation model for long-distance cross-border business trips

Parameter name	Explanation	Alternative	Parameter value	t-value
$\beta_{LowMedInc.0}$	Low/medium income segment	No trip	1.207	5.34
β_{CarHH}	Availability of car in household	Business trip	0.181	4.53
β_{Female}	Traveller is female	Business trip	-1.019	-11.21
β_{Age31_64}	Traveller age 31-64	Business trip	0.729	4.64
$\beta_{Age>64}$	Traveller age >64	Business trip	-1.127	-4.67
$\beta_{HighInc}$	High income segment	Business trip	1.156	8.01
β_{Summer}	Summer time	Business trip	-0.837	-5.68
$\beta_{Christmas}$	Christmas time	Business trip	-0.945	-4.01
ASC	Alternative specific constant	Business trip	-4.740	-23.81
Number of observations			39996	
Number of parameters			9	
Log-likelihood			-3066.4	
Log-likelihood all parameters=0			-27723.1	
McFadden rho			0.889	

It is found that low income is an important explanatory factor that contributes to not conducting any long-distance cross-border trips, which is expected. High income is a positive factor for conducting business trips. Pensioners (age >64) and female travellers are less likely to conduct business trips. Number of cars in the household is positively correlated with the likelihood of conducting business trips. It is as expected that there are fewer business trips in summer and Christmas.

Elasticities for the level-of-service attributes for train are derived to provide a first look into magnitudes of the impacts. Elasticities are calculated using sample enumeration. The elasticity shows the unit percentage change of the likelihood given a unit percentage change of a level-of-service attribute. The following scenarios are adopted for the elasticity calculations: 10% increase in travel cost by train and 10% decrease in train in-vehicle time. The results are presented in Table 3.

Table 3: Elasticity results of changes in level-of-service attributes for train for business trips.

		Car	Bus	Train	Air	Total
Baseline	Likelihood	0.468%	0.076%	0.068%	1.210%	1.823%
10% increase travel cost by train	Likelihood	0.469%	0.076%	0.065%	1.213%	1.823%
	Elasticity	0.021	0.020	-0.475	0.017	0.000
10% decrease in train in-vehicle time	Likelihood	0.466%	0.076%	0.075%	1.205%	1.823%
	Elasticity	-0.038	-0.041	1.069	-0.043	0.000

The elasticity of increased train travel cost is -0.475, which is similar to that of private trips. The business elasticity of decreased train in-vehicle time is much higher than that of private trips, 1.069, suggesting that business travellers are more inclined to take high-speed trains due to the travel time saving. Since the logsum variable is not significant and not included in the trip generation model for business trips, changes in level-of-service variables will not result in a change in the overall likelihood of business trip generation.

4. CONCLUSIONS

Long-distance international travel, although low in number of trips compared to regional travel, contributes significantly to total distance travelled and thus externalities from the transport sector. Despite the abundant literature on analysing tourist demand and long-distance travel, most developed models are direct demand models that focus on a specific mode or specific origin-destination pair. The absence of such disaggregated models indicates a lack of ability to calculate modal shift for long-distance international travel for large infrastructure investments such as high-speed rail.

In this study trip generation, mode and destination choice are modelled in Multinomial Logit models and Nested Logit models respectively. Swedish national travel survey data is used as observations of long-distance cross-border travel. European networks for road, train, and ferry and a world-wide network for air are developed at a reasonable level of detail. Models for private and business trips are developed where the ones for private trips are further segmented by number of nights away. The estimation results reveal the effects of individual socio-economic variables, level-of-service attributes, and destination variables. Income and access to car in household are found important explanatory factors in trip generation models for business trips. The derived VOT suggest that VOT for long-distance cross-border travel may differ significantly from VOT for domestic long-distance travel.

Elasticities of level-of-service attributes for trains are also derived to provide a first impression of high-speed rail scenarios. The most elastic attribute for private long-distance cross-border trips is travel cost, while for business long-distance cross-border trips it is in-vehicle time. The induced demand, i.e., those who previously did not conduct a long-distance cross-border trip and now travel by train due to the improved train service is however found to be neglectable.

ACKNOWLEDGEMENTS

The research has been conducted within the research project “Forecast models for international travel” funded by the Swedish Transport Administration under grant TRV 2019/98241.

REFERENCES

- Divisekera, S. (2010). Economics of leisure and non-leisure tourist demand: A study of domestic demand for Australian tourism. *Tourism Economics*, *16*(1), 117–136.
- Gelhausen, M. C., Berster, P., & Wilken, D. (2018). A new direct demand model of long-term forecasting air passengers and air transport movements at German airports. *Journal of Air Transport Management*, *71*, 140–152.
- Kim, S., & Shin, D. H. (2016). Forecasting short-term air passenger demand using big data from search engine queries. *Automation in Construction*, *70*, 98–108.
- Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, *59*, 57–66.
- Pieters, M., de Jong, G., & van der Hoorn, T. (2012). Cross-border car traffic in Dutch mobility models. *European Journal of Transport and Infrastructure Research*, *12*(2).
- Rich, J., Bröcker, J., Overgård, C. H., Korzenewych, A., Nielsen, O. A., & Vuk, G. (2009). *Report on scenario, traffic forecast and analysis of traffic on the ten-t, taking into consideration the external dimension of the union: Trans-tools version 2; model and data improvements*. <https://www.semanticscholar.org/paper/Report-on-Scenario%2C-Traffic-Forecast-and-Analysis-Rich-Broecker/eb43addc54da5362319169437188db7c09260214>
- Rich, J., & Mabit, S. L. (2012). A long-distance travel demand model for Europe. *European Journal of Transport and Infrastructure Research*, *12*(1), 1–20. http://orbit.dtu.dk/fedora/objects/orbit:42232/datastreams/file_6461773/content
- Santana-Jiménez, Y., & Hernández, J. M. (2011). Estimating the effect of overcrowding on tourist attraction: The case of Canary Islands. *Tourism Management*, *32*(2), 415–425.

- Suh, D. Y., & Ryerson, M. S. (2019). Forecast to grow: Aviation demand forecasting in an era of demand uncertainty and optimism bias. *Transportation Research Part E: Logistics and Transportation Review*, *128*, 400–416.
- TRT Trasporti e Territorio. (2018). *Description of the Trust model*. <http://www.trt.it/wp/wp-content/uploads/2016/09/TRUST-model-detailed-description-1.pdf>
- Witlox, F., Zwanikken, T., Jehee, L., Donners, B., & Veeneman, W. (2022). Changing tracks: Identifying and tackling bottlenecks in European rail passenger transport. *European Transport Research Review*, *14*(1), 1–12.

On the pricing of pool ride-hailing services in bus lanes

Lynn Fayed*¹, Gustav Nilsson¹, and Nikolas Geroliminis¹

¹Urban Transport Systems Laboratory (LUTS) École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. {lynn.fayed,gustav.nilsson,nikolas.geroliminis}@epfl.ch

SHORT SUMMARY

Ride-hailing vehicles contribute to traffic congestion in urban areas, where spatial constraints and uneven multi-modal distribution of infrastructure are a constant problem. In fact, roaming empty vehicles cause additional delays to other concurrent network users without delivering passengers to their destinations. Ride-splitting is one potential solution to counteract the negative impact of ride-hailing on traffic. In this work, we provide a dynamic non-equilibrium modelling framework for ride-splitting, where pool passengers are allowed to use dedicated bus lanes and can potentially travel faster than solo users. The objective is to develop a ride-splitting pricing policy between solo and pool options to encourage trip sharing with the goal of minimizing overall delays in multi-modal networks with bus lanes. Therefore, a Model Predictive Control (MPC) framework is set forward to investigate the price difference between the two ride-hailing alternatives with the objective of reconsidering the space allocation between the available modes. The results show that the proposed strategy is able to adjust the time-dependent fare changes based on the multi-modal demand and speeds in different parts of the network.

Keywords: Model predictive control, multi-modal networks, network delays, public transportation, regulations, ride-splitting services, space allocation.

1. INTRODUCTION

Ride-hailing has quickly established itself as a transportation alternative in its own because of the myriad of benefits it brings. It is a flexible and affordable door-to-door service with relatively low wait times compared to public transit. Despite its advantages, the impact of ride-hailing services on multi-modal demand distribution and traffic congestion is significant at multiple levels. Ride-splitting has the potential to mitigate these impacts by reducing the number of drivers required to achieve the same level of service (Ma, Zheng, & Wolfson, 2015), and cutting down the number of Vehicle Kilometer Travelled (VKT) to serve the same level of demand (Ke, Zheng, Yang, & Ye, 2021). However, the critical mass for pooling is rarely fulfilled, and this is due to the longer travel time that this option entails. The need to regulate ride-hailing services is therefore becoming more substantiated, and much work in this area has examined the efficiency of setting a cap on the fleet size or the maximum empty VKT allowed to be travelled by ride-hailing drivers (Yu, Tang, Max Shen, & Chen, 2020; Schaller, 2018). Other lines of research address the gains from sending empty ride-hailing vehicles to available parking spaces in urban areas in dynamic non-equilibrium settings (Beojone & Geroliminis, 2021), or in static equilibrium frameworks (Li, Qin, Yang, Poolla, & Varaiya, 2020; Xu, Yin, & Zha, 2017). The commonality among these proposed solutions is that they all fall into an enforcement-based regulatory approach. In a previous paper, Fayed, Nilsson, and Geroliminis (2023) demonstrated the potential of an incentive-based policy to

reduce the impact of ride-hailing on traffic congestion. By proposing an occupancy-based spatial assignment policy in a multi-modal network with dedicated bus lanes, we showed that it is possible to mitigate the impact of ride-hailing by encouraging users to share their trips. By giving pool drivers the right to utilize the bus lanes, the number of vehicles in the mixed traffic portion of the network is reduced. A similar modal-dependent allocation strategy is used for Autonomous Vehicles (AV) in Lamotte, de Palma, and Geroliminis (2017), or for buses in Geroliminis and Daganzo (2008); Geroliminis, Zheng, and Ampountolas (2014). Nevertheless, a very high number of pool vehicles in bus lanes deteriorates network conditions by causing significant delays for bus users. A pricing strategy is hence needed to encourage pooling when bus lane capacity allows it, and to deter pooling when bus delays are significant. Therefore, the main contribution of this work is to develop an aggregate multi-modal dynamic model that incorporates the proposed allocation policy. This model serves as a basis for establishing a control pricing scheme between the two ride-hailing alternatives in an attempt to contain overall network delays.

2. METHODOLOGY

The following section deals with the dynamic macroscopic modelling of multi-modal networks with bus lanes. Within this scope, we assume that private vehicles and solo ride-hailing users travel in the vehicle network, while buses and pool ride-hailing users travel in the bus network. We begin first by characterizing the aggregate traffic model that we use to determine speeds in both networks. We then use this model to describe the dynamics of private and ride-splitting vehicles, and bus passengers. Finally, we introduce the MPC scheme with the goal of determining the pricing gap between the solo and pool alternatives that minimizes multi-modal user delays. The full list of notations used in this paper is displayed in Table 1.

Network model

In the network under consideration, travellers perform their trips by one of the set of available modes \mathcal{M} : private vehicles pv , ride-splitting rs , or buses b , so that $\mathcal{M} := \{pv, b, rs\}$. Therefore, the exogenous and time-dependent hourly demand for each mode is given by $Q_j(k)$ for $j \in \mathcal{M}$ where k is the time-step such that $k \in \mathcal{K} := \{0, \dots, k_{max}\}$. Moreover, we assume that the time interval between two consecutive time-steps is equal to Δ . Ride-hailing users choose either a solo trip s in the vehicle network \mathcal{V} or a pool trip p in the bus network \mathcal{B} . The fraction of infrastructure allocated to the vehicle network \mathcal{V} is denoted by α where $\alpha \in [0, 1]$. It can be inferred that buses and pool users are allocated a space equal to $1 - \alpha$ of the total available network infrastructure. The ride-splitting fleet N is constant, and drivers belong exclusively to one of the following categories at each time-step:

- i) idling or dispatching with no passengers inside the vehicle that we denote by n_e ,
- ii) performing a solo trip in the vehicle network \mathcal{V} which we denote by n_s , and
- iii) performing a pool trip in the bus network \mathcal{B} which we denote by n_p .

In addition to the ride-splitting fleet, we denote the number of private vehicles in network \mathcal{V} at time-step $k \in \mathcal{K}$ by $n_{pv}(k)$, and the number of buses in the bus network \mathcal{B} by n_b . Note that the number of buses is assumed to be constant, but the bus occupancy o_b is time-dependent. Accordingly, the total accumulation in the vehicle network $n_{\mathcal{V}}$ at time-step $k \in \mathcal{K}$ is $n_{\mathcal{V}}(k) = n_{pv}(k) + n_e(k) + n_s(k)$ where the idle ride-hailing vehicles travel exclusively in the vehicle network. Similarly, the accumulation in the bus network $n_{\mathcal{B}}$ at time-step $k \in \mathcal{K}$ is $n_{\mathcal{B}}(k) = n_p(k) + n_b$. The fleet size N of ride-hailing vehicles is constant and provided by the platform operator such that $N = n_e(k) + n_s(k) + n_p(k)$ for all $k \in \mathcal{K}$.

Table 1: Notations

Variable name	Description
\mathcal{M}	Set of transportation modes where $\mathcal{M} = \{pv, rs, b\}$
$n(k)$	Total vehicle accumulation at time-step k
$P(n(k))$	Total production in the network at time-step k
$v(k)$	Network speed at time-step k
α	Fraction of space allocated to \mathcal{V}
$n_{\mathcal{V}}(k), n_{\mathcal{B}}(k)$	Accumulation in \mathcal{V} and \mathcal{B} respectively at time-step k
N	Ride-hailing fleet size
$n_{pv}, n_e(k), n_s(k), n_p(k)$	Number of private, idle, solo and pool ride-hailing vehicles respectively at time-step k
n_b	Number of buses in the bus network
$P_{\mathcal{V}}(n_{\mathcal{V}}(k)), P_{\mathcal{B}}(n_{\mathcal{B}}(k))$	Production in \mathcal{V} and \mathcal{B} respectively at time-step k
$P_p(n_p(k), n_b)$	Production of pool vehicles in \mathcal{B} at time-step k
$P_b(n_p(k), n_b)$	Production of buses in \mathcal{B} at time-step k
$v_{\mathcal{V}}(n_{\mathcal{V}}(k)), v_{\mathcal{B}}(n_{\mathcal{B}}(k))$	Speed in \mathcal{V} and \mathcal{B} respectively at time-step k
$v_p(n_p(k), n_b)$	Speed of pool vehicles at time-step k
$v_b(n_p(k), n_b)$	Speed of buses at time-step k
$Q_i(k)$	Demand for mode $i \in \mathcal{M}$ at time-step k
$c(k)$	Number of ride-hailing customers waiting to be assigned at time-step k
$\beta(k)$	Fraction of ride-hailing requests opting for solo trips at time-step k
$c_s(k), c_p(k)$	Number of ride-hailing customers opting for a solo and a pooled trip respectively at time-step k
$u_s(k), u_p(k)$	Travel time cost for solo and pooled respectively expressed in monetary terms at time-step k
$U_s(k), U_p(k)$	Total cost for a solo and pooled respectively at time-step k
$\bar{F}_s(k), \bar{F}_p(k)$	Dynamic fare for pooled and solo trips at time-step k
F_s, F_p	Ride-hailing basic fare for solo and pooled trips at time-step k
N	Ride-hailing fleet size
$o_b(k)$	Actual occupancy per bus at time-step k
\bar{o}_{pv}, \bar{o}_p	Average private and pool vehicles occupancy respectively
\bar{v}_b	Target speed for buses
$\bar{l}_{pv}, \bar{l}_s, \bar{l}_b$	Average private vehicle, solo, and bus trip lengths respectively
$\Delta d, \Delta p$	Driver and passenger pooled trip detour distance
Δ	Length of simulation time-step
$M(k)$	Matching rate at time-step k
μ	Mode choice scale parameter
κ	Value of time
a_0, α_e, α_c	Meeting function parameters
\bar{t}_d	Bus dwell time at stops
\bar{s}	Average spacing between bus stops
w_{\max}	Maximum passenger waiting time
$\xi(k)$	Control variable integrated in discrete mode choice at time-step k
$\phi(k)$	Additional controlled fee or discount that pooled vehicle incur at time-step k
$A(k)$	Ride-splitting request abandonment at time-step k

Traffic dynamics

In the following section, we address the macroscopic traffic dynamics that allow us to specify the relationships between demand and accumulation for each mode. Thus, let $P : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ be the total network production, n its accumulation, and v its speed, then we know that for all time-steps, the relation $P(n(k)) = n(k)v(n(k))$ holds, where $\frac{\partial v}{\partial n} \leq 0$. If the entire network space is partitioned into a vehicle and bus network according to the fractional split α , then the production functions in the vehicle network P_V and bus network P_B satisfy the conditions $\alpha P(n(k)) = P_V(\alpha n(k))$ and $\bar{\alpha} P(n(k)) = P_B(\bar{\alpha} n(k))$ respectively where $\bar{\alpha} = 1 - \alpha$ (Ni & Cassidy, 2019; Sirmatel, Tsitsokas, Kouvelas, & Geroliminis, 2021). Similarly, the speed in the vehicle network v_V and in the bus network v_B are given by $v(n(k)) = v_V(\alpha n(k))$ and $v(n(k)) = v_B(\bar{\alpha} n(k))$ respectively. Rewriting the production functions in terms of speeds, we obtain that $P_V(n_V(k)) = n_V(k)v_V(n_V(k))$ and $P_B(n_B(k)) = n_B(k)v_B(n_B(k))$.

To account for the fact that the marginal effects on the speed of buses and pool vehicles are not equivalent, we divide the bus network production into pool vehicle production P_b and a bus production P_b . This is because buses, unlike pool vehicles, frequently stop at stations to allow passengers to board and disembark. We capture this recurrent action by reducing the speed in the bus network v_B by a factor $r(n_b)$ where $r : \mathbb{R}_{\geq 0} \rightarrow (0, 1]$ and $\frac{dr}{dn_b} < 0$. Consequently, the running speed of the pool vehicles in the bus network is $v_p(n_p(k), n_b) = v_B(n_B(k))r(n_b)$. Taking into account the time that buses spend boarding and alighting passengers, the operational bus speed is

$$v_b(n_p(k), n_b) = \left(\frac{1}{1 + v_p(n_p(k), n_b) \frac{\bar{t}_d}{\bar{s}}} \right) v_p(n_p(k), n_b), \quad (1)$$

where \bar{t}_d and \bar{s} are the average time of buses and the spacing between stops, respectively. Therefore, after defining the individual pool vehicles and bus speeds, the production functions become $P_p(n_p(k), n_b) = n_p(k)v_p(n_p(k), n_b)$ and $P_b(n_p(k), n_b) = n_b(k)v_b(n_p(k), n_b)$ respectively for all $k \in \mathcal{K}$.

Private vehicles dynamics

According to the modal-dependent spatial allocation policy proposed in this framework, private vehicles utilize the vehicle network \mathcal{V} . The change in the accumulation of private vehicles between any two successive time-steps is given by the difference between the exogenous arrival of private vehicle users Q_{pv} and the completion rate of private vehicle trips O_{pv} . The latter is derived based on the accumulation n_V and the network production function P_V such that n_{pv} at time-step k is equal to

$$\begin{aligned} n_{pv}(k) &= n_{pv}(k-1) + \Delta \left[\frac{Q_{pv}(k)}{\bar{o}_{pv}} - O_{pv}(k-1) \right] \\ &= n_{pv}(k-1) + \Delta \left[\frac{Q_{pv}(k)}{\bar{o}_{pv}} - \frac{n_{pv}(k-1)}{n_V(k-1)} \frac{P_V(n_V(k-1))}{\bar{l}_{pv}} \right], \quad \forall k \in \mathcal{K} \setminus \{0\}, \end{aligned} \quad (2)$$

where \bar{l}_{pv} is the average trip length of private vehicles and \bar{o}_{pv} is their average occupancy. Note that the accumulation n_V itself depends on the number of private vehicles n_{pv} , but also on the number of solo ride-hailing drivers n_s .

Ride-splitting dynamics

As highlighted earlier, the total ride-splitting demand in this work is exogenous, but the demand split between solo and pool trips is the result of user choice, and is determined endogenously using the total travel cost for each alternative. This cost is the sum of two elements: the alternative-dependent travel fare and the travel time in each network. Accordingly, if $U_s(k)$ and $U_p(k)$ are the

disutilities for a solo or pool trip at time-step $k \in \mathcal{K}$ respectively, then their expressions are given by

$$U_s(k) = \tilde{F}_s(k) + \kappa \frac{\bar{l}_s}{v_{\mathcal{V}}(n_{\mathcal{V}}(k))}, \quad (3)$$

$$U_p(k) = \tilde{F}_p(k) + \kappa \frac{\bar{l}_s + \Delta l_p}{v_p(n_p(k), n_b)}, \quad (4)$$

where κ is the value of time, \bar{l}_s is the average trip length for a solo trip, and Δl_p is the additional travel distance that pool users incur to pick up and/or drop off another passenger. The variable $\tilde{F}_s(k)$ represents the fare collected from a solo user and $\tilde{F}_p(k)$ represents the fare collected from a pool user at time-step k . For simplicity, we assume that the solo fare is static such that $\tilde{F}_s(k) = F_s$ for all $k \in \mathcal{K}$ where F_s is the fare set by the platform operator. The pool fare is defined by the expression $\tilde{F}_p = F_p + \phi(k)$ where F_p is the static pool charges collected by the platform, and $\phi(k)$ is the control fare set by the network regulator to steer the system toward its optimum.

If the choice of ride-hailing users is the outcome of a binary logit choice model, then the fraction of solo trips that we denote by β is

$$\beta(k) = \frac{\exp(-\mu U_s(k))}{\exp(-\mu U_s(k)) + \exp(-\mu U_p(k))}, \quad (5)$$

where $\mu > 0$ is the binary logit scale parameter. Rewriting the disutilities for solo and pool that we denote by u_s and u_p respectively in terms of static fare only, we obtain the following

$$u_s(k) = F_s + \kappa \frac{\bar{l}_s}{v_{\mathcal{V}}(n_{\mathcal{V}}(k))}, \quad (6)$$

$$u_p(k) = F_p + \kappa \frac{\bar{l}_s + \Delta l_p}{v_p(n_p(k), n_b)}. \quad (7)$$

Therefore, the expression for β as function of u_s and u_p is

$$\beta(k) = \frac{\exp(-\mu u_s(k))}{\exp(-\mu u_s(k)) + \underbrace{\exp(-\mu \phi(k)) \exp(-\mu u_p(k))}_{\xi(k)}}, \quad (8)$$

where $\xi(k) \in (0, +\infty)$ is an auxiliary variable that paves the way for a more pragmatic implementation of the MPC framework. From the auxiliary variable $\xi(k)$, it is straightforward to derive the control price $\phi(k)$ by using the expression

$$\phi(k) = \frac{\log(\xi(k))}{-\mu}. \quad (9)$$

Once the choice of users is determined, it is straightforward to divide the waiting passengers into two categories. If $c(k)$ is the number of requests waiting to be assigned at time k , then the number of requests that choose to ride solo is $c_s(k) = \beta(k)c(k)$ and those that choose to pool is $c_p(k) = (1 - \beta(k))c(k)$. These passengers are matched to idling vehicles according to the following bilateral meeting rate M , so that

$$M(k) = a_0 n_e(k)^{\alpha_e} \left(c_s(k) + \frac{1}{2} c_p(k) \right)^{\alpha_c}, \quad (10)$$

where $a_0 > 0$, $\alpha_e > 0$, and $\alpha_c > 0$ are the parameters of the Cobb-Douglas meeting function in (10). Note that a factor $\frac{1}{2}$ is added to c_p to indicate that pool waiting passengers are assigned to a single unoccupied vehicle.

Once all these elements are defined, it becomes possible to determine the dynamics of the ride-splitting market. We first start with idle vehicles, so that the value of n_e at time-step k is given by

$$\begin{aligned} n_e(k) &= n_e(k-1) + \Delta[O_s(k-1) + O_p(k-1) - M(k-1)] \\ &= n_e(k-1) + \Delta \left[\frac{n_s(k-1)}{n_V(k-1)} \frac{P_V(n_V(k-1))}{\bar{l}_s} + \frac{P_p(n_p(k-1), n_b)}{\bar{l}_s + \Delta l_d} - M(k-1) \right], \forall k \in \mathcal{K} \setminus \{0\}, \end{aligned} \quad (11)$$

where O_s and O_p are the completion rates for a solo and a pool trip respectively, \bar{l}_s is the average solo trip length, and Δl_d is the driver detour, i.e., the extra distance that a driver travels to deliver a pool trip. Note that all drivers who complete their trips become idling again, and thus represent the inflow for the idling vehicle category. In contrast, the outflow for this specific category consists of every vehicle that has been matched, which is derived from $M(k)$.

Moving to the solo vehicle category, the number of solo vehicles at every time-step is

$$n_s(k) = n_s(k-1) + \Delta \left[\beta(k-1)M(k-1) - \frac{n_s(k-1)}{n_V(k-1)} \frac{P_V(n_V(k-1))}{\bar{l}_s} \right], \quad \forall k \in \mathcal{K} \setminus \{0\}. \quad (12)$$

Since only a fraction β of the total ride-hailing requests choose a solo trip, the inflow for the solo vehicle category is determined by a portion β of the total matching rate, and the outflow is computed using vehicle network production P_V .

In a similar manner, we calculate the change in the number of pooling vehicles n_p by taking the difference between the number of drivers matched to a pool trip and the number of pool drivers that completed their trips. Thus, the expression of n_p is given by

$$n_p(k) = n_p(k-1) + \Delta \left[(1 - \beta(k-1))M(k-1) - \frac{P_p(n_p(k-1), n_b)}{\bar{l}_s + \Delta l_d} \right], \quad \forall k \in \mathcal{K} \setminus \{0\}. \quad (13)$$

The trip completion rate for the pool vehicle category is computed using the pool vehicle production P_p , which itself is a function of the time-dependent number of pool vehicles n_p and the static number of buses n_b .

Finally, the dynamics of waiting passengers remain to be defined. Knowing that the demand for ride-splitting at time-step k is given by $Q_{rs}(k)$, the number of waiting passengers $c(k)$ is

$$c(k) = c(k-1) + \Delta[Q_{rs}(k) + (\beta(k-1) - 2)M(k-1)] - A(k), \quad \forall k \in \mathcal{K} \setminus \{0\}. \quad (14)$$

where $A(k)$ is the number of abandoning requests that are not served within reasonable waiting times. We point out here that the calculation of the passenger outflow takes into account that each pool trip results in two passengers leaving the queue of waiting requests. The number of abandonments $A(k)$ is computed as follows

$$A(k) = \max \left(c(k-1) - \frac{1}{k-1} \sum_{\tilde{k}=1}^{k-1} M(\tilde{k}-1)w_{\max}, 0 \right), \quad (15)$$

where w_{\max} is a measure of the maximum wait tolerance of ride-hailing requests. This equation is an approximation of the number of waiting requests when the wait tolerance is set to w_{\max} . It computes the number of requests that leave the platform due to a poor level of service.

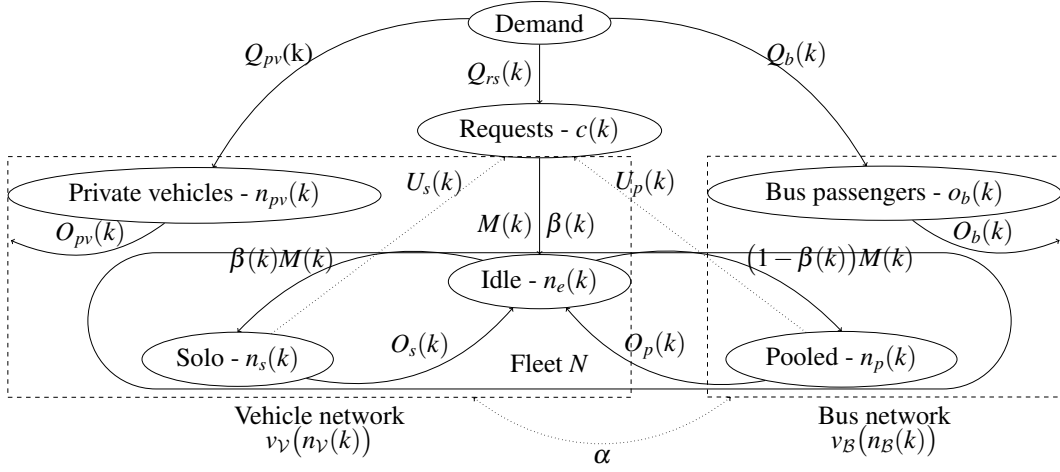


Figure 1: Modelling framework of the proposed dynamic and modal-dependent space allocation policy

Bus dynamics

Moving to the bus dynamics, we assume in this framework that the number of buses circulating in the bus network \mathcal{B} is constant. This number is computed using

$$n_b = \frac{\bar{Q}_b \bar{l}_b}{\bar{o}_b \bar{v}_b}, \quad (16)$$

such that \bar{Q}_b is the average expected bus demand per hour, \bar{l}_b is the average bus trip length, \bar{o}_b is the target bus occupancy, and \bar{v}_b is the expected bus operating speed. Assuming that n_b remains constant in time, we track the bus dynamics by following the variation of the average bus occupancy o_b , which at time k , is given by

$$o_b(k) = o_b(k-1) + \Delta \frac{1}{n_b} \left[Q_b(k) - \frac{P_b(n_p(k-1), n_b)}{\bar{l}_b} o_b(k-1) \right], \quad \forall k \in \mathcal{K} \setminus \{0\}. \quad (17)$$

We assume in (17) that bus demand $Q_b(k)$ is uniformly distributed over the available bus fleet. The trip completion rate is computed using the bus production function P_b , and is converted to passenger trips by multiplying the completion rate by the average bus occupancy.

Figure 1 summarizes the full network model including the dynamics that we previously described.

Model predictive control

The modelling framework proposed captures the full impact on the multi-modal commuters when pool users are allowed to use the bus lanes. In Fayed et al. (2023), it was shown that the overall network situation deteriorates when the number of pool vehicles in the bus lanes is relatively high, causing additional delays for buses. Therefore, a pricing scheme is needed to encourage or discourage ride-hailing users to utilize the bus lanes depending on the speeds in the vehicle network \mathcal{V} and in the bus network \mathcal{B} . To provide such a pricing scheme, we integrate the network dynamics into a MPC framework, and use the latter to determine the fare control variable $\phi(k)$ that minimizes the total Passenger Hour Travelled (PHT) such that PHT at time k is equal to the sum of the individual PHT of each commuter category in the network

$$\text{PHT}(k) = \Delta [n_{pv}(k) \bar{o}_{pv} + n_b(k) o_b(k) + n_s(k) + n_p(k) \bar{o}_p].$$

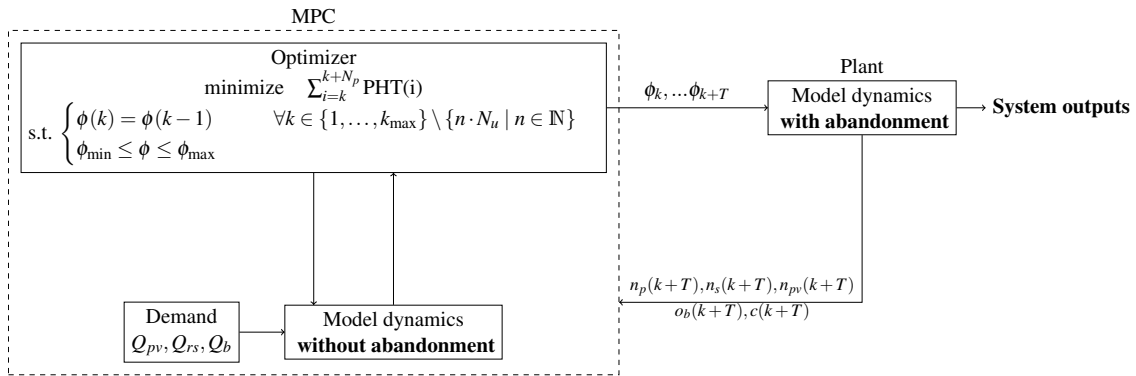


Figure 2: MPC framework with distinct plant and MPC dynamics

The variable \bar{o}_p is the average occupancy of the pool vehicles, where $\bar{o}_p \in [1, 2]$. The objective of MPC is to solve the following minimization problem

$$\begin{aligned}
& \text{minimize} && \sum_{k \in \mathcal{K}} \text{PHT}(k) \\
& \{\phi(k)\}_{k \in \mathcal{K}} && \\
& \text{subject to} && (1) - (17) \\
& && \phi(k) = \phi(k-1), \quad \forall k \in \{1, \dots, k_{\max}\} \setminus \{n \cdot N_u \mid n \in \mathbb{N}\} \\
& && \phi_{\min} < \phi(k) < \phi_{\max}, \quad \forall k \in \{1, \dots, k_{\max}\}
\end{aligned} \tag{18}$$

where the second constraint in (18) ensures that the control variable is updated only every $N_u \in \mathbb{N}$ time-steps. The third constraint sets a minimum and maximum range for the control variable.

The minimization framework presented in (18) is straightforward if we disregard abandonment from (14). Nevertheless, including abandonment in our framework introduces an additional computational complexity since $A(k)$ is calculated based on prior matching rates. To circumvent this, we adapt the MPC framework according to the approach displayed in Figure 5. In this scope, we neglect the abandonment in the MPC dynamics and solve the optimization problem to obtain the values of the control fare for a prediction horizon $N_p \in \mathbb{N}$. However, we extract only the control variables up to an update horizon, such that $T \in \mathbb{N}$ and $T < N_p$, and incorporate them into the actual plant dynamics with abandonment to estimate the new state variables. The state variables are then fed as inputs to the MPC dynamics and the optimizer is relaunched.

3. RESULTS AND DISCUSSION

In this section, we will numerically show the performance of the proposed control strategy and quantify the influence of abandonment on the model dynamics. To do so, we consider a network which production function is given by $P(n) = A_0 n^3 + B_0 n^2 + C_0 n$, such that $A_0 = 5.74 \cdot 10^{-9}$, $B_0 = -1.02 \cdot 10^{-3}$, and $C_0 = 36$ for $n \in [0, 58536]$. Assuming the fraction of the total space allocated to the vehicle network α is equal to 0.8, it becomes straightforward to derive the production functions P_V and P_B . The function capturing the marginal effect of buses on traffic is given by $r(n_b)$ such that $r(n_b) = e^{-6.5 \cdot 10^{-4} n_b}$. This function, in addition to $\bar{s} = 0.8$ km and $\bar{t}_d = 30$ s, are used to compute the bus speed v_b .

Next, we list the different constant values used in the modelling framework. The average private vehicle and solo trip lengths are set to be equal such that $\bar{l}_{pv} = \bar{l}_s = 3.86$ km. The trip length by bus is generally larger, and therefore \bar{l}_b is set to be equal to $1.4\bar{l}_{pv}$. The values of the driver and passenger detour are $\Delta l_d = 0.7\bar{l}_s$ and $\Delta l_p = 0.15\bar{l}_s$, and the average occupancies for private vehicles \bar{o}_{pv} , pool vehicles \bar{o}_p , and bus \bar{o}_b are equal to 1.2, 1.5, and 20 respectively. The average design bus speed \bar{v}_b used to compute the number of buses n_b is 18 km/hr. The a_0 , α_e , and α_c parameters for

Table 2: PHT for different simulation frameworks with no abandonment

Scenario	β	N_p	T	PHT [pax.km/hr]
Benchmark	logit	-	-	171300
Pool trips only	0	-	-	182037
Solo trips only	1	-	-	175798
MPC	logit	3600	3600	168437
MPC	logit	900	450	168729

Table 3: PHT for different simulation frameworks with abandonment

Scenario	β	N_p	T	PHT [pax.km/hr]	Abandonment
Benchmark	logit	-	-	170007	8590
Pool trips only	0	-	-	177376	20463
Solo trips only	1	-	-	174952	21626
MPC	logit	3600	3600	167635	10271
MPC	logit	900	450	167938	13961

the Cobb-Douglas matching function are 0.025, 0.93, and 0.98 respectively. With respect to the computation of the mode choice between solo and pool, we set the scale parameter $\mu = 1$ and the value of time $\kappa = 30$ CHF/hr. The constant solo and pool trip fare F_s and F_p are equal to 5 and 4 CHF respectively. The ride-hailing fleet size N is equal to 3500 and the discretized time interval Δ is equal to 6 s.

The aggregate simulation that we advance spans over a 6-hour period covering the afternoon on-peak in-between two off-peak periods. The demand profiles for private vehicles, ride-splitting services, and buses are shown in Figure 3. To start with, we test our controller for a scenario where abandonment is always set to 0, and we show the results in Table 2. We compare different simulation settings, including a benchmark scenario with no external intervention and an MPC framework with different prediction and update horizon settings, with $N_u = 180$. The same simulations are then repeated when the plant (but not the dynamics in the MPC) has abandonment rates as given by (15), and the results are displayed in Table 3. Logically, the values of PHT are lower for scenarios with abandonment mainly because the travel time of abandoning requests is not accounted for in the computation of PHT. Moreover, irrespective of the abandonment settings, the MPC with a full prediction horizon returns the lowest PHT, and decreasing the prediction horizon slightly deteriorates the results. Finally, for instances with abandonment, even if the MPC framework reduces delays, the increase in abandonment relative to the benchmark scenario does not allow a fair comparison basis. Therefore, designing an MPC that accommodates abandonment in its dynamic framework is a research area that we plan to investigate in the future.

Lastly, we plot the variations of the main model variables for the benchmark scenario with no abandonment in Figure 4, and for the MPC framework with abandonment and a prediction horizon of 900 in Figure 5. Compared to a scenario with no control, i.e., $\phi(k) = 0$ for all $k \in \mathcal{K}$, the MPC imposes high charges for pooling to prevent the deterioration of the conditions in bus lanes, and to improve the overall network situation. During on-peak periods however, we observe that the charges decrease, and the control variable ϕ assumes negative values to further encourage pooling as the speed in the vehicle network v_V drops significantly. The aim in these particular time spots is to alleviate congestion in the vehicle network.

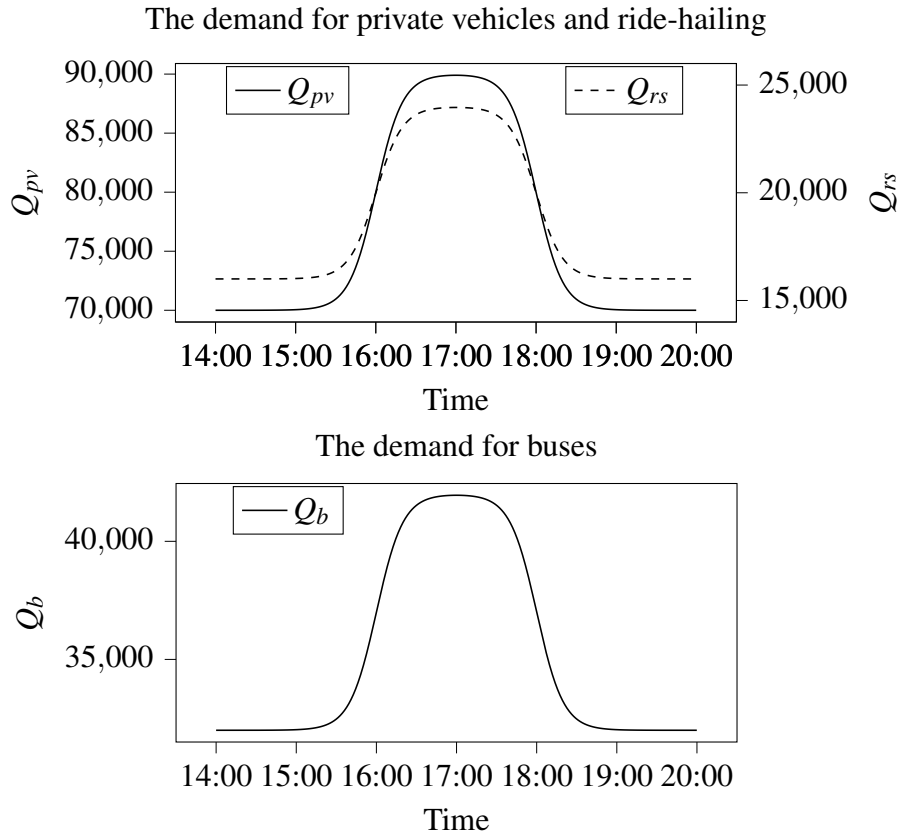


Figure 3: Demand profiles for Q_{pv} , Q_{rs} and Q_b

4. CONCLUSIONS

In this work, we designed a dynamic space and occupancy-dependent allocation policy for multi-modal networks with ride-splitting services. We then build upon this model to construct a pricing control strategy that encourages or discourages pooling in bus lanes according to the overall user delays. Using an MPC framework, we solve the minimization problem, and determine what is the additional fare or discount that ride-splitting users should incur to improve the total network situation. By comparing many different scenarios, we demonstrate that our control scheme is indeed capable of reducing total Passenger Hour Travelled in scenarios with or without ride-hailing request abandonment. Future research direction however will focus on a more reasonable integration of abandoning requests within the MPC framework through reintroducing them to the system as bus or private vehicle users.

ACKNOWLEDGMENT

This work was supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

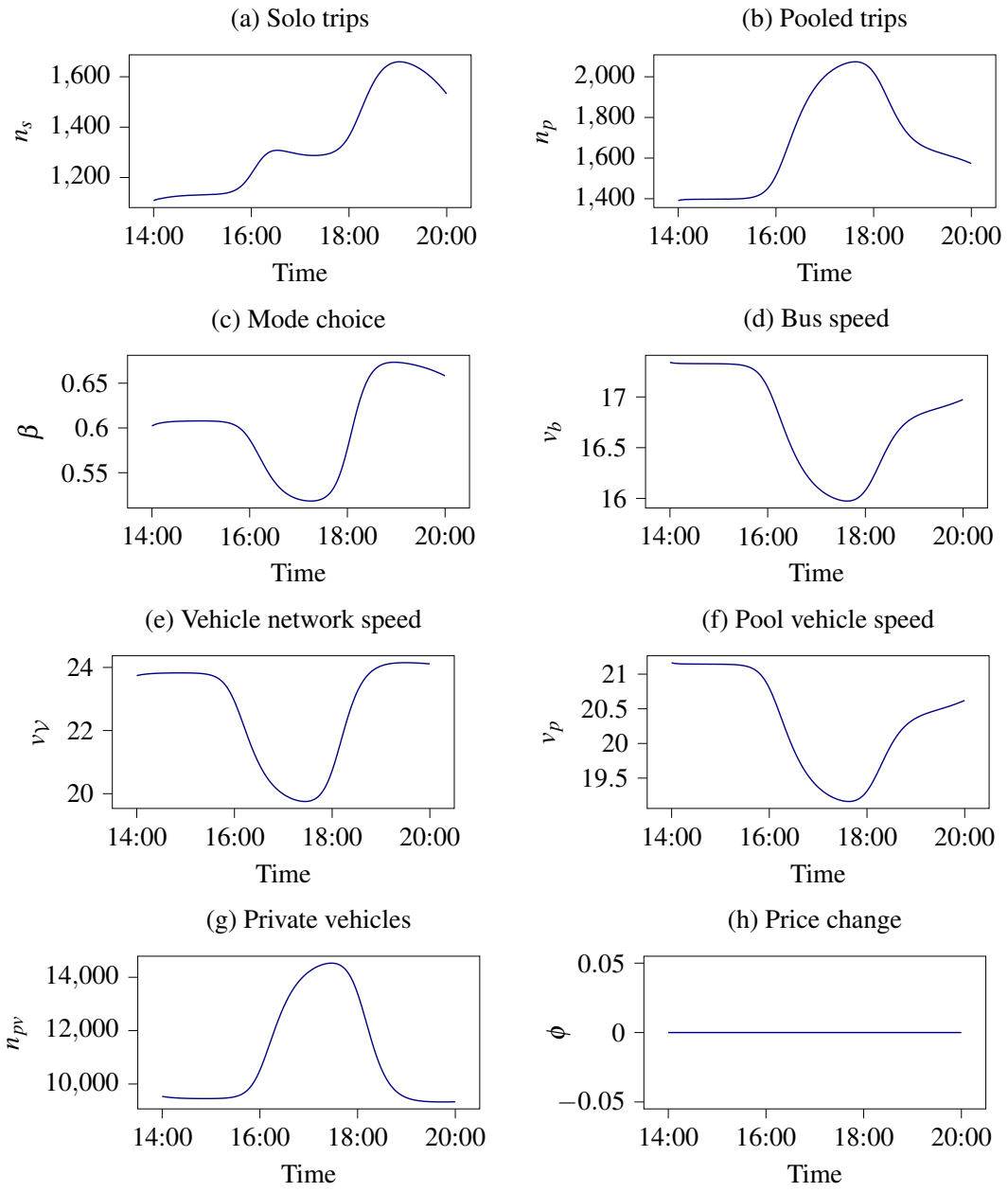


Figure 4: Time-dependent model variables for the benchmark scenario with no request abandonment

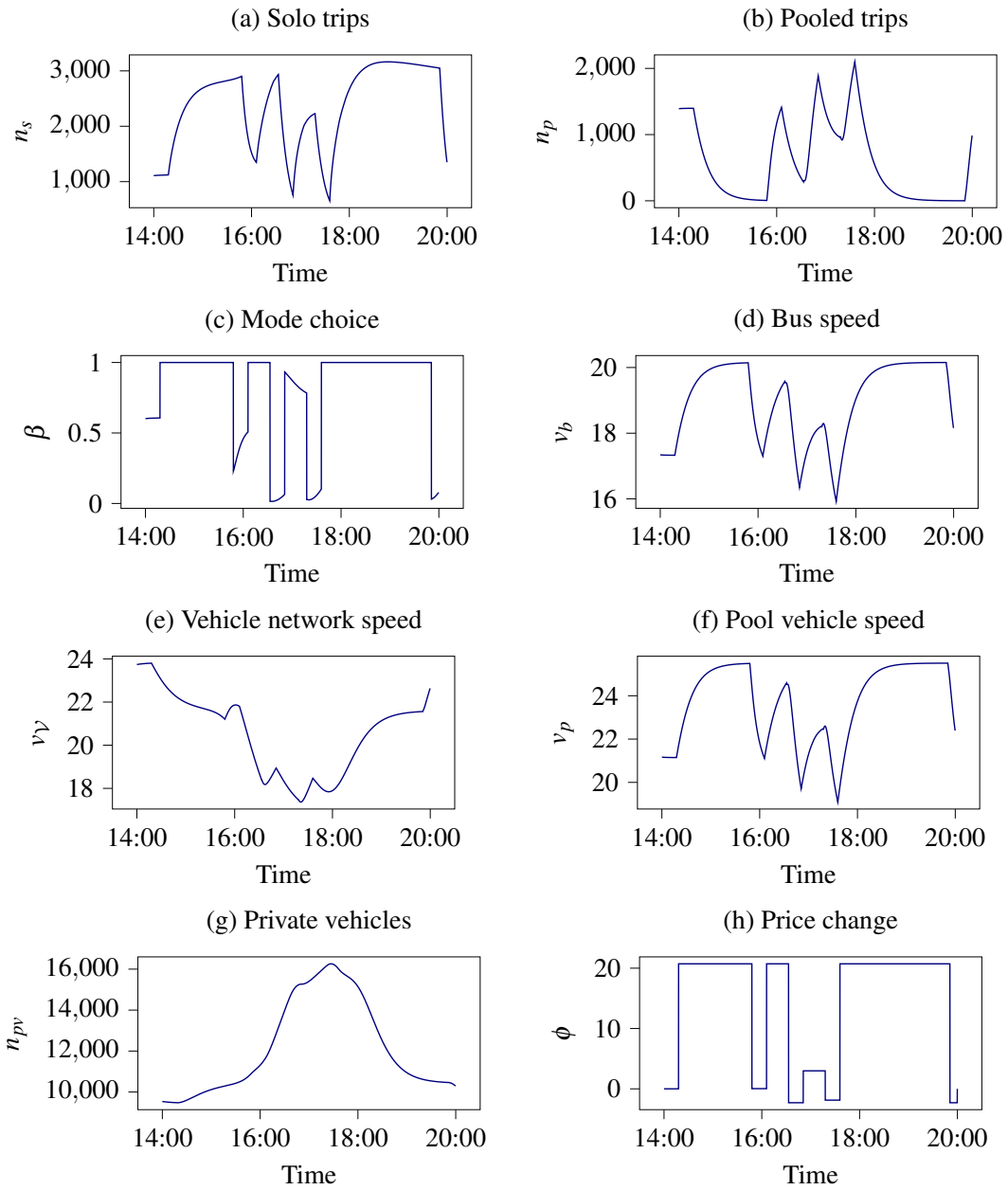


Figure 5: Time-dependent model variables for the MPC scenario with abandonment, $N_p = 900$, and $T = 450$

REFERENCES

- Beojone, C. V., & Geroliminis, N. (2021). On the inefficiency of ride-sourcing services towards urban congestion. *Transportation Research Part C: Emerging Technologies*, 124, 102890. doi: 10.1016/j.trc.2020.102890
- Fayed, L., Nilsson, G., & Geroliminis, N. (2023). On the utilization of dedicated bus lanes for pooled ride-hailing services. *Transportation Research Part B: Methodological*, 169, 29-52. doi: 10.1016/j.trb.2023.01.005
- Geroliminis, N., & Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9), 759-770. doi: 10.1016/j.trb.2008.02.002
- Geroliminis, N., Zheng, N., & Ampountolas, K. (2014). A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks. *Transportation Research Part C: Emerging Technologies*, 42, 168-181. doi: 10.1016/j.trc.2014.03.004
- Ke, J., Zheng, Z., Yang, H., & Ye, J. (2021). Data-driven analysis on matching probability, routing distance and detour distance in ride-pooling services. *Transportation Research Part C: Emerging Technologies*, 124, 102922. doi: 10.1016/j.trc.2020.102922
- Lamotte, R., de Palma, A., & Geroliminis, N. (2017). On the use of reservation-based autonomous vehicles for demand management. *Transportation Research Part B: Methodological*, 99, 205-227. doi: 10.1016/j.trb.2017.01.003
- Li, S., Qin, J., Yang, H., Poolla, K., & Varaiya, P. (2020). Off-street parking for TNC vehicles to reduce cruising traffic. *2020 59th IEEE Conference on Decision and Control (CDC)*, 2585-2590.
- Ma, S., Zheng, Y., & Wolfson, O. (2015). Real-time city-scale taxi ridesharing. *IEEE Transactions on Knowledge and Data Engineering*, 27, 1782-1795.
- Ni, W., & Cassidy, M. (2019). City-wide traffic control: Modeling impacts of cordon queues. *Transportation Research Part C: Emerging Technologies*, 113. doi: 10.1016/j.trc.2019.04.024
- Schaller, B. (2018). *The new automobility: Lyft, Uber, and the future of American cities*.
- Sirmatel, I. I., Tsitsokas, D., Kouvelas, A., & Geroliminis, N. (2021). Modeling, estimation, and control in large-scale urban road networks with remaining travel distance dynamics. *Transportation Research Part C: Emerging Technologies*. doi: 10.1016/j.trc.2021.103157
- Xu, Z., Yin, Y., & Zha, L. (2017). Optimal parking provision for ride-sourcing services. *Transportation Research Part B: Methodological*, 105, 559-578. doi: 10.1016/j.trb.2017.10.003
- Yu, J. J., Tang, C. S., Max Shen, Z.-J., & Chen, X. M. (2020). A balancing act of regulating on-demand ride services. *Management Science*, 66(7), 2975-2992. doi: 10.1287/mnsc.2019.3351

Who is ready to live a car-independent lifestyle? A latent class cluster analysis of attitudes towards car ownership and usage

Jaime Soza-Parra*¹, Oded Cats²

¹ Assistant Professor in Transportation Analysis and Data Science, Human Geography and Spatial Planning Department, Utrecht University, The Netherlands

¹ External Researcher, Transportation & Planning Department, Delft University of Technology, The Netherlands

² Professor Passenger Transport Systems, Transportation & Planning Department, Delft University of Technology, The Netherlands

SHORT SUMMARY

The excess use of private cars for transportation has multiple negative effects on our society, and therefore, determining the underlying factors driving car usage among different groups of travellers could contribute to a more sustainable future. In this paper, we aim to identify and characterise traveller groups in terms of their car-related attitudes and how different sociodemographic attributes, behavioural characteristics (such as using cars as the primary mode of transportation), and their residential location accessibility vary amongst different population groups. Through Confirmatory Factor Analysis and Latent Profile Analysis we identify five different classes, namely *car detractors*, *hesitants*, *positives*, *friends*, and *lovers*. Overall, . We also see that the farther away households tend to be located from urban areas and public transportation facilities, the closer the relationship with cars. The results of this analysis will provide valuable insights into how to discourage the use of cars and promote more sustainable mobility.

Keywords: Car independence, car ownership, multi-modal transportation, latent profile analysis, sustainability

1. INTRODUCTION

Car usage, ranging from day-to-day commuting to weekend gateways, is playing a negative role in terms of traffic and climate change mitigation. This problematic situation is expected to continue increasing as road motor vehicles possession has been steadily growing in the last decade in OECD countries (OECD, 2022), in contrast to the hypothesis that car usage had reached a saturation point one decade ago (Goodwin & Van Dender, 2013). To address this problem, transport authorities improve and promote existing transportation alternatives such as public and active transportation options considering car drivers' perceptions (Abenoza et al., 2017; de Oña, 2022; Van Exel & Rietveld, 2010). The inherent idea behind this approach is that promoting better alternatives might make car users shift to more sustainable ways of mobility.

However, taking The Netherlands as an example, a country with excellent public transport infrastructure and plenty of safe and well-connected bike paths, private cars account for the majority (~40%) of the modal share in terms of the primary transport mode. Moreover, ~23% of those trips

are shorter than 2.5 km, and 40% are shorter than 5 km. In addition, there are several psychological aspects associated with the different dimensions related to travelling by car that influence the travel mode choice decision.

In this work, we aim to identify and characterise traveller groups in terms of their car-related attitudes and how different socio-demographic attributes, behavioural characteristics (such as using cars as the primary mode of transportation), and their residential location accessibility vary amongst different population groups.

2. METHODOLOGY

Our analysis is based on the Netherlands Mobility Panel (MPN) data, a representative sample of the adult Dutch population panel which every year gathers information at a personal, household, and a three-day travel diary (Hoogendoorn-Lanser et al., 2015). Given the Corona pandemic, we focused on the last even year before the start of the pandemic, 2018. In terms of the attitudinal questions, for each of the five modes analysed - car, train, bus/tram/metro (referred to as BTM, representing all urban public transport), bike, and walk - participants are asked about their overall opinion and about how they evaluate them in terms of being comfortable, relaxing, saving time, safe, flexible, pleasurable, and prestige. In addition, a set of 26 related to car usage and ownership experience were included in this specific wave in 2018. The final set of gathers 73 different attitudinal statements and 6,502 respondents answered all these questions in the sample. To handle the large number of attitudinal questions present in our dataset, we first carry out an exploratory factor analysis (EFA) to determine potential factors structures which we then use as a first step in composing a confirmatory factor analysis (CFA).

The Confirmatory Factor analysis resulted in a structure consisting of ten factors. These factors are related to the (i) convenience of cars, (ii) experience of driving, (iii) social status of car owners, (iv) own-car ideas, (v-ix) opinions on each of the five modes and (x) attitude to modal prestige. In general, car convenience is associated with ideas such as freedom, safety and how cars facilitate daily and personal activities, while driving experience is associated with the sense of control and adrenaline during driving. The social status factors are related to how people feel about having a car and the image it conveys to others, and the own-car ideas factor bundles those statements that pertain specifically to the possession of a car. Finally, there are five factors with similar characteristics, in which each respondent evaluates a series of seven attributes for the five different modes included in this study, while the prestige attributes are collected in the final factor.

Based now on these ten factors, we aim at estimating models for different numbers of subpopulations in order to study how these different groups differ in their relationship with cars, their sociodemographic characteristics, and the urban environment where they live. Since these factors are continuous variables (and not ordinal as the initial set of 73 statements), we estimate these groups through Latent Profile Analysis (Spurk et al., 2020). In the following, we use the term "classes" to refer to the different groups identified in the analysis. Afterwards, we characterize

each class based on its sociodemographic composition, and we explore if there are differences regarding their residential location choice and the corresponding accessibility.

3. RESULTS AND DISCUSSION

Since the constructed factors are continuous variables (and not ordinal as the initial set of 73 statements), we estimate these groups through Latent Profile Analysis (Spurk et al., 2020). In the following, we use the term "classes" to refer to the different groups identified in the analysis. As a result, all of the analyses presented in this study are based on probabilistic calculations. We select a five classes model because the marginal decrease in BIC goes under 2%, the smallest class is big enough to study, and also the class membership is stable.

Overall, the five different classes vary in the way they relate to cars. We arrange them so that their attitudes towards cars become increasingly positive from left to right. We name these five classes “*car detractors*”, “*car hesitants*”, “*car positives*”, “*car friends*”, and “*car lovers*”, respectively. These names are based on the distribution of the ten different factors, which are described next. When we calculate the expected share of these five groups, we see a bell-shaped distribution centered around the third class, “the positives”, as presented in Figure 1.

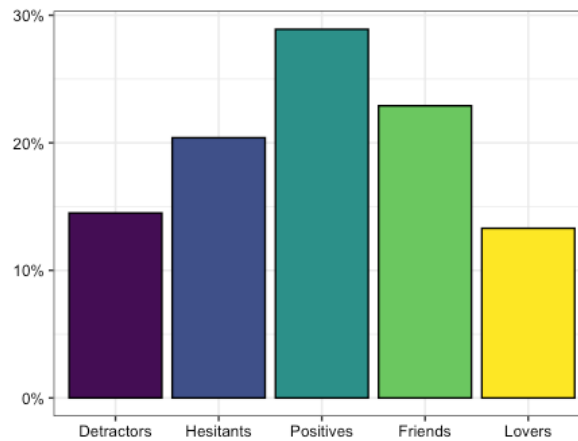


Figure 1. Classes' shares of sample respondents

In Figure 2, we present the 10th to 90th quantile range (grey line), the 25th and 75th quantile range (blue line), and the average (bullet) for each attribute and each class. The ten different factors obtained through the Confirmatory Factor analysis have the property that their respective average over the entire sample is fixed and equal to zero, denoted through a red dotted line in the figure. This property eases the comparison and allows us to identify differences among the latent groups. We also calculate and display the z-value for each attribute and class, assuming the null hypothesis of the mean being zero. We highlight in bold the z-values which imply significant differences from the sample average at the 95% confidence level.

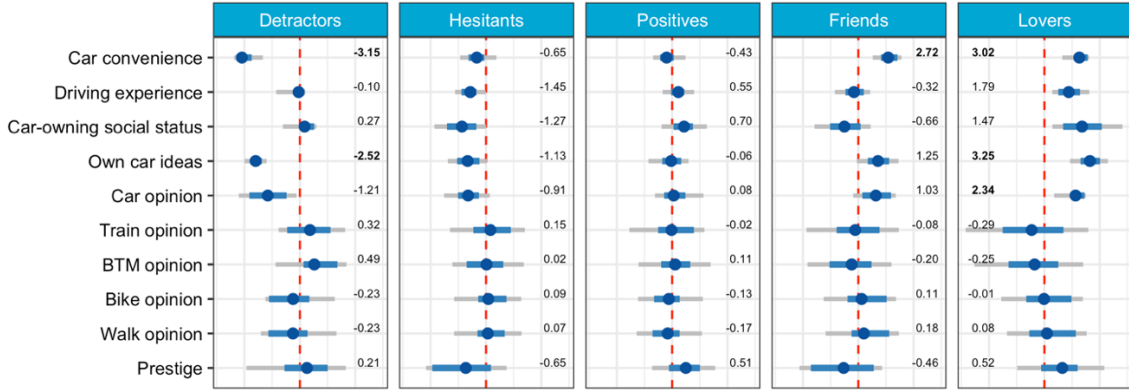


Figure 2. Factors distribution and share for each class

In general, we find that car convenience and attitudes towards both private cars and cars in general become more positive as we move from the *detractors* to *lovers*. *Detractors*, in comparison to other groups and the sample as a whole, report particularly low levels of convenience and positive opinions about cars. However, there are no significant differences between *detractors* and the overall sample with respect to driving experience and social status. Although not significantly different from the sample average, we observe that this group exhibits the highest appreciation for both trains and BTM among the different classes. *Hesitants* exhibit negative attitudes toward cars, but to a lesser degree than *detractors*. However, they have more negative attitudes towards the driving experience, social status, and the relevance of prestige factors. Together, these two groups make up 35% of the sample.

In addition, the distribution of these five classes is centred around what we denoted as *car positives*, as these individuals hold slightly positive attitudes towards driving experience, car-owning social status, and also towards modal prestige. This fact suggests that this class of users have a positive impression of what owning and using a car means. Interestingly, we found no significant differences between this class and the sample for any of the attributes.

The *car friends* group exhibits significantly more positive attitudes towards car convenience and car opinions compared to the other three groups already described and the sample as a whole. However, they do not share this positive attitude towards driving and social status, suggesting a potential willingness to consider alternative modes if they are similarly convenient. Conversely, *car lovers* display very positive attitudes towards cars across all variables and report more negative impressions of other modes, particularly trains and buses. These two groups together make up for just over 36% of the sample, and when including the more neutral but still pro-car *car positives* class, they account for 65% of the sample.

Afterwards, we are interested in the analysis of the socio-demographic composition of each identified latent class. This way, we can identify particular differences which might help us understand better who is represented by each of these attitudinal groups. To aid our analysis, we use Figure 3 to present the distribution of 13 attributes for each class, with a color code indicating the percentage-wise difference from the sample average. This visualization facilitates the identification of those attributes' levels that are either under- or over-represented in each class. Looking at

Figures 3a and 3b, which show *detractors* and *hesitants*, respectively, we see that there are no gender differences for detractors, but women are overrepresented amongst hesitants. The share of people older than 60 years old in both classes and younger people in the case of *detractors* is higher than in the sample as a whole. Both classes have a higher proportion of people who own a public transport card and fewer who never use bike as a means of transportation. Most *car detractors* do not own a car, whereas households who have only one car are overrepresented in the *hesitants* class. In terms of car access, *detractors* tend not to have a driving license, while households who have a car available but not freely accessible are overrepresented amongst *hesitants*. Both *detractors* and *hesitants* are characterized by a higher proportion of smaller households: with one and two persons per household, respectively. Regarding the central class, *car positives*, there is no evident difference between their socio-demographical distribution and the sample's average (Figure 3c).

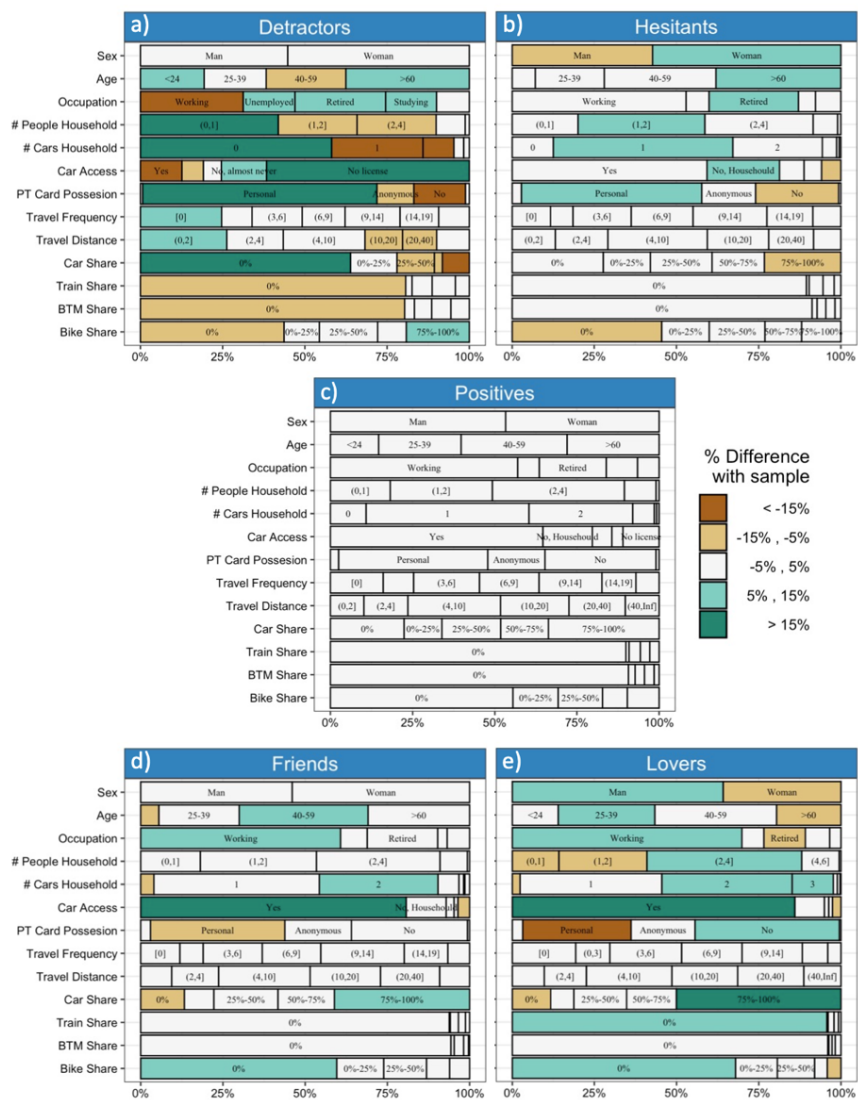


Figure 3. Socio-demographic characterisation of a) *car detractors*, b) *car hesitants*, c) *car positives*, d) *car friends*, and e) *car lovers*.

Figures 3d and 3e present the socio-demographic characteristics of *car friends* and *lovers*, respectively. *Car lovers* are predominantly male, while there is no gender difference for *car friends*. Both groups have a higher proportion of working-age people and households with at least two cars, and they also have greater car access than the overall sample. They are also less likely to own a personal public transport card and have a higher percentage of individuals who never bike. *Car lovers* also have a higher share of people who never use the train. There are no significant differences between these classes and the overall sample in terms of their travel distance and travel frequency.

The previous analysis does not take into account where each respondent lives. Therefore, we will examine the geographic distribution of the five latent classes across the Netherlands and how various urban environment variables are distributed within each class. This analysis is important as we expect that the factors that influence car usage affinity will vary depending on the built environment conditions. We calculate the distance between all households and the nearest urban area, train station, metro or express tram stop, tram stop, and bus stop (based on different frequency thresholds) and then average those based on class membership rates. The results of this analysis are presented in Table 1.

Table 1: Average and 10th to 90th quantile range of the distance to different locations for each the classes

Distance to	Detractors	Hesitants	Positives	Friends	Lovers
Urban area [km]	15.06 (1.7 - 35.1)	15.35 (2.2-35.0)	17.07 (2.5-37.1)	16.39 (2.8-35.4)	18.56 (3.3-38.3)
Train station [km]	3.07 (0.5-7.8)	3.55 (0.6-8.9)	3.87 (0.7-9.5)	3.79 (0.7-8.8)	4.09 (0.7-9.6)
Metro, express tram stop [km]	50.10 (0.9-132.4)	48.36 (1.6-122.8)	53.62 (2.0-128.0)	51.25 (2.0-125.4)	58.83 (2.3-133.4)
Tram stop [km]	54.37 (0.4-134.7)	53.857 (2.1-128.4)	59.09 (2.6-130.6)	57.36 (3.3-129.8)	64.50 (4.6-138.9)
Bus stop >4 /hr [km]	1.07 (0.1-2.7)	1.26 (0.1-3.3)	1.48 (0.2-4.2)	1.48 (0.2-4.4)	1.71 (0.2-5.4)
Bus stop >2 /hr [m]	460.04 (99.0-836.8)	549.48 (110.0-1057.6)	607.21 (113.0-1313.3)	612.87 (114.0-1345.7)	659.60 (113.0-1378.4)
Bus stop >1 /hr [m]	297.19 (94-515.94)	348.20 (103-587)	345.33 (104-603.4)	359.95 (103-610)	354.16 (99-656)
Bus stop [m]	275.21 (92.7-474.5)	307.93 (98.0-520.9)	304.64 (101.0-542.0)	313.01 (99.0-545.0)	307.51 (95.0-579.5)

The proximity of households to urban areas and public transportation facilities varies depending on the degree of affinity with cars: *detractors* and *hesitants* tend to be closest to urban areas, followed by *positives* and *friends*, while *lovers* tend to reside farther away. In terms of distance to the closest train station, *car lovers* are, on average, one kilometre (+33%) farther away than *detractors*. The situation is different for metro and tram stops, as *car lovers* are significantly

farther away from these facilities, while the other classes are comparatively closer to each other. Regarding bus stops, the frequency threshold is an important factor. Although there are no major differences between the classes for any bus stop, *car detractors* are significantly closer to high-frequency bus stops. This distance increases, as expected, for the other classes, particularly for *car lovers*.

Finally, we are also interested in the distribution of household location urbanization level and also the respondents' perceptions about parking and accessibility in their neighbourhoods. These results are presented in Table 2. The urbanization level varies from non-urbanized to very highly urbanized, while the scale used by respondents to indicate their opinions varies from strongly disagree to strongly agree.

Table 2: Responses distribution for different urban variables for each of the classes

Urban variable		Detractors	Hesitants	Positives	Friends	Lovers
<i>Urbanization Level</i>	Non urbanized	6.5%	7.3%	8.5%	9.2%	11.5%
	Low	15.9%	20.7%	23.2%	22.1%	24.4%
	Moderate	15.2%	17.3%	19.2%	18.3%	17.6%
	High	33.7%	31.0%	30.1%	32.0%	31.7%
	Very High	28.8%	23.7%	19.0%	18.4%	14.9%
<i>My neighbourhood has a sufficient number of parking places</i>	Strongly Disagree	5.7%	4.5%	3.8%	5.5%	5.5%
	Disagree	15.0%	15.4%	14.1%	14.9%	16.1%
	Neutral	18.1%	13.2%	13.4%	13.1%	11.7%
	Agree	32.6%	36.8%	39.2%	29.0%	23.6%
	Strongly Agree	24.0%	30.0%	29.0%	37.4%	42.6%
<i>My neighbourhood is easily accessible by car</i>	Unknown	4.6%	0.1%	0.5%	0.1%	0.5%
	Strongly Disagree	1.6%	1.1%	0.9%	1.1%	1.1%
	Disagree	4.1%	2.0%	1.9%	1.4%	2.1%
	Neutral	12.9%	4.0%	3.4%	2.3%	2.1%
	Agree	39.7%	38.9%	41.0%	20.8%	19.7%
<i>My neighbourhood is easily accessible by bicycle</i>	Strongly Agree	37.8%	54.0%	52.5%	74.4%	74.7%
	Unknown	4.0%	0.0%	0.2%	0.0%	0.4%
	Strongly Disagree	1.3%	0.9%	0.6%	0.5%	0.7%
	Disagree	2.0%	1.0%	1.6%	0.8%	1.4%
	Neutral	8.4%	0.9%	1.4%	0.8%	1.6%
<i>My neighbourhood is easily accessible by Public Transport</i>	Agree	34.2%	31.5%	37.9%	16.9%	17.7%
	Strongly Agree	52.9%	65.5%	58.1%	80.7%	78.2%
	Unknown	1.2%	0.1%	0.4%	0.3%	0.4%
	Strongly Disagree	5.6%	7.2%	5.3%	8.9%	7.9%
	Disagree	9.3%	13.1%	14.6%	14.2%	10.8%
<i>My neighbourhood is easily accessible by Public Transport</i>	Neutral	16.8%	15.0%	16.7%	15.3%	17.4%
	Agree	37.8%	37.4%	38.4%	27.9%	27.1%
	Strongly Agree	29.1%	26.7%	23.4%	32.2%	35.0%
	Unknown	1.5%	0.6%	1.6%	1.5%	1.8%

Based on these results, *car detractors* and *hesitants* are more likely to live in highly- and very highly-urbanized areas, while *friends* and *lovers* are more likely to live in non- or low-

urbanization areas. In the case of parking facilities, respondents' answers are generally similar, except for *friends* and *lovers* who strongly agree more frequently with the statement that there are enough parking spaces in their neighbourhoods. When asked about accessibility by car, a considerable gap can be observed between *car lovers* and other classes. About 75% of *car lovers* or *friends* strongly agree that their neighbourhood is easily accessible by car, while this figure drops to under 40% for *car detractors*. A similar trend is observed for bike accessibility, where approximately 80% of *car friends* and *lovers* strongly agree that their neighbourhood has good bike accessibility, whereas only 53% of *detractors* do.

4. CONCLUSIONS

Based on the analysis of the distribution of ten latent attitudinal factors, we identify five different sub-population groups which vary in terms of car ownership and usage ideas. In addition, we can also observe differences in their sociodemographic characteristics. As expected, more positive car attitudes are associated with higher car ownership and access and reduced use of public transportation modes. Noticeably, there are no significant differences in terms of travel frequency and trip length distribution, which suggests the differences come mostly from modal preference and not from the associated activities.

The place where people live and their personal circumstances are also relevant variables when studying car ownership and usage. In summary, we observe that the farther away households tend to be located from urban areas and public transportation facilities, the closer the relationship with cars, and vice-versa. Thus, policies that either aim to restrict or reduce car use need to acknowledge that not every car user behaves the same: reactions will vary depending on their attitudinal characteristics.

We aim to continue this research by now analysing how these different groups of the population may differ in terms of relative access to opportunities. For example, we intend to identify groups of people who hold negative attitudes towards cars but don't have sufficient access to alternative modes of transportation. These individuals may feel like they have no choice but to rely on cars, even though they would prefer not to. By conducting this analysis, we aim to gain valuable insights into how to identify and address the barriers that prevent people from choosing more sustainable transportation options and ultimately lead to a more environmentally friendly and equitable society.

ACKNOWLEDGEMENTS

This research was supported by the Amsterdam Institute for Advanced Metropolitan Solutions (AMS Institute).

REFERENCES

- Abenoza, R. F., Cats, O., & Susilo, Y. O. (2017). Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. *Transportation Research Part A: Policy and Practice*, 95, 64–84. <https://doi.org/10.1016/j.tra.2016.11.011>
- de Oña, J. (2022). Service quality, satisfaction and behavioral intentions towards public transport from the point of view of private vehicle users. *Transportation*, 49(1), 237–269. <https://doi.org/10.1007/s11116-021-10175-7>
- Goodwin, P., & Van Dender, K. (2013). ‘Peak Car’—Themes and Issues. *Transport Reviews*, 33(3), 243–254. <https://doi.org/10.1080/01441647.2013.804133>
- Hoogendoorn-Lanser, S., Schaap, N. T. W., & OldeKalter, M.-J. (2015). The Netherlands Mobility Panel: An Innovative Design Approach for Web-based Longitudinal Travel Data Collection. *Transportation Research Procedia*, 11, 311–329. <https://doi.org/10.1016/j.trpro.2015.12.027>
- OECD. (2022). Transport, Performance indicators. <https://stats.oecd.org/index.aspx?queryid=73639>
- Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and “how to” guide of its application within vocational behavior research. *Journal of Vocational Behavior*, 120, 103445. <https://doi.org/10.1016/j.jvb.2020.103445>
- Van Exel, N. J. A., & Rietveld, P. (2010). Perceptions of public transport travel time and their effect on choice-sets among car drivers. *Journal of Transport and Land Use*, 2(3). <https://doi.org/10.5198/jtlu.v2i3.15>

On fair discounted charging in electric ride-hailing markets with limited budgets

Marko Maljkovic^{*1}, Gustav Nilsson¹, and Nikolas Geroliminis¹

¹Urban Transport Systems Laboratory (LUTS) École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. {marko.maljkovic,gustav.nilsson,nikolas.geroliminis}@epfl.ch

SHORT SUMMARY

Coordinated charging of electric vehicles (EVs) has the potential to provide significant benefits to both electric vehicle owners and the wider community. In fact, intelligent, coordinated charging of large electric fleets, such as the ones operated by ride-hailing companies, could be essential in preventing a collapse of the energy market. We study a scenario in which a central body, e.g., the power-providing company or the government, wants to influence how the EVs of different ride-hailing companies spread among different charging stations by offering discounted prices of charging. Compared to previous works in this domain, we investigate a Stackelberg-based mechanism that takes into account potentially limited discount budgets available to the companies. We propose an iterative method to compute the local Stackelberg equilibrium that guarantees fairness in the sense that we have equal prices of charging for all ride-hailing companies. Finally, we test the proposed method in a simulated case study based on taxi data from the city of Shenzhen.

Keywords: Stackelberg game, Electric vehicle charging, Ride-hailing operation

1 INTRODUCTION

The increased popularity of electric vehicles (EV) and the steep incline in their number International Energy Agency (2021) opened a significant amount of questions in the domain of electric energy management and electric mobility. On one hand, this widespread adoption of electric vehicles has led to a growing need for efficient and reliable charging infrastructure. Combining this with the increasing demand for electricity from EV charging puts the spotlight on the problem of efficient, coordinated charging. On the other hand, combining smart mobility systems and intelligent charging management could pave the way for gaining an opportunity to trade different services to achieve societal optimum, e.g., by providing discounted charging in off-peak hours, it could help improve the stability of the electricity grid by reducing the variability in demand. Moreover, given that ride-hailing companies now constitute the central part of the services offered within a city, it is to be expected that they also electrify their fleets. Hence, the impact of coordinated charging of a large number of vehicles operated by ride-hailing companies can be significant in preventing the collapse of the energy market. Ride-hailing companies already offer access to different service facilities to their drivers so it is not unlikely that they would also offer discounted charging in accordance with the received monetary subsidies aimed at incentivizing the drivers to follow the management's desires. That way, the companies, in collaboration with external financiers, could hope to improve the overall utilization rate of their fleets by increasing the availability of the vehicles or to motivate the drivers to charge in distant areas in an attempt to increase the coverage.

With this in mind, we study a scenario in which a central body, e.g., the power-providing company or the government, has a desire about how the vehicles should spread out among different charging stations in a region where the operators of several ride-hailing companies try to minimize their operational costs, depicted in Figure 1. We assume the charging infrastructure is shared, so the ride-hailing companies are inherently interested in directing their vehicles to different charging stations so as to minimize the queuing time at the stations. Moreover, we assume the central authority has the power to set the prices of charging at different stations and hence, tries to influence how the companies behave in an attempt to attain a personal objective. Since all the agents in the system compete for the resources, this opens the door for a game-theoretic analysis.

There is extensive literature on game theoretic models based on congestion, mean-field, Stackelberg and Inverse Stackelberg games used to solve different problems in the domain of smart mobility systems Basar & Srikant (2002); Brown & Marden (2020); Groot et al. (2012); Laha et al. (2019); Ma et al. (2013); Paccagnan et al. (2016, 2019); Staňková et al. (2011); Tushar et al. (2012); J. Zhang et al. (2018); L. Zhang et al. (2019). This paper, in particular, is a continuation of our previous works Maljkovic et al. (2022a,b), where we analyzed a game-theoretic model of a similar structure as the ones listed in the literature. We did so, however, from the perspective of designing demand-based, feedback pricing policies that guaranteed the exact minimization of the central authority’s objective, making our work fall in the category of Inverse Stackelberg games. With this work, however, we aim to address the problem of potentially unfair prices induced by the optimal charging policies proposed in Maljkovic et al. (2022a). Namely, we study a Stackelberg game setup that assumes a game with limited budgets is played between the central authority and the ride-hailing companies. Based on Maljkovic et al. (2023), we propose an iterative method to compute a fixed discount for the ride-hailing companies that aligns well with the individual discount budget constraints. That way, we provide complete fairness in the sense of having the same prices for every company at the expense of being able to guarantee convergence only to the local Stackelberg equilibrium of the game. At the end, we test the proposed method in a simulated case study based on real taxi data from the city of Shenzhen in China.

The paper is outlined as follows: the rest of this section is devoted to introducing some basic notation. In Section 2, we revise the structure of the model with limited budgets and introduce the iterative method for computing the local Stackelberg equilibrium. In the following section, Section 3, we demonstrate the effectiveness of the proposed method in a simulated case study based on the city of Shenzhen. The final section contains the concluding remarks and some ideas for future research.

Notation

Let \mathbb{R} denote the set of real numbers, \mathbb{R}_+ the set of non-negative reals, and \mathbb{Z}_+ the set of non-negative integers. Let $\mathbf{0}_m$ and $\mathbf{1}_m$ denote the all zero and all one vectors of length m respectively, and \mathbb{I}_m the identity matrix of size $m \times m$. For a finite set \mathcal{A} , we let $\mathbb{R}_{(+)}^{\mathcal{A}}$ denote the set of (non-negative) real vectors indexed by the elements of \mathcal{A} and $|\mathcal{A}|$ the cardinality of \mathcal{A} . Furthermore, for finite sets \mathcal{A}, \mathcal{B} and a set of $|\mathcal{B}|$ vectors $x^i \in \mathbb{R}_{(+)}^{\mathcal{A}}$, we define $x := \text{col}((x^i)_{i \in \mathcal{B}}) \in \mathbb{R}^{|\mathcal{A}||\mathcal{B}|}$ to be their concatenation. For $A \in \mathbb{R}^{n \times n}$, $A \succ 0 (\succeq 0)$ is equivalent to $x^T A x > 0 (\geq 0)$ for all $x \in \mathbb{R}^{n \times n}$. We let $A \otimes B$ denote the Kronecker product between two matrices and for a vector $x \in \mathbb{R}^n$, we let $\text{Diag}(x) \in \mathbb{R}^{n \times n}$ denote a diagonal matrix whose elements on the diagonal correspond to vector x . For a differentiable function $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, we let $\mathbf{D}_x f \in \mathbb{R}^{m \times n}$ denote the Jacobian matrix of f defined as $(\mathbf{D}_x f)_{ij} := \frac{\partial f_i}{\partial x_j}$. Finally, for a set-valued mapping $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $\text{gph}(\mathcal{F}) := \{(y, x) \in \mathbb{R}^n \times \mathbb{R}^m \mid x \in \mathcal{F}(y)\}$ denotes its graph.

2 METHODOLOGY

Model

We begin explaining the proposed methodology by introducing the system model. Let us consider a region where multiple shared charging stations are available for the EV drivers, i.e., let \mathcal{M} be the set of all charging stations such that $|\mathcal{M}| = m$ and $M_j > 0$ denotes the capacity of the charging station $j \in \mathcal{M}$. Let the set of all ride-hailing companies be denoted as \mathcal{I} . Let the cardinality of the set of companies be $|\mathcal{I}| = N$ and for every $i \in \mathcal{I}$, let the number $N_i > 0$ represent the number of vehicles that want to recharge. For every company $i \in \mathcal{I}$, let the vector $x^i \in \mathcal{X}_i \subseteq \mathbb{R}^m$ describe the ride-hailing fleet split among charging stations. Namely, let $\|x^i\|_1 = N_i$ and $x_j^i \geq 0$ denote the number of vehicles to be directed to station $j \in \mathcal{M}$. Moreover, if we define the sets $\mathcal{X} := \prod_{i \in \mathcal{I}} \mathcal{X}_i$ and $\mathcal{X}_{-i} := \prod_{j \in \mathcal{I} \setminus i} \mathcal{X}_j$, then the joint strategy of all followers can be denoted as $x := \text{col}((x^i)_{i \in \mathcal{I}}) \in \mathcal{X}$ and for every agent $i \in \mathcal{I}$ we can define $x^{-i} := \text{col}((x^j)_{j \in \mathcal{I} \setminus i}) \in \mathcal{X}_{-i}$.

Let the nominal prices of charging at different stations be encoded in vector $\pi_{\text{base}} \in \mathbb{R}^M$. Furthermore, let us assume that the central authority is interested in determining the optimal discount $\Delta \pi \in \mathbb{R}^M$, such that for every $i \in \mathcal{I}$, the total monetary discount that the company i receives does not exceed a predefined value $B_i \in \mathbb{R}$. Here, B_i represents the limited discount budget of company i , which corresponds to the level of external subsidies that the company is entitled to.

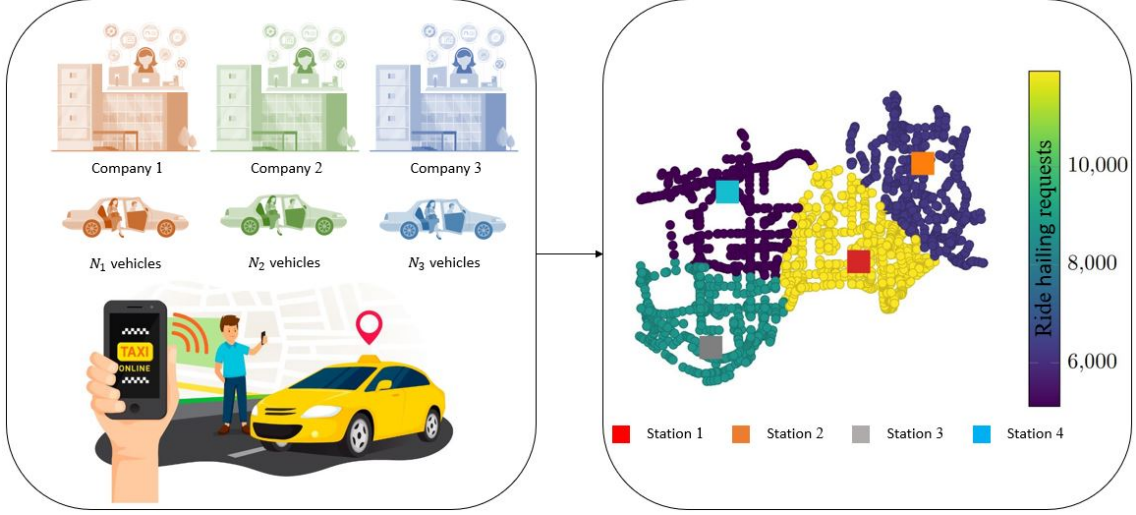


Figure 1: **Ride-hailing market in a region with 4 charging stations** - $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{M}_4\}$ and the network topology used in the case study consists of 1858 intersections connected by 2013 road segments, divided in 4 regions based on the Voronoi partitioning of the city with the centroids located at the charging stations.

Similar to the cost models introduced in Maljkovic et al. (2022a,b), we assume that based on the announced pricing policy $\pi \in \mathcal{P} \subseteq \mathbb{R}^M$, such that $\pi := \pi_{\text{base}} - \Delta\pi$, the ride-hailing companies choose their strategies in an attempt to minimize personal objective functions $J^i(x^i, x^{-i}, \pi)$ by playing the best response to the other agents' strategies, as illustrated in Figure 2. For every company $i \in \mathcal{I}$, let $\sigma(x^{-i}) := \sum_{j \in \mathcal{I} \setminus i} x^j$ be the aggregate decisions of the other players and let $\sigma(x) := \sum_{j \in \mathcal{I}} x^j$. Then, every company operator is interested in minimizing its operational cost under the feasibility constraints imposed by the battery status of its vehicles. Inspired by the objective functions analyzed in Maljkovic et al. (2022a,b); Tushar et al. (2012); Yu et al. (2021); Zavvos et al. (2022), here, we analyze the operator's cost that consists of three terms

$$J^i(x^i, \sigma(x^{-i}), \pi) = J_1^i(x^i, \sigma(x^{-i})) + J_2^i(x^i) + J_3^i(x^i, \pi),$$

such that $J_1^i(x^i, \sigma(x^{-i}))$ denotes the expected queuing cost, $J_2^i(x^i)$ denotes the negative expected revenue and $J_3^i(x^i, \pi)$ denotes the charging cost.

Expected queuing cost model is governed by the cost term of the form:

$$\begin{aligned} J_1^i(x^i, \sigma(x^{-i})) &= (x^i)^T Q (x^i + \sigma(x^{-i}) - M) \\ &= (x^i)^T Q (\sigma(x) - M), \end{aligned} \quad (1)$$

where $M \in \mathbb{R}^M$ is the vector of charging station capacities, i.e., $M = \text{col}((M_j)_{j \in \mathcal{M}})$, and $Q = \text{Diag}(q) \in \mathbb{R}^{M \times M}$ is a positive definite scaling matrix such that every element $q_j > 0$ depicts how expensive it is for a vehicle to queue in the region around charging station $j \in \mathcal{M}$. The charging stations located in the city's more busy areas should experience higher queuing costs and hence have a higher corresponding diagonal entry in the Q matrix. Moreover, the more the capacity of the station is exceeded, the higher the cost per vehicle should be, which is directly enabled through the inner product with the vector $\sigma(x) - M$. To take into account the total queuing cost for the whole fleet, we calculate the inner product between the vector describing the fleet's distribution, i.e., x^i , and the incurred cost per vehicle for choosing a particular station, i.e., $Q(\sigma(x) - M)$.

The negative expected revenue is modeled as:

$$J_2^i(x^i) = (e_i^{\text{arr}})^T x^i - (e_i^{\text{pro}})^T x^i, \quad (2)$$

where $e_i^{\text{arr}} \in \mathbb{R}^M$ is the average cost of a vehicle being unoccupied while traveling to a station and the vector $e_i^{\text{pro}} \in \mathbb{R}^M$ is the expected profit per vehicle, should the vehicle choose to stay in a region around a particular charging station, estimated from historical data.

The charging cost is modeled as:

$$J_3^i(x^i, \pi) = \pi^T S_i x^i, \quad (3)$$

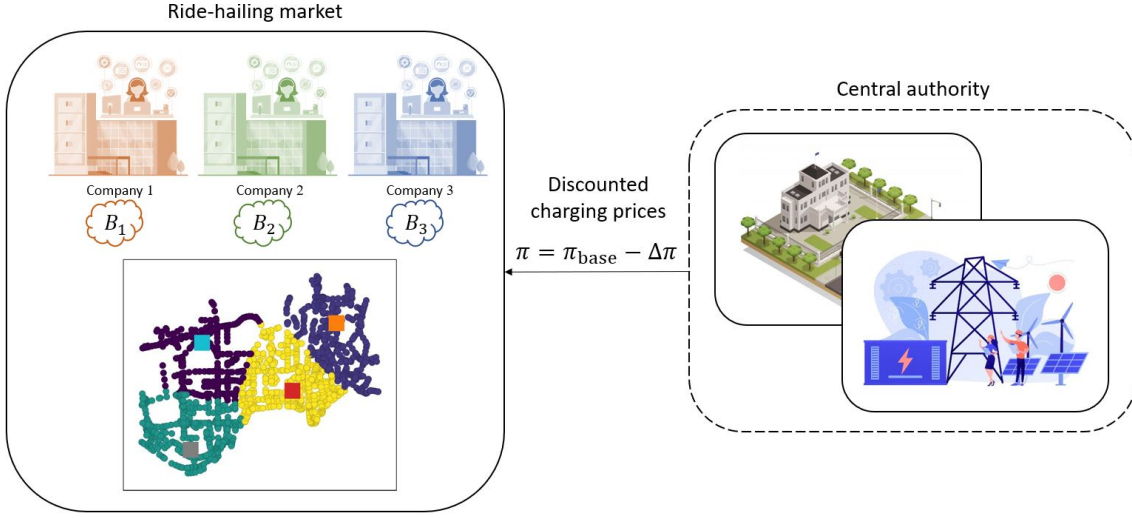


Figure 2: Schematic overview of the interaction between the ride-hailing market and the central authority, i.e., the power-providing company or the government.

where, the matrix $S_i \in \mathbb{R}^{M \times M}$ is diagonal, i.e., $S_i = \text{Diag}(d^i) \succeq 0$, and every element $d_k^i \in \mathbb{R}_+$ of the vector $d^i \in \mathbb{R}_+^M$ can be interpreted as the expected average charging demand per vehicle when choosing the station k . Therefore, the total operational cost of the company can be written in a general form given by:

$$J^i(x^i, \sigma(x^{-i}), \pi) = \frac{1}{2} (x^i)^T P_i x^i + (x^i)^T Q_i \sigma(x^{-i}) + r_i^T x^i + \pi^T S_i x^i, \quad (4)$$

where $P_i := 2Q$, $Q_i := Q$ and $r_i := e_i^{\text{arr}} - e_i^{\text{pro}}$. Regarding the constraint sets of the ride-hailing companies, it has been shown that for a particularly constructed polytopic constraint, it will always be possible to match every ride-hailing vehicle with exactly one charging station in an attempt to respect the optimal allocation given by the optimal split x^* . For every $i \in \mathcal{I}$, the matching constraints in accordance with Maljkovic et al. (2022a) are given by

$$\mathcal{X}_i^m := \{x^i \in \mathbb{R}^M \mid A_i x^i = b_i \wedge G_i x^i \leq h_i\}. \quad (5)$$

Apart from them, in this paper we also focus on budget constraints. As previously mentioned, for every company $i \in \mathcal{I}$, the total discount budget for the electricity prices that the central authority can provide is B_i . Taking into account that the discount is given by $\Delta\pi = \pi_{\text{base}} - \pi$, for a particular pricing strategy $\pi \in \mathcal{P}$, the budget constraint can be described by $\mathcal{X}_i^b(\pi)$

$$\mathcal{X}_i^b(\pi) := \left\{x^i \in \mathbb{R}^M \mid (\pi_{\text{base}} - \pi)^T S_i x^i - B_i \leq 0\right\}, \quad (6)$$

which is also a polytopic constraint in x^i . Hence, for a pricing strategy $\pi \in \mathcal{P}$ and for every $i \in \mathcal{I}$, the resulting constraint set is given by $\mathcal{X}_i = \mathcal{X}_i^m \cap \mathcal{X}_i^b(\pi)$. It is worth mentioning that for particular choices of parameters S_i , π_{base} and B_i the set $\mathcal{X}_i^b(\pi)$ could be empty and, hence, the optimization problem would be infeasible. Therefore, for future analysis, we assume that $\mathcal{X}_i^b(\pi) \neq \emptyset$.

On the other hand, we assume the central authority is interested in balancing the vehicles so as to minimize the personal objective of the form:

$$\min_{\sigma(x)} J_G(\sigma(x)) = \min_{\sigma(x)} \frac{1}{2} \sigma(x)^T A_G \sigma(x) + b_G^T \sigma(x), \quad (7)$$

for some diagonal matrix $A_G \succ 0$ and $b_G \in \mathbb{R}^M$. In particular, in this paper, we focus on minimizing a special case of (7) that corresponds to balancing the vehicles so as to match a predefined vehicle distribution given by vector $\mathcal{Z} \in [0, 1]^M$ with $\mathbf{1}^T \mathcal{Z} = 1$, i.e., to minimize

$$J_G(\sigma(x)) = \frac{1}{2} \|\sigma(x) - \mathbf{1}^T n \mathcal{Z}\|_2^2, \quad (8)$$

where $n = \text{col}((N_i)_{i \in \mathcal{I}})$ is the vector containing the number of vehicles per company that need to be recharged. Having defined the system model, we continue to present the theoretical preliminaries.

Theoretical preliminaries

The central authority and the ride-hailing companies admit a single-leader, multiple-follower Stackelberg game with the leader being the central authority. Upon the announcement of the leader's strategy, the aggregative game between the ride-hailing companies is described by

$$G_0(\pi) := \left\{ \min_{x^i \in \mathcal{X}_i} J^i(x^i, \sigma(x^{-i}), \pi), \forall i \in \mathcal{I} \right\}, \quad (9)$$

whose Nash equilibrium x^* is given in the definition below:

Definition 1 (Nash equilibrium) For any pricing strategy $\pi \in \mathcal{P}$ of the central authority, a joint strategy $x^* \in \mathcal{X}$ is a Nash equilibrium of the game G_0 , if for all $i \in \mathcal{I}$ and all $x^i \in \mathcal{X}_i$ holds

$$J^i(x^{i*}, x^{-i*}, \pi) \leq J^i(x^i, x^{-i*}, \pi).$$

We focus our attention on the subset of general Nash equilibria given by Definition 1, called the Variational Nash equilibria (v-NE), because different methods in the literature facilitate their decentralized computation Grammatico et al. (2016). Based on the theory of variational inequalities Harker & Pang (1990), if we define a map $F : \mathcal{X} \times \mathcal{P} \rightarrow \mathbb{R}^{NM}$ as

$$F(x, \pi) := \text{col} \left((\nabla_{x^i} J^i(x^i, x^{-i}, \pi))_{i \in \mathcal{I}} \right),$$

then the set of v-NE of the game $G_0(\pi)$ is given by $\mathcal{V}_0(\pi)$:

$$\mathcal{V}_0(\pi) := \left\{ x \in \mathcal{X} \mid (y - x)^T F(x, \pi) \geq 0, \forall y \in \mathcal{X} \right\}.$$

With this in mind, we now proceed to state the existence and uniqueness result for $G_0(\pi)$.

Proposition 1 For any $\pi \in \mathcal{P}$, let the game $G_0(\pi)$ between the ride-hailing companies be defined as in (9). Moreover, for every company $i \in \mathcal{I}$, let the constraint sets \mathcal{X}_i be defined as $\mathcal{X}_i = \mathcal{X}_i^m \cap \mathcal{X}_i^b(\pi)$, with \mathcal{X}_i^m defined in (5) and $\mathcal{X}_i^b(\pi)$ defined in (6). If the company operator's objective is defined by (1), (2), (3) and (4), then the game $G_0(\pi)$ admits a unique v-NE joint strategy $x^* \in \mathcal{X}$.

Proof Since $P_i \succ 0$ for every $i \in \mathcal{I}$, the agents' cost functions are convex in x^i . For \mathcal{X}_i defined as $\mathcal{X}_i = \mathcal{X}_i^m \cap \mathcal{X}_i^b(\pi)$, with \mathcal{X}_i^m as in (5) and $\mathcal{X}_i^b(\pi)$ as in (6), based on (Rosen, 1965, T.1), there exists a Nash equilibrium of the game $G_0(\pi)$. A sufficient condition for the uniqueness of the Nash equilibrium is that the operator $F(x, \pi)$ be strictly monotone in x (Facchinei & Pang, 2007, Ch.2). The pseudo gradient can be written as $F(x, \pi) = F_1 x + F_2$, such that $F_1 = \mathbb{I}_N \otimes Q + \mathbf{1}_N \mathbf{1}_N^T \otimes Q$ and $F_2 = \text{col}((r_i + S_i \pi)_{i \in \mathcal{I}})$. To show that $F(x, \pi)$ is strictly monotone, it suffices to prove that $F_1 \succ 0$ Bauschke & Combettes (2017). This is true as for any $x \in \mathcal{X}$, it holds that $x^T F_1 x = \sum_{i \in \mathcal{I}} (x^i)^T Q x^i + (\sum_{i \in \mathcal{I}} x^i)^T Q (\sum_{i \in \mathcal{I}} x^i) > 0$.

Since the unique v-NE can be computed using the Picard-Banach fixed point iteration Berinde (2004), we can now formally define the Stackelberg game played between the central authority and the ride-hailing market. Moreover, we can introduce the notion of the local Stackelberg equilibrium (l-SE) that we wish to compute for this hierarchical game structure.

Definition 2 (Stackelberg game) Let the game between the ride-hailing companies be defined as in Proposition 1 and the central authority's objective be defined as (8). Then the Stackelberg game is defined by a bi-level optimization problem

$$G_L := \left\{ \begin{array}{l} \min_{\pi \in \mathcal{P}} J^L(x^*, \pi) = \frac{1}{2} \|\sigma(x^*) - \mathbf{1}^T n \mathcal{Z}\|_2^2 \\ \text{s.t. } x^* \in \mathcal{V}_0(\pi) \end{array} \right\}. \quad (10)$$

In general, there could exist multiple Stackelberg equilibria that solve the game given by (10). Therefore, we shift our focus towards computing the local Stackelberg equilibria given by the following definition and previously analyzed in Fabiani et al. (2022).

Definition 3 (Local Stackelberg equilibrium) Let G_L be a game as in Definition 2. A pair $(\hat{x}^*, \hat{\pi}) \in \text{gph}(\mathcal{V}_0) \cap (\mathcal{X} \times \mathcal{P})$ is a local Stackelberg equilibrium of G_L if there exist open neighborhoods $\Omega_{\hat{x}^*}, \Omega_{\hat{\pi}}$ of \hat{x}^* and $\hat{\pi}$ respectively, such that

$$J^L(\hat{x}^*, \hat{\pi}) \leq \inf_{(x^*, \pi) \in \text{gph}(\mathcal{V}_0) \cap \Omega} J^L(x^*, \pi),$$

where $\Omega := \Omega_{\hat{x}^*} \times (\mathcal{P} \cap \Omega_{\hat{\pi}})$.

In the following section, we will present a bi-level, iterative method for computing the local Stackelberg equilibria.

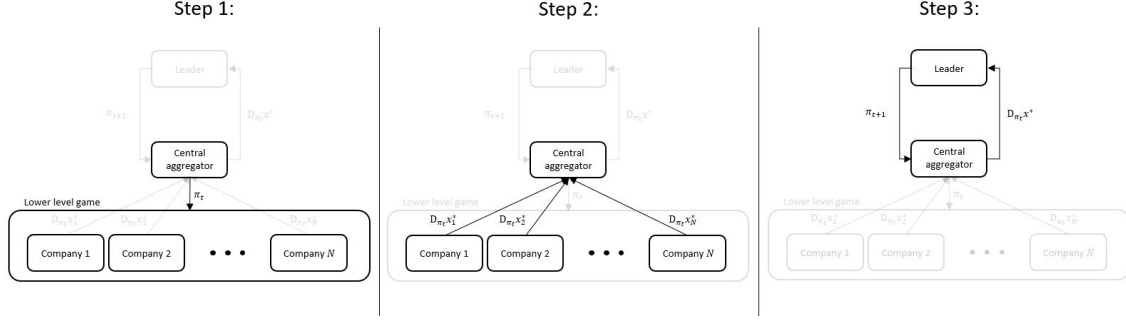


Figure 3: Overview of the three-step iterative procedure used for calculating the local Stackelberg equilibrium.

Computing the l-SE

To compute the local Stackelberg equilibria, we build on top of the standard iterative structure used for finding the v-NE of the aggregative game played between the ride-hailing companies. As previously mentioned, based on Proposition 1, we can utilize the Picard-Banach iteration to find the v-NE of $G_0(\pi)$ for a particular $\pi \in \mathcal{P}$. Based on the aggregative structure of the operator's cost, standard methods leverage the entity typically referred to as the 'central aggregator' to transmit the information about $\sigma(x)$ to all the agents in the game during the procedure.

To compute the l-SE, we will perform a gradient-based, iterative procedure that requires communication just between the 'central aggregator' and the central authority. Namely, we aim to update the central authority's pricing strategy according to

$$\pi_{t+1}(\pi_t, s) := \Pi_{\mathcal{P}} \left[\pi_t - s \frac{dJ^L(x^*(\pi), \pi)}{d\pi} \Big|_{\pi=\pi_t} \right], \quad (11)$$

where Π denotes the projection operator, π_t is the current value of the central authority's pricing policy and s is the step size carefully determined according to the Armijo step-size rule Bertsekas (1999). If we restrict ourselves to compact and convex pricing spaces $\mathcal{P} \subseteq \mathbb{R}^M$, then the complexity of each update step defined by (11) boils down to estimating how the Nash equilibrium of $G_0(\pi)$ reacts to any change in π , i.e., calculating the gradient

$$\frac{dJ^L(\cdot)}{d\pi} = \frac{\partial J^L(\cdot)}{\partial \pi} + \sum_{i \in \mathcal{I}} \mathbf{D}_{\pi}^T x^{i*} \frac{\partial J^L(\cdot)}{\partial x^{i*}}. \quad (12)$$

Here, it is of paramount importance to show that the Jacobians $\mathbf{D}_{\pi}^T x^{i*}$ are well defined. To do so, we take into account that the unique v-NE, $x^* \in \mathcal{X}$, pre-computed for the current value of π_t , has to satisfy the KKT optimality conditions of the best-response optimization problem given by (9) for each ride-hailing company. For every $i \in \mathcal{I}$, let us describe the constraint set $\mathcal{X}_i := \mathcal{X}_i(\pi)$ as

$$\mathcal{X}_i(\pi) = \left\{ x^i \in \mathbb{R}^M \mid \begin{bmatrix} G_i \\ (\pi_{\text{base}} - \pi) S_i \end{bmatrix} x^i \leq \begin{bmatrix} h_i \\ B_i \end{bmatrix} \right\} = \{x^i \in \mathbb{R}^M \mid \Gamma_i x^i \leq \delta_i\}. \quad (13)$$

Moreover, let us partition Γ_i and δ_i into active and inactive inequality constraints described by $\bar{\Gamma}_i$, $\underline{\Gamma}_i$, $\bar{\delta}_i$ and $\underline{\delta}_i$ such that

$$\bar{\Gamma}_i x^{i*} = \bar{\delta}_i \text{ and } \underline{\Gamma}_i x^{i*} < \underline{\delta}_i. \quad (14)$$

Then, applying the Implicit Function theorem Dontchev & Rockafellar (2009) to the KKT mapping of an equivalent best-response optimization problem with partitioned constraints described by (14), directly yields

$$\mathbf{D}_{\pi} x^{i*} = - \begin{bmatrix} \frac{\partial^2}{\partial x_i \partial x_i} J^i & \underline{\Gamma}_i^T & \bar{A}_i^T \\ \mathbf{0} & \text{Diag}(\underline{\Gamma}_i x^{i*} - \underline{\delta}_i) & \mathbf{0} \\ \bar{A}_i & \mathbf{0} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} S_i \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad (15)$$

where \bar{A}_i is full row-rank and obtained by removing redundant constraints from $A_{\text{total}}^T = [A_i^T, \bar{\Gamma}_i^T]$. The schematic representation of the three-step procedure is presented in Figure 3. Finally, we can summarize the convergence results in the following proposition.

Proposition 2 *Let the Stackelberg game between the central authority and the ride-hailing companies be defined as (10). Moreover, let the update step of the central authority’s pricing strategy be defined by equations (11), (12), (13) and (15) and the step size $s > 0$ in (11) be chosen according to Armijo step-size rule. Then, for \mathcal{P} compact and convex, the following convergence result holds:*

$$\lim_{t \rightarrow +\infty} [J^L(\cdot, \pi_{t+1}) - J^L(\cdot, \pi_t)] = 0,$$

Proof The proof follows directly from applying the Implicit function theorem Dontchev & Rockafellar (2009) and applying the properties of the Armijo step-size rule Bertsekas (1999).

In the following section we illustrate the performance of our algorithm in a simulated case study based on taxi data from the city of Shenzhen.

3 RESULTS AND DISCUSSION

Case study

We begin this section by introducing the case study that was previously analyzed in Maljkovic et al. (2022a). We consider 3 ride-hailing companies $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3\}$ with fleet sizes given by $N_{\text{fleet}} = [450, 400, 350]^T$ that operate in the Shenzhen region with 4 public charging stations $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{M}_4\}$. The stations are described by the vector of their capacities $M = [15, 60, 35, 50]^T$ and are located in parts of Shenzhen with different demands for ride-hailing services as shown in the color-coded map depicted in Figure 1. We consider a 3 hour long simulation that represents one of the two peak-hour periods during the day. New passengers constantly arrive in the system and either increase the number of private vehicles in the system or request a ride-hailing vehicle to be assigned to them. The demand profile represents the real taxi demand that we assume is now served by the ride-hailing companies Beojone & Geroliminis (2021). The congested conditions in the city are represented by modeling the space mean speed of the vehicles as a decreasing function of the total vehicle accumulation n_v in the region and according to the network Macroscopic Fundamental Diagram (MFD) Geroliminis & Daganzo (2008) obtained from Ji et al. (2014). Under the assumption of homogeneous congestion in the city, the MFD of the region is given by:

$$v_{\text{space}}(n_v) = \begin{cases} 36 \exp\left(-\frac{29n_v}{60000}\right), & \text{if } \frac{n_v}{1000} \leq 36 \\ 6.31 - 0.28\left(\frac{n_v}{1000} - 36\right), & \text{if } 36 < \frac{n_v}{1000} \leq 60 \\ 0, & \text{if } \frac{n_v}{1000} > 60 \end{cases}.$$

To prevent the ride-hailing vehicles from flocking in the busiest parts of the city, the desired distribution of the ride-hailing vehicles \mathcal{Z} is formed so as to match the spatial distribution of the ride-hailing service requests. To approximate this distribution, the city region is divided into 4 cells according to the Voronoi Kang (2008) partitioning of the map. The charging stations are chosen as the centroids of the Voronoi cells, the number of vehicles per company that want to recharge after a 3 hour simulation is given by $n = [194, 181, 157]$, and \mathcal{Z} is chosen to correspond to the total number of requests in each cell. For the analyzed case study, this results in obtaining \mathcal{Z} such that $\mathbf{1}^T n \mathcal{Z} = [198, 103, 144, 87]$ and we set $\mathcal{P} := [p_{\min}, p_{\max}]^4$, such that $p_{\min} = 0.0$ and $p_{\max} = 5.0$. All the remaining parameters in the simulation are kept identical as in Maljkovic et al. (2022a).

System performance

For the Picard-Banach fixed point iteration procedure used to compute the v-NE before each update step of the central authority’s pricing strategy, we used $k_{v\text{-NE}} = 5000$ iterations whereas for the iterative procedure between the central authority and the ‘central aggregator’ we used $k_{l\text{-SE}} = 350$ iterations. The evolution of the achieved total vehicle accumulations and the central authority’s objective are shown in Figure 4. For the given number of iterations, the system manages to achieve perfect matching with respect to desired vehicle distribution.

This is further supported by the plot on the right-hand side of Figure 4, which shows that the objective function converges to the global minimum value of 0. Finally, we can investigate the trend in the evolution of the actual discount budget used for each of the ride-hailing companies by looking at Figure 5. The evolution of the discount budget used is presented for the base price of $\pi_{\text{base}} =$

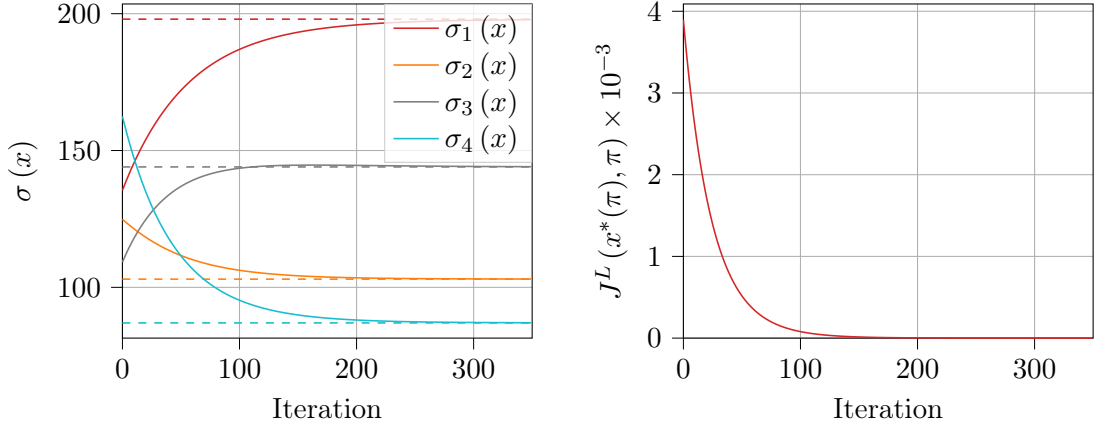


Figure 4: The left plot shows the evolution of the total vehicle accumulation $\sigma(x)$, whereas the right plot shows the evolution of the central authority's objective during the procedure.

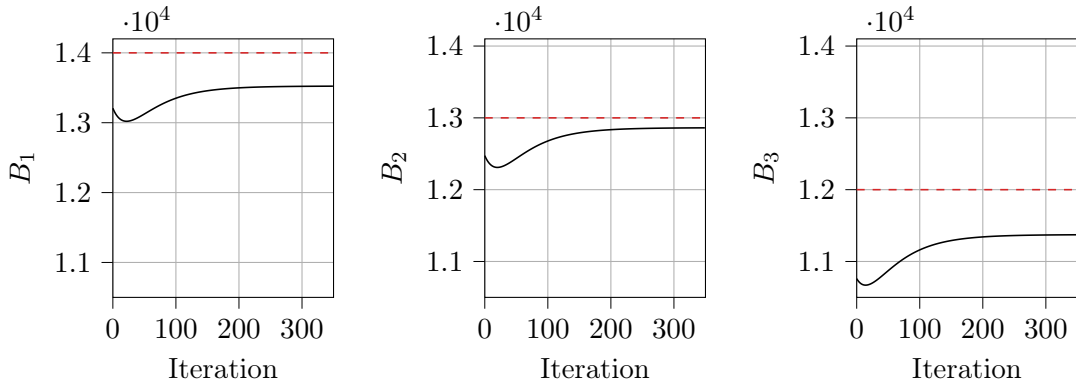


Figure 5: The three plots show the evolution of the value of the used discount budget. The black lines represent the achieved values whereas the red line represents the maximum possible value of the discount budget.

[5.0, 3.0, 5.0, 3.0] and the attained charging discount is given by $\Delta\pi = [1.560, 0.809, 2.179, 1.440]$. Note that in this case, the system was able to recover a solution that matches the global minimum of the central authority's objective. However, achieving such a local Stackelberg equilibrium is not always necessarily possible. In fact, for some cases, it could be that the chosen initial value of the pricing policy π_{init} largely determines which discount policy the algorithm converges to. Therefore, we plan to investigate in the future how different initial values of the pricing policy influence the result of the iterative procedure.

4 CONCLUSIONS

In this paper, we presented an iterative framework for computing a pricing strategy corresponding to a local Stackelberg equilibrium in a pricing game with one leader, i.e., the central authority and multiple followers, i.e., the ride-hailing companies, where the ride-hailing companies are constrained by fixed, a priori defined, discount budgets. We provided theoretical convergence guarantees and demonstrated the performance of the system in a simulated case study based on taxi data from the city of Shenzhen. However, what remains a promising research direction for the future is the question of how to choose the initial conditions in order to converge to a global optimum of the central authority's objective whenever such an optimum exists. Moreover, we aim to increase the complexity of the model in an attempt to better describe the reality.

ACKNOWLEDGEMENTS

This work was supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

REFERENCES

- Basar, T., & Srikant, R. (2002). Revenue-maximizing pricing and capacity expansion in a many-users regime. In *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies* (Vol. 1, p. 294-301 vol.1). doi: 10.1109/INFCOM.2002.1019271
- Bauschke, H. H., & Combettes, P. L. (2017). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer New York, NY. doi: 10.1007/978-3-319-48311-5
- Beojone, C. V., & Geroliminis, N. (2021). On the inefficiency of ride-sourcing services towards urban congestion. *Transportation Research Part C: Emerging Technologies*, 124, 102890. doi: 10.1016/j.trc.2020.102890
- Berinde, V. (2004). *Iterative approximation of fixed points*. Springer Berlin, Heidelberg. doi: 10.1007/978-3-540-72234-2
- Bertsekas, D. (1999). *Nonlinear programming*. Athena Scientific.
- Brown, P. N., & Marden, J. R. (2020). Can taxes improve congestion on all networks? *IEEE Transactions on Control of Network Systems*, 7(4), 1643-1653. doi: 10.1109/TCNS.2020.2992679
- Dontchev, A. L., & Rockafellar, R. T. (2009). *Implicit functions and solution mappings*. Springer New York, NY. doi: 10.1007/978-0-387-87821-8
- Fabiani, F., Tajeddini, M. A., Kebriaei, H., & Grammatico, S. (2022). Local Stackelberg equilibrium seeking in generalized aggregative games. *IEEE Transactions on Automatic Control*, 67(2), 965-970. doi: 10.1109/TAC.2021.3077874
- Facchinei, F., & Pang, J.-S. (2007). *Finite-dimensional variational inequalities and complementarity problems*. Springer New York, NY. doi: 10.1007/b97543
- Geroliminis, N., & Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9), 759-770. doi: 10.1016/j.trb.2008.02.002
- Grammatico, S., Parise, F., Colombino, M., & Lygeros, J. (2016). Decentralized convergence to Nash equilibria in constrained deterministic mean field control. *IEEE Transactions on Automatic Control*, 61(11), 3315-3329. doi: 10.1109/tac.2015.2513368
- Groot, N., De Schutter, B., & Hellendoorn, H. (2012). Reverse Stackelberg games, Part I: Basic framework. *2012 IEEE International Conference on Control Applications*, 421-426. doi: 10.1109/CCA.2012.6402334
- Harker, P., & Pang, J.-S. (1990). Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Math. Program.*, 48, 161-220. doi: 10.1007/BF01582255
- International Energy Agency. (2021). Global EV outlook 2021.
- Ji, Y., Luo, J., & Geroliminis, N. (2014). Empirical observations of congestion propagation and dynamic partitioning with probe data for large-scale systems. *Transportation Research Record*, 2422(1), 1-11. doi: 10.3141/2422-01
- Kang, J. M. (2008). Voronoi diagram. In S. Shekhar & H. Xiong (Eds.), *Encyclopedia of GIS* (pp. 1232-1235). Boston, MA: Springer US. doi: 10.1007/978-0-387-35973-1_1461
- Laha, A., Yin, B., Cheng, Y., Cai, L. X., & Wang, Y. (2019). Game theory based charging solution for networked electric vehicles: A location-aware approach. *IEEE Transactions on Vehicular Technology*, 68(7), 6352-6364. doi: 10.1109/TVT.2019.2916475

- Ma, Z., Callaway, D. S., & Hiskens, I. A. (2013). Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Transactions on Control Systems Technology*, 21(1), 67-78. doi: 10.1109/TCST.2011.2174059
- Maljkovic, M., Nilsson, G., & Geroliminis, N. (2022a). Hierarchical pricing game for balancing the charging of ride-hailing electric fleets. (Under review. <https://arxiv.org/abs/2210.08496>)
- Maljkovic, M., Nilsson, G., & Geroliminis, N. (2022b). A pricing mechanism for balancing the charging of ride-hailing electric vehicle fleets. In *2022 European Control Conference (ECC)* (p. 1976-1981). doi: 10.23919/ECC55457.2022.9838005
- Maljkovic, M., Nilsson, G., & Geroliminis, N. (2023). On finding the leader's strategy in quadratic aggregative stackelberg pricing games. (Under review. https://www.dropbox.com/s/epkiprr9fm8a44z/ECC_2023.pdf?dl=0)
- Paccagnan, D., Gentile, B., Parise, F., Kamgarpour, M., & Lygeros, J. (2016). Distributed computation of generalized Nash equilibria in quadratic aggregative games with affine coupling constraints. , 6123-6128. doi: 10.1109/CDC.2016.7799210
- Paccagnan, D., Gentile, B., Parise, F., Kamgarpour, M., & Lygeros, J. (2019). Nash and Wardrop equilibria in aggregative games with coupling constraints. *IEEE Transactions on Automatic Control*, 64(4), 1373-1388. doi: 10.1109/TAC.2018.2849946
- Rosen, J. B. (1965). Existence and uniqueness of equilibrium points for concave N -person games. *Econometrica*, 33(3), 520-534.
- Staňková, K., Olsder, G., & Bliemer, M. (2011). Bi-level optimal toll design problem solved by the inverse stackelberg games approach. *WIT Transactions on The Built Environment*, 89, 871-880. doi: 10.2495/UT060841
- Tushar, W., Saad, W., Poor, H. V., & Smith, D. B. (2012). Economics of electric vehicle charging: A game theoretic approach. *IEEE Transactions on Smart Grid*, 3(4), 1767-1778. doi: 10.1109/TSG.2012.2211901
- Yu, Y., Su, C., Tang, X., Kim, B., Song, T., & Han, Z. (2021). Hierarchical game for networked electric vehicle public charging under time-based billing model. *IEEE Transactions on Intelligent Transportation Systems*, 22(1), 518-530. doi: 10.1109/TITS.2020.2994192
- Zavvos, E., Gerding, E. H., & Brede, M. (2022). A comprehensive game-theoretic model for electric vehicle charging station competition. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 12239-12250. doi: 10.1109/TITS.2021.3111765
- Zhang, J., Lu, J., Cao, J., Huang, W., Guo, J., & Wei, Y. (2018). Traffic congestion pricing via network congestion game approach. *Discrete and Continuous Dynamical Systems - S*, 14. doi: 10.3934/dcdss.2020378
- Zhang, L., Gong, K., & Xu, M. (2019). Congestion control in charging stations allocation with Q-learning. *Sustainability*, 11(14). doi: 10.3390/su11143900

Contributions of Can, May and Want to the Home Office Frequency Decision

Daniel Heimgartner*¹ and Kay W. Axhausen¹

¹Institute for Transport Planning and Systems, ETH Zurich, Switzerland

*Corresponding author: daniel.heimgartner@ivt.baug.ethz.ch

SHORT SUMMARY

Camawa. Can, may and want. These are the constituents of the home office frequency decision. Not every job can be done from home nor is it a matter of all or nothing. Every job profile can be positioned on a continuum reflecting its home office feasibility. Further, those who can, might not may: Firms call back employees to the office or set constraints such as a home office budget. Last but not least *cama* does not mean anything without *wa* - the preference dimension. This work tries to account for all three dimensions simultaneously by means of a structural equation model (SEM). We find that the *may* dimension is of most substance and an employee's perception of her employer's point of view plays a crucial role in it. Meanwhile, preferences are governed by several suitability considerations. *Personal suitability*, *residential suitability* and the suitability of the *home office workstation* play into the decision, perceived personal suitability being the most important of the three.

Keywords: Home office, Preferences, Structural equation modeling.

1 INTRODUCTION

The ability to shift work from the office to home varies greatly across industries, cities and countries (Dingel & Neiman, 2020). While the question of how many jobs can be done from home has been widely discussed, it should be acknowledged that a job's home office feasibility is not binary. Sener & Bhat (2011) argue that when modeling the home office frequency, one should first estimate whether or not a job can be done from home. But even if the characteristics of work would allow for home office it is not guaranteed that the employee may shift to remote nor is it given that the employee wants to do so. After all, observed home office frequencies reflect a labor market equilibrium and should therefore account for both home office supply and demand.

This work tries to quantify the contributions of *can*, *may* and *want* to the home office frequency decision with a structural equation modeling (SEM) approach. Is it a supply-driven (home office supply of the employer) market or is it demand-driven? Is a job's home office feasibility accounted for in the current market or is there an inefficiency arising from too much home office (as the pandemic and current full employment shifted the momentum and bargaining power to the workforce, asking for unreasonable high levels of remote work)?

It can be argued, that before the pandemic, home office was the exception rather than the rule. "Shirking from home" was stigmatized and perceived to be bad for career advancements and therefore workers were afraid to postulate their desire (Brewer & Hensher, 2000). However, this perception has drastically changed in recent years, but might still play a role. This work tries to elicit whether or not the perceived viewpoint of the employer matters in the employee's decision-making process.

Further, modeling a person's preference for home office as a latent construct, allows us to elicit the constituents of that preference by differentiating perceived personal suitability, the suitability of the residential environment as well as the home office workstation.

2 METHODOLOGY

The data was collected as part of a pre-test fielded in February 2023 in the German-speaking part of Switzerland. 886 respondents were invited by mail. The response rate was 24%, however, after the exclusion criteria, a sample of 148 participants remained. For the modeling part, only people

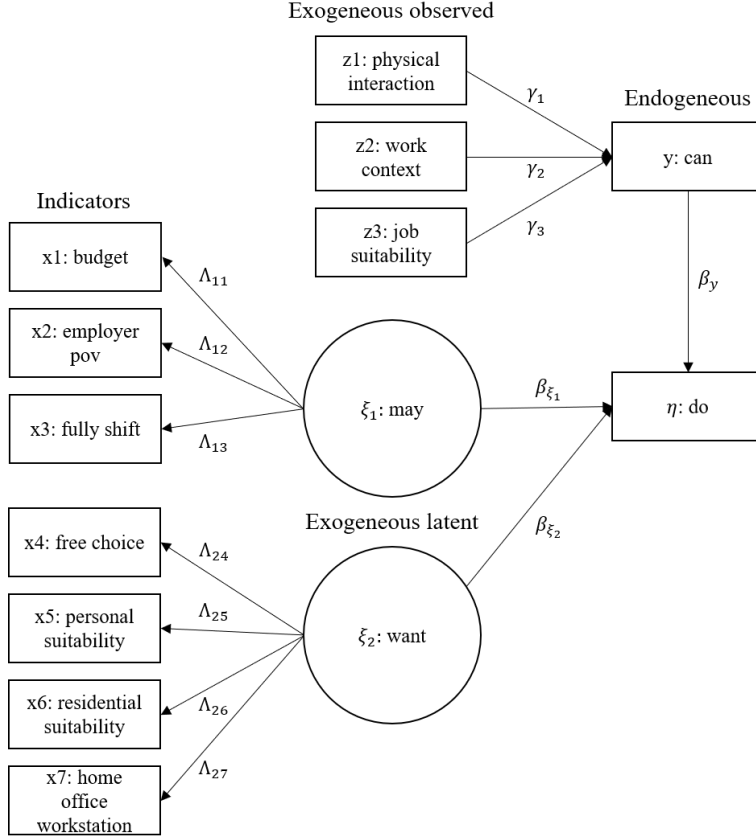


Figure 1: SEM path diagram with two latent variables and one explanatory endogenous variable

currently working from home were included. This was necessary because the questions related to home office were only asked the respondents currently working from home.

Table 1 describes the 12 variables which were derived from the survey answers for modeling purposes. The five-point Likert questions were simplified to binary indicators where the median value was chosen for the cutoff. For example, the variable *personal suitability* discriminates people into the following two classes: Higher or equal personal suitability (for home office) than the median person's perceived suitability or below. This ensures sufficient variation in the indicators.

We employ a structural equation model with two latent exogenous variables (*may* and *want*) and one endogenous observed variable (*can*). The final regression of interest is concerned with how these three dimensions impact the home office frequency decision. Figure 1 shows the envisioned path diagram.

The SEM equations are depicted below and consist of measurement equations eq. (1), an equation capturing the explanatory endogenous variable eq. (2) as well as linking the two in a structural equation eq. (3). SEM allows the modeler to simultaneously estimate these equations and account for complex correlation patterns. The model reads

$$\mathbf{x}_i = \Lambda \boldsymbol{\xi}_i + \boldsymbol{\delta}_i \quad (1)$$

$$y_i = \mathbf{z}_i^t \boldsymbol{\gamma} + \epsilon_i \quad (2)$$

$$\eta_i = \beta_y y_i + \beta_{\xi_1} \xi_{i1} + \beta_{\xi_2} \xi_{i2} + \zeta_i \quad (3)$$

where \mathbf{x}_i is the vector of measurement indicators, Λ the corresponding matrix of factor loadings (with 0 for some elements), $\boldsymbol{\xi}_i$ the vector of latent exogenous variables (*may* and *want*), \mathbf{z}_i the vector of (observed) exogenous variables, $\boldsymbol{\gamma}$ the corresponding coefficients, explaining y_i , an explanatory endogenous variable (*can*) and η_i the endogenous target of interest (*do*: the home office frequency choice). The vector $\boldsymbol{\beta}$ captures the main effects of interest, i.e., the impact of *can*, *may* and *want* on the observed home office frequency choice (η_i). $\boldsymbol{\delta}_i$, ϵ_i and ζ_i are random errors.

At this point, it should be noted that ordinal scaled variables (e.g., *budget*: maximum number of days allowed to work from home $\{1, \dots, 5+\}$) were treated as continuous. While we tested ordered

Table 1: Model variables.

SEM	Variable	Binary	Question
Exogenous observed	physical interaction	Yes	My job requires physical/interpersonal interaction which cannot be compensated by digital channels.
	work context	Yes	My job requires a specific work environment (e.g., equipment, safety precautions, working outdoors, etc.).
	job suitability	Yes	How suitable do you consider your main occupation for home office?
Indicators	budget	No	Does your employer/manager set a maximum number of days per week where you can do home office?
	employer pov	Yes	Does your employer like the idea of home office?
	fully shift	Yes	Could you shift all work that can be done remote to home office, without feeling pressured to return to your regular work place more often?
	free choice	Yes	If you could choose freely, how many days per week would you work from home?
	personal suitability	Yes	How suitable do you consider yourself as a person for home office?
	residential suitability	Yes	How suitable do you consider your residential environment (distraction through family, noise, number of rooms, etc.) for home office?
	home office workstation	Yes	How suitable do you consider your home office workstation for home office?
Endogenous observed	can	No	How much of your work time could you shift to home office?

Table 2: Cross tables.

Can	May			Budget					Want					
	No	Yes	COVID-19	1	2	3	4	5+	0	1	2	3	4	5+
0%	16.22	2.03	2.7	0	0	0	0	0	0	0	0	0	0	0
1 – 25%	1.35	13.51	1.35	3.7	3.7	0.93	0	9.26	1.71	11.97	5.98	0.85	0	0
26 – 50%	1.35	25	0.68	4.63	15.74	0	0	12.96	5.98	4.27	17.09	6.84	0	0
51 – 75%	0	10.81	0	0	5.56	1.85	0	7.41	2.56	0.85	2.56	5.13	2.56	0
76 – 99%	0	18.92	0	3.7	3.7	5.56	0.93	12.04	0	1.71	1.71	7.69	10.26	2.56
100%	0	6.08	0	0.93	1.85	2.78	0	2.78	1.71	0.85	2.56	0.85	0.85	0.85

Want	May			Budget					Do					
	No	Yes	COVID-19	1	2	3	4	5+	0	1	2	3	4	5+
0	0.85	11.11	0	0.93	5.56	1.85	0	3.7	11.82	7.27	4.55	0	0	0
1	0.85	16.24	2.56	3.7	5.56	0.93	0	7.41	0	8.18	13.64	3.64	1.82	0
2	0.85	29.06	0	5.56	10.19	0.93	0	12.96	0	1.82	10	6.36	0	0
3	0.85	20.51	0	1.85	6.48	1.85	0	12.04	0	0	2.73	7.27	5.45	0
4	0	13.68	0	0.93	2.78	4.63	0.93	5.56	0	0	0	2.73	2.73	0.91
5+	0	3.42	0	0	0	0.93	0	2.78	0	0	0	1.82	4.55	2.73

logit models (e.g. for the before-mentioned measurement equation as well as all the others), we chose not to, as the additional cutoff parameters to be estimated in an ordered logit would yield very few observations per parameter. However, the specifications were tested and did not lead to alternative conclusions.

The model was estimated with the `lavaan` package (Rosseel, 2012) in R using the maximum likelihood approach.

3 RESULTS AND DISCUSSION

In what follows, the cross tables in table 2 are abbreviated: For example *cama* stands for *can* x *may* and the reported values reflect percentage numbers. As a side note, the dimensions *may* and *want* should not be mistaken for the latent variables. *May* indicates whether or not an individual is currently allowed to do home office and *want* is approximated by the variable *free choice* from table 1.

Cama clearly shows that those who can at least partially do some work from home, also may. *Cabu* indicates that roughly half of the home office population have agreements, fixing the maximum number of home office days (5+ means no constraints). The budget seems not correlated with a job's ability to be performed remotely. A budget of 2 days is the most common constraint. In the *wa* of *cawa*, the respondents were asked to realistically factor in their job characteristics. Therefore the two dimensions are correlated. In each row, the modulus shifts to the right and matches the job's ability to be performed remotely. This hints that people generally would like to shift all the work that can be productively completed in the home office to remote.

Shifting attention to the second row of table 2: Those who want, may. Only very few people are not allowed to do home office. *Wabu* hints that the employer decides on the budget more or less unilaterally. This leads to 33% of the employees being constrained in their frequency choice. However, most of the workforce still can shift their desired number of days to the home office: *cado* shows a strong correlation with most of the mass clustering around the diagonal. Still, the upper triangle has slightly more mass (which makes sense given the previously discussed "budget" constraint).

We now discuss the modeling results presented in table 3. It should be noted, that there are relatively few observations per estimated parameter (roughly 5 per parameter). Nevertheless, standard errors are small. The goodness of fit statistics indicates mediocre fit (for a nice discussion of how to interpret these measures, see Lin (2021)). This is not surprising given the very simplified model (in terms of model specification, binary feature engineering and linear approximation of ordered scales).

With this in mind, the factor loadings and regression coefficients all have the expected signs and most of them are significant. *May* shows the most substantial contribution to the home office frequency choice followed by *want* and *can*.

We tested to include the allowed home office budget directly as an exogenous explanatory variable (now solely reflecting the *may* dimension), the effect of *want* becomes dominant. However, modeling *may* as a latent variable (as we did here), including measurements of whether or not the

Table 3: Model coefficients and goodness of fit indicators.

Parameter Estimates:				
	Latent Variable	Indicator	Loading	Std.Err
Latent Variables:	may	budget	1.000	
		employer pov	0.197*	0.082
		fully shift	0.234*	0.101
	want	free choice	1.000	
		personal suitability	1.256***	0.235
		residential suitability	0.803***	0.182
		homeoffice workstation	0.787***	0.176
Regressions:	Dependent Variable	Predictor	Estimate	Std.Err
	can	physical interaction	-0.378	0.237
		work context	-0.344	0.244
		job suitability	1.007***	0.223
	do	can	0.402***	0.092
		may	2.439*	1.333
		want	1.628	1.120
Goodness of Fit:				
	Comparative Fit Index (CFI)	0.702		
	Tucker-Lewis Index (TLI)	0.616		
	RMSEA	0.122		
Model Characteristics:				
	Number of model parameters:	23		
	Number of observations:	108		

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

person feels pressured to return to the office more frequently, as well as the employer’s point of view about home office, the substance changes in favor of *may*. This could indicate that contractual agreements play a minor role whereas an individual’s perception of her employer’s viewpoint still plays an important role.

The regression on *can* could suffer from endogeneity (with *job suitability* accounting for both *work context* as well as *physical interaction*). Therefore it is unsurprising that only the *job suitability* was found to be significant. In future versions of the model (when more data is available), *can* should be treated as latent too, including the proposed predictors as measurements.

We now discuss the latent variables and their factor loadings. The variable *budget* shows the highest loading for *may* which reflects the previously discussed fact, that 33% of the employees are constrained in their free choice. On the other hand and as already noted, the employees’ perception of the employer’s standpoint is important too.

Interestingly, the preference dimension (*want*) reveals that the *personal suitability* loads most heavily. In future research, we will include (latent) personality traits as predictors. Meanwhile, *residential suitability*, as well as the suitability of the *homeoffice workstation*, load with similar magnitudes. The questions of what makes a residence or workstation suitable is left to future planned research.

4 CONCLUSIONS

We used a SEM model with two exogenous latent variables and one endogenous predictor to understand the home office frequency decision. The center of attention was placed on disentangling the contributions of *can*, *may*, and *want*.

We find that the *may* dimension is most decisive and an employee’s perception of her employer’s point of view plays a crucial role in it. On individuals’ preferences, we can note, that all three suitability dimensions, *personal suitability*, *residential suitability* and the suitability of the *home office workstation*, are equally important (with a slightly higher loading of *personal suitability*).

To our best awareness, this is the first model, that accounts for *can*, *may* and *want* simultaneously. The model should be extended, once more data is available: *Can* should be modeled as a latent variable, ordered logit models should be used where appropriate and the latent variables should be treated as endogenous rather than exogenous (allowing us to delve more deeply into the questions of *why* a person has certain home office preferences or what industries and employer characteristics explain the *may* or *can* dimension). We showed that all three dimensions *can*, *may* and *want* matter and should be accounted for when modeling observed home office frequencies.

ACKNOWLEDGEMENTS

This work was inspired by the discussions with Benjamin Gramsch and his learnings on structural equation modeling. The paper could benefit from his hEART contribution.

REFERENCES

- Brewer, A. M., & Hensher, D. A. (2000). Distributed work and travel behaviour: The dynamics of interactive agency choices between employers and employees. *Transportation*, *27*, 117–148.
- Dingel, J. I., & Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, *189*, 104235.
- Lin, J. (2021). *Introduction to structural equation modeling (SEM) in R with lavaan*. Los Angeles, California,.
- Rosseel, Y. (2012). lavaan: An r package for structural equation modeling. *Journal of statistical software*, *48*, 1–36.
- Sener, I. N., & Bhat, C. R. (2011, January). A copula-based sample selection model of telecommuting choice and frequency. *Environ Plan A*, *43*, 126–145.

Household fleet adaptation as reaction to price regulations: A stated adaptation experiment on the promotion of electric vehicles

Gutjar, Margarita*¹, Calastri, Chiara², Kowald, Matthias³

¹Msc, Department of Architecture and Civil Engineering,
RheinMain University of Applied Sciences, Germany

²Dr., Institute of Transport Studies, University of Leeds

³Prof., Department of Architecture and Civil Engineering,
RheinMain University of Applied Sciences, Germany

SHORT SUMMARY

The goal to limit global warming requires a shift to electric vehicles and a reduction of vehicles in total. To achieve this transition, governments could design price regulations effectively. The potential effect of different price regulations has been assessed by surveying 466 respondents. After providing detailed information on all mobility tools in the household, respondents were faced with four scenarios with varying price regulations concerning prices for fuel, CO₂, electricity, and public transport. Given the reported mobility tools and supported by live calculation of resulting cost changes, respondents were asked to adapt their household fleet while being allowed to choose the mobility tools at a high level of detail. Results of a multinomial logit model show that increasing fuel prices, very low electricity prices, high EV subsidies and low public transport prices have the potential to decarbonize household fleets (remove conventional vehicles and/or replace by an electric vehicle).

Keywords: sustainable mobility, electric vehicle, political incentives, multinomial logit model, stated adaptation

1. INTRODUCTION

To reduce greenhouse gas emissions in the transport sector a shift from vehicles with internal combustion engine (ICE) to electric vehicles (EV) is required. For this purpose, governments are implementing policies to promote EVs. However, several studies investigate either isolated economic interventions such as fuel prices (see e.g. Erath and Axhausen, 2010; Jäggi et al., 2012; Liao et al., 2017). They show that only a great increase of fuel prices have the potential to increase the market share of alternative fuel vehicles (Jäggi, 2015; Lebeau et al., 2012). However, not only fuel price is of relevance to promote the shift from ICEs to EVs. Research has shown that higher operating costs have a negative effect on the preference of a vehicle (Beck et al., 2017; Helveston et al., 2015; Higgins et al., 2017; Jensen et al., 2020; Li et al., 2020). Free charging has a positive effect on adoption of EV and is after free parking the incentive with the highest willingness-to-pay for a vehicle (Langbroek et al., 2016). Further, persons intending to buy a new vehicle prefer lower purchase costs (Helveston et al., 2015). Studies show that price subsidies have a positive effect on the choice of EVs (Bjerkkan et al., 2016; Higgins et al., 2017; Lebeau et al., 2012) and on the diffusion of EVs in general (Buchmann et al., 2021; Melton, 2020). Therefore, not only effects of fuel prices, but also EV purchase subsidies and electricity prices need to be investigated.

However, large majority of research either conduct stated preference studies to show, which prices have an effect on the choice of a vehicle type (Beck et al., 2017; Bjerkan et al., 2016; Helveston et al., 2015; Higgins et al., 2017; Jäggi, 2015; Jensen et al., 2020; Langbroek et al., 2016; Lebeau et al., 2012; Li et al., 2020) or simulation studies on market diffusion potentials (Buchmann et al., 2021; Melton, 2020). Revealed preferences studies mostly investigate the effect of socio-demographics (Brückmann et al., 2021; Jakobsson et al., 2016). However, to our best knowledge, there are no studies, which would model the effects of price regulations on fostering households' adaption of their household fleet. Therefore, this study is aiming to analyze the effectiveness of fuel prices, CO₂ surcharge on fuel prices, electricity prices, and EV purchase subsidies on household's stated decisions to either adopt an EV, replace an ICE by an EV, and to remove an ICE. Besides that, the effect of reduced prices for public transport will be considered, since the decarbonization of the transport sector requires not only the adoption of EVs but a general reduction in vehicle usage.

METHODOLOGY

Data and sample

Adults of 18 years and older were recruited via an introduction letter and a follow-up recruitment phone call from a sample of 6,107 addresses in the South-West of Germany. Computer-assisted personal interviews were conducted from January to December 2020. In total, 466 individuals completed the survey providing information on the household (for details on fieldwork see Gutjar et al. (2021); Gutjar and Kowald (2021a)). After data cleaning and exclusion of households without persons owning a driver license, data from 444 respondents will be considered for analyses.

Survey design

Firstly, the respondents provided information on the household (e.g. household income, housing type), the sociodemographic characteristics, and mobility behavior of every household member (e.g. age, gender, car availability). Further, to answer the research question a two-stage process was created:

In the first stage, revealed preferences (RP) for mobility tools in the household fleet were collected. Respondents provided detailed information on all vehicles (e.g. vehicle type, engine type, annual vehicle kilometers traveled (VKM)), motorcycles, and public transport subscriptions available.

Next, a stated adaptation experiment (Lee-Gosselin, 1996) was designed to assess the effect of price regulations, which are presented in **Table 1** together with their variation. Based on these price attributes and variation levels, an efficient experiment design (Rose and Bliemer, 2014) was created in Ngene (Rose et al., 2018) resulting in 20 scenarios divided into five blocks, so that every respondent was faced with four different tasks. Each task was designed as an iterative adaptation of the household fleet under a given scenario: Employing the RP data on mobility tools in the household, the survey program initially calculated the actual household fleet costs and additionally presented changes in monthly and annual costs for the actual household fleet as a consequence of the hypothetical price regulations given in the scenario. Respondents were asked to react to the scenario by adapting the household fleet under consideration of the financial restrictions and the mobility needs of the household. They could e.g. remove present and/or add new vehicle(s), motorcycle(s), and public transport subscriptions and adjust the annual VKM. For every adopted vehicle, a vehicle type and an engine type (gasoline, diesel, BEV, PHEV) had to be specified. Respondents were supported by a real-time calculation of the monthly and annually

household fleet costs to allow a comparison of the resulting and current costs after every adaptation. Thus, they were able to adjust the mobility tools until they found the optimal household fleet under given price regulations (for details on the study design see Gutjar and Kowald (2021b) and Reckermann et al. (2021)). Finally, n=1,737 observations (stated adaptations) from 466 individuals will be analyzed.

Table 1. Stated adaptation experiment: price attributes and variation levels (Gutjar and Kowald, 2021b)

price attribute	variation levels	model changes
fuel price (€/l)	1.50* / 3.00 / 4.50	sum (continuous): 0 / 0.20 / 0.60
CO ₂ surcharge (€/liter fuel)	0.00* / 0.20 / 0.60	/ 1.50 / 1.70 / 2.10 / 3.00 / 3.20 / 3.60
electricity price (€/100km)	0.00 / 3.50* / 7.00	-3.50 / 0 (reference) / +3.50
purchase bonus for EVs (€)	2,000 / 6,000* / 10,000	-4,000 / 0 (reference) / +4,000
public transport prices relative to today	free / 50% of today's price / as today*	-100 / -50 / 0 (reference)

Note: * = value at the time of fieldwork

Estimation

Since the aim of this study is to model adaptation (changes) of the household fleet, changes in price regulations in comparison to the reference values at the time of fieldwork will be modelled. For this purpose, fuel prices and CO₂ surcharge were summed to one continuous variable, while dummy variables were created for the remaining price regulations with no change in prices incorporated as reference category. The created variables to model changes in correspondence to the variation of the price regulations are presented in **Table 1**.

The changes between RP vehicle ownership and the final adapted household fleet as reaction to price scenarios will be modelled as the outcome with reference to no change (no adaptation to vehicle ownership):

- **add an EV** (if an EV was added as an additional vehicle to the household fleet; the alternative is always available)
- **remove ICE** (if at least one existing ICE was removed from the household fleet; the alternative is available to households with min. one ICE)
- **replace ICE by EV** (if at least one existing ICE was removed but an EV was included instead; the alternative is available to households with min. one ICE)
- **remove and replace ICE** (if at least two existing ICEs were removed and an EV included instead; the alternative is available to households with min. two ICEs)

The frequencies (absolute and relative) of alternative availability and actual choices (n=1,737 observations) are presented in **Table 2**.

Table 2. Description of choices (adaptations)

choice	n available	n chosen	% chosen overall	% chosen when available
no change	1,737	1,191	68.57	68.57
add an EV	1,737	74	4.26	4.26
remove ICE(s)	1,605	123	7.08	7.66
replace ICE(s) by EV	1,605	306	17.62	19.07
remove & replace ICE(s) by EV	701	43	2.48	6.13

Total	1,737	100
<i>Note: changes not modeled due to a small number of observations (n≤10): add ICE, remove EV, replace ICE by EV & add EV</i>		

Since the presented adaptation alternatives are discrete choices, they will be analyzed by applying the random utility maximization theory, which assumes respondents rationally choose the alternative with the highest associated utility (Adamowicz et al., 1994; Louviere et al., 2010). An individual n confronted with j alternatives in t choice tasks associates an indirect utility U_{njt} for an alternative and chooses the alternative with the highest utility. The indirect utility U_{njt} of an alternative j is decomposed as

$$U_{njt} = V_{njt} + \varepsilon_{njt} = x'_{njt}\beta + \varepsilon_{njt} \quad (1)$$

where U_{njt} is not observed, V_{njt} is the deterministic utility of alternative j , and ε_{njt} is a random component not included in V_{njt} ; V_{njt} can be specified by $x'_{njt}\beta$, where x is a vector of explanatory variables (e.g. attribute levels, socio-demographics), and β are the coefficients to be estimated. Further, alternative-specific constants (ASC) (Train, 2009) were estimated.

2. RESULTS AND DISCUSSION

Data analyses were performed with R using the package *apollo* (Hess and Palma, 2019; R Core Team, 2020) for the step-wise estimation of a multinomial logit model (MNL) (Louviere et al., 2000; Train, 2009). The results of the current MNL model are presented in **Table 3** (previous steps and sample descriptions are available upon request).

As expected, increasing fuel price (including CO₂ surcharge) has a positive effect on the decision to remove an ICE, replace an ICE with EV, and do both. A drastic reduction in electricity prices (for free) increases the utility to replace an ICE with an EV and to do both, remove and replace an ICE. However, it has also a positive effect on the adoption of an EV as an additional vehicle, which needs to be considered (rebound effect). Correspondingly, increased electricity prices decrease the utility to add an EV but also to adopt one as a replacement for an ICE. Interestingly, for the purchase bonus, no strong effects have been found. However, the reduction of the EV subsidy by 4,000€ reduces the utility to add an EV and to choose both remove and replace an ICE, while an increase by 4,000€ has a positive utility on the replacement of an ICE with an EV. Free public transport (-100%) has a positive impact on the removal of an ICE and to do both remove and replace existing ICEs.

In comparison to no change, the utility of adding an EV and doing both removing and replacing an ICE is decreasing with age, while it is firstly positive but becomes negative with higher age for the alternatives to remove an ICE or to replace an ICE. Interestingly, while highly educated respondents (in comparison to low-middle education) prefer to remove an ICE, replace an ICE and do both, low-educated persons prefer to adopt an EV as an additional vehicle (e.g. BEV might serve as a status symbol). Higher equalized household income (considering the number and age of household members) decreases the utility of removing an ICE and to do both replacing and removing an ICE, but no remarkable effect was shown for the replacement of an ICE with an EV. Further, with increasing income, the utility to add an EV increases. To sum up, with greater income households associate disutility with alternatives required for transport decarbonization, because they can afford to keep their status quo. Intuitively, households with an equal or greater number of vehicles than persons with driver's license associate greater utility with removing and replacing an ICE, while they show a lower preference for adding an EV than persons in a household with fewer vehicles than drivers. Similarly, households with greater VKM associate increasing utility with removing, replacing, and doing both removing and replacing an ICE by EV.

Table 3. Results of the MNL

	add EV			remove ICE			replace ICE			remove & replace ICE		
	β	r. se.	r. t-val.	β	r. se.	r. t-val.	β	r. se.	r. t-val.	β	r. se.	r. t-val.
fuel price	-			0.61	0.10	5.94	0.50	0.06	8.78	0.68	0.14	4.96
electricity price (Ref. no change)												
minus 3.50€	0.64	0.29	2.23	-			0.85	0.17	5.04	1.29	0.31	4.14
plus 3.50€	-0.68	0.39	-1.74	-			-0.40	0.22	-1.78	<i>fixed</i>		
purchase bonus (Ref: no change)												
minus 4,000€	-0.50	0.23	-2.16				<i>fixed</i>			-0.48	0.31	-1.56
plus 4,000€	<i>fixed</i>						0.21	0.14	1.47	<i>fixed</i>		
public transport (Ref: no change)												
minus 100%	-			0.21	0.19	1.10	-					
plus 100%	-			0.45	0.20	2.30	-			0.47	0.35	1.35
ASC	-1.72	0.79		-7.41	2.44	-3.04	-4.19	1.13	-3.72	-2.46	1.07	-2.30
age	-0.03	0.01	-2.18	0.15	0.09	1.68	0.08	0.05	1.68	-0.03	0.02	-1.43
age2	<i>fixed</i>		-1.99	-0.00	0.00	-1.78	-0.00	0.00	-2.06	<i>fixed</i>		
education												
low	<i>Ref.</i>											
middle				<i>Ref.</i>			<i>Ref.</i>			<i>Ref.</i>		
high	-0.66	0.42	-1.56	0.54	0.30	1.78	0.32	0.221	1.44	0.80	0.57	1.41
equivalised household income	0.48	0.16	2.92	-0.25	0.14	-1.76	<i>fixed</i>			-0.61	0.30	-2.03
n vehicles \geq n drivers (Ref: less)	-0.76	0.35	-2.19	0.45	0.31	1.44	0.23	0.206	1.09	<i>fixed</i>		
VKM	<i>fixed</i>			0.01	0.01	1.19	0.01	0.008	1.74	0.03	0.01	2.52
Number of individuals	444											
LL (final)	-1452.59											
Adj.Rho-square	0.3976											
AIC	2979											
BIC	3181											

Note: r. se. = robust standard error; r. t-val. = robust t-value; - = not estimated (e.g. no previous hypotheses); fixed = parameter fixed to zero during the step-wise estimation procedure (e.g. small and insignificant parameter, small number of observations (n<20)), Ref.= Reference category

3. CONCLUSIONS

Given study contributes to previous research by modeling potential adaptations of vehicle ownership as a consequence of political incentives and other price regulations relevant to transport decarbonization. Preliminary results have been presented. Next, this model will be extended by interactions of price attributes with socio-demographic characteristics to explain taste heterogeneity (e.g. are households with greater household income less sensitive towards price increases?). Further, an integrated choice latent variable (ICLV) (Abou-Zeid and Ben-Akiva, 2014) model will implement the effect of the latent factor intention to buy an EV as a direct predictor of behavior (Ajzen, 1991). All results along with policy implications will be presented at the conference if accepted.

ACKNOWLEDGEMENTS

This work is part of the project "Electric City Rüsselsheim" funded by the German Federal Ministry for Economic Affairs and Energy; Grant number: 01MZ18008B.

REFERENCES

- Abou-Zeid, M., Ben-Akiva, M., 2014. Hybrid choice models, in: Handbook of Choice Modelling. Edward Elgar Publishing.
- Adamowicz, W., Louviere, J., Williams, M., 1994. Combining revealed and stated preference methods for valuing environmental amenities. *J. Environ. Econ. Manag.* 26, 271–292. <https://doi.org/10.1006/jeem.1994.1017>
- Ajzen, I., 1991. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211.
- Beck, M.J., Rose, J.M., Greaves, S.P., 2017. I can't believe your attitude: a joint estimation of best worst attitudes and electric vehicle choice 21.
- Bjerkan, K.Y., Nørbech, T.E., Nordtømme, M.E., 2016. Incentives for promoting Battery Electric Vehicle (BEV) adoption in Norway. *Transp. Res. Part Transp. Environ.* 43, 169–180. <https://doi.org/10.1016/j.trd.2015.12.002>
- Brückmann, G., Willibald, F., Blanco, V., 2021. Battery Electric Vehicle adoption in regions without strong policies. *Transp. Res. Part Transp. Environ.* 90, 102615. <https://doi.org/10.1016/j.trd.2020.102615>
- Buchmann, T., Wolf, P., Fidaschek, S., 2021. Stimulating E-Mobility Diffusion in Germany (EMOSIM): An Agent-Based Simulation Approach. *Energies* 14, 656.
- Gutjar, M., Kowald, M., 2021a. Political Incentives to Promote Electric Vehicles: Results From a Stated Adaptation Experiment, in: Ryan C. P., W., Jiangping, Z., W. Y., S. (Eds.), Sustainable Mobility - Proceedings of the 25th International Conference of Hong Kong Society for Transportation Studies (HKSTS). Hong Kong Society for Transportation Studies Limited, Hong Kong, pp. 215–222.
- Gutjar, M., Kowald, M., 2021b. Electro mobility acceptance: The influence of political bonus and malus factors. Presented at the Symposium of the European Association for Research in Transportation, Symposium of the European Association for Research in Transportation, Lyon, France.
- Gutjar, M., Plotz, N.E., Kistner, R., Kowald, M., 2021. Feldbericht zur sozialwissenschaftlichen Begleitforschung in „Electric City Rüsselsheim (Arbeitsberichte Fachgruppe Mobilitätsmanagement No. 014). Hochschule RheinMain, Wiesbaden.
- Helveston, J.P., Liu, Y., Feit, E.M., Fuchs, E., Klampfl, E., Michalek, J.J., 2015. Will subsidies drive electric vehicle adoption? Measuring consumer preferences in the US and China. *Transp. Res. Part Policy Pract.* 73, 96–112.
- Hess, S., Palma, D., 2019. Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *J. Choice Model.* 32, 100170. <https://doi.org/10.1016/j.jocm.2019.100170>
- Higgins, C., Mohamed, M., Ferguson, M., 2017. Size Matters: How Vehicle Body Type Affects Consumer Preferences for Electric Vehicles. *Transp. Res. Part Policy Pract.* 100, 182–201. <https://doi.org/10.1016/j.tra.2017.04.014>
- Jäggi, B., 2015. Decision Modelling on Household Level for Energy, Fleet Choice and Expenditure (PhD Thesis). ETH Zurich.
- Jakobsson, N., Gnann, T., Plötz, P., Sprei, F., Karlsson, S., 2016. Are multi-car households better suited for battery electric vehicles? – Driving patterns and economics in Sweden and

- Germany. *Transp. Res. Part C Emerg. Technol.* 65, 1–15.
<https://doi.org/10.1016/j.trc.2016.01.018>
- Jensen, A.F., Thorhauge, M., Mabit, S.L., Rich, J., 2020. Analyses of EV buying preferences SP method and model (DTU Management Rapport). Technical University of Denmark.
- Langbroek, J.H., Franklin, J.P., Susilo, Y.O., 2016. The effect of policy incentives on electric vehicle adoption. *Energy Policy* 94, 94–103.
- Lebeau, K., Van Mierlo, J., Lebeau, P., Mairesse, O., Macharis, C., 2012. The market potential for plug-in hybrid and battery electric vehicles in Flanders: A choice-based conjoint analysis. *Transp. Res. Part Transp. Environ.* 17, 592–597.
- Lee-Gosselin, M.E.H., 1996. Scope and potential of interactive stated response data collection methods, in: *Conference on Household Travel Surveys: New Concepts and Research Needs*. pp. 12–15.
- Li, L., Wang, Zhiqiang, Chen, L., Wang, Zixuan, 2020. Consumer preferences for battery electric vehicles: A choice experimental survey in China. *Transp. Res. Part Transp. Environ.* 78, 102185.
- Louviere, J.J., Flynn, T.N., Carson, R.T., 2010. Discrete choice experiments are not conjoint analysis. *J. Choice Model.* 3, 57–72.
- Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. *Stated choice methods: analysis and applications*. Cambridge university press.
- Melton, N., 2020. Which plug-in electric vehicle policies are best? A multi-criteria evaluation framework applied to Canada. *Soc. Sci.* 15.
- R Core Team, 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Reckermann, H., Kistner, R., Gutjar, M., Kowald, M., 2021. Die Berechnung von Pkw-Kosten: Vorstellung bestehender Ansätze und einer eigenen Methode (Arbeitsberichte Fachgruppe Mobilitätsmanagement No. 007). Hochschule RheinMain, Wiesbaden.
- Rose, J.M., Bliemer, M.C.J., 2014. Stated choice experimental design theory: the who, the what and the why, in: *Handbook of Choice Modelling*. Edward Elgar Publishing, pp. 152–177.
- Rose, J.M., Collins, A.T., Bliemer, M.C.J., Hensher, D.A., 2018. *Ngene (Version 1.2.1)*.
- Train, K.E., 2009. *Discrete choice methods with simulation*. Cambridge university press, New York, NY, US.

Modeling Visit Probabilities within Space-Time Prisms of Daily Activity-Travel Patterns

Jing Lyu*¹, Feixiong Liao², Soora Rasouli³

¹PhD candidate, Urban Planning and Transportation Group, Eindhoven University of Technology, the Netherlands

²Assistant professor, Urban Planning and Transportation Group, Eindhoven University of Technology, the Netherlands

³Professor, Urban Planning and Transportation Group, Eindhoven University of Technology, the Netherlands

SHORT SUMMARY

Space-time prism (STP) delimits the space-time opportunities reachable by a moving object and is widely applied to measure the ability of individuals to travel and participate in activities. The majority STPs are binary measures in that all locations are considered equally accessible if within the prisms. A few probabilistic STP models discussed heterogeneous interiors, but they focus on the trip level and have not addressed daily activity programs with flexible activity sequences. This study proposes a model framework to construct and estimate the state-dependent probabilistic STP of daily activity-travel patterns based on the multi-state supernetwork representation. Utilizing GPS trajectories, the estimation and simulation results of visit probabilities in the STPs demonstrate the validity of the model framework.

Keywords: activity-travel patterns; multi-state supernetwork; space-time prism; visit probability

1. INTRODUCTION

Individuals' travel and activity participation are subject to space-time constraints. As a central time geographic concept, space-time prism (STP) delimits the space-time opportunities that can be reached by a moving object (Miller, 2017), and provides a measurement of potential mobility. The classic STP is determined by the known anchor points, time budget, the maximum attainable travel speed, and the time available for a flexible activity that can be conducted at one of multiple locations (Hägerstrand, 1970). By constructing the STP over a transportation network, the network-time prism delimits the accessible locations with respect to the spatial network (Miller, 1991). The boundary of an STP has been widely used as a space-time accessibility measurement, indicated by a spatial location set at instant times within time intervals.

The majority STPs are deterministic binary measures such that all locations are considered equally accessible if within the prisms, otherwise not accessible. In reality, a prism does not have homogeneous interiors. To capture the variations within a prism, Winter and co-workers (Winter, 2009; Winter and Yin, 2010a, 2010b) introduced probabilistic time geography and modeled probability distributions of an individual's potential visit locations at a time moment from a stochastic perspective. Song and Miller (2013) formulated discrete and continuous stochastic models for the movement of an individual within the potential path area (PPA). In Song et al. (2016, 2017), the visit probability within a directed STP is modeled using the approach of continuous-time semi-Markov process, describing the likelihood of visiting different locations within the prism.

In previous studies, STP is predominantly modeled at the trip level with single activities rather than for daily activity programs (APs). To capture the dependencies in an activity chain and improve realism, Chen and Kwan (2012) identified the location choice set based on STPs in the presence of all possible

activity-travel chains. Kang and Chen (2016) constructed the feasible space-time region for a daily AP by intersecting a set of feasible space-time regions for single activities. Liao (2021) applied bidirectional searches of full activity-travel patterns (ATPs) in multi-state supernetworks (SNK) (Liao, Arentze, and Timmermans, 2010, 2013) for delineating the exact STP for an AP. However, the probabilistic characteristics of accessibility within a prism have not been investigated for an AP with multiple activities and flexible activity sequences.

The aim of this study is, therefore, to propose a model framework to construct and estimate the probabilistic STP of an AP. The model framework includes three steps. First, based on the multi-state supernetwork, we construct the activity-based STP for a daily AP with flexible activity sequences (Liao, 2021). Second, we model the visit probability within the activity-based STP using semi-Markov techniques (Howard, 1971). Third, we estimate a visit probability model using extracted GPS trajectories and simulate a typical AP to demonstrate the validity of the model framework. The visit probabilities of an AP provide quantitative descriptions of the activity-based STP interiors and evaluations of the accessibility for an individual participating in multiple activities with flexible sequences.

2. METHODOLOGY

Multi-state supernetwork (SNK) representation and STP

Multi-state supernetworks are capable of representing ATPs of conducting an individual's AP. A daily AP's implementation is a path choice through networks of different states, including activity states specifying which activities have been conducted, and vehicle states specifying where the private vehicles are (in use or parked somewhere).

Denote a multi-state supernetwork as $SNK(N, E)$. The set of nodes N indicates locations in SNK , including road intersections, activity locations, and parking locations. A set of links E includes travel links of road segments, transaction links for conducting activities at activity locations, transition links for parking and picking up PVs, and boarding and alighting.

Given origin H_0 and destination H_1 as two anchors and the corresponding time budget $[t_{H_0}, t_{H_1}]$, the STP is constructed by deriving the potential path area (PPA) of an AP. The temporal feasibility of STP in SNK is formulated as follows (Liao, 2021):

$$\min\{g_{H_0}(n|_s) + g_{H_1}(n|_s)\} \leq t_{H_1} - t_{H_0} \quad (1)$$

where $n|_s$ denote node n at activity-vehicle state s , $g_{H_0}(n|_s)$ and $g_{H_1}(n|_s)$ are the actual activity-travel times from H_0 to $n|_s$ and $n|_s$ to H_1 respectively.

Visit probability within the STP

We define a status as a movement starting from node $i|_s$ to node $j|_{s'}$ along link $l_{ij|ss'}$ in SNK that has not arrived at $j|_{s'}$ yet, $l_{ij|ss'} = (i|_s, j|_{s'})$ for $\forall s, s'$. The status space includes the movements on all possible links within the STP. We formulate the holding time density functions of SNK links and the visit probability of each status at a moment in time.

(1) Holding time density functions

A holding time density function describes the probability that a transition from $i|_s$ to $j|_{s'}$, corresponding to the movement on link $l_{ij|ss'}$, will take extra time τ over the minimum time on $l_{ij|ss'}$, denoted as $f_{ij|ss'}(\tau)$ when $\tau \geq 0$, otherwise $f_{ij|ss'}(\tau) = 0$. The extra time τ on $l_{ij|ss'}$ can be calculated using

$\tau = \tau' - t_{ij|ss'}$, where τ' is the total time one traverse $l_{ij|ss'}$ and $t_{ij|ss'}$ is the minimum time expense from $i|_s$ to $j|_{s'}$. For travel link $l_{ij|ss'}$ that $i \neq j$ when $s = s'$, $t_{ij|ss'}$ is the minimum travel time. If conducting activity $\alpha \in A$ at location i ($i = j$) at state s results in a new state s' , that is $s \neq s'$, $l_{ij|ss'}$ is an activity link and $t_{ij|ss'}$ corresponding to the minimum activity duration d_α .

Considering parameters of $f_{ij|ss'}(\tau)$ are heterogeneous for travel and transaction links, we use the latent class models to capture the latent heterogeneity of holding times. Suppose there exist K different homogeneous latent classes in the heterogeneous population of extra travel times and activity durations. Let P_{qk} denote the class membership probability that an individual q belongs to latent class k :

$$P_{qk} = \frac{\exp(\boldsymbol{\beta}_k \mathbf{x}_q)}{\sum_{k=1}^K \exp(\boldsymbol{\beta}_k \mathbf{x}_q)}, \quad k = 1, \dots, K, \boldsymbol{\beta}_K = 0 \quad (2)$$

where \mathbf{x}_q is the vector of sociodemographic variables of individual q , $\boldsymbol{\beta}_k$ is the parameter vector of \mathbf{x}_q of latent class k . The latent class holding time density function for extra time is formulated as

$$f^q(\tau) = \sum_{k=1}^K P_{qk} \cdot f^q(\tau|k) \quad (3)$$

where $f^q(\tau|k)$ denotes the probability of individual q that belongs to class k spending extra time τ on a specific link in SNK , and $f^q(\tau)$ is the probability of unconditional holding time density.

(2) Visit probability formulation

For each link $l_{ij|ss'}$ in SNK , we can calculate a feasible time range as $(t_{i|s}^-, t_{j|s'}^+)$, where $t_{i|s}^-$ is the earliest arrival time and $t_{j|s'}^+$ is the latest departure time at $i|_s$ and $j|_{s'}$ in SNK , respectively. The visit probability in SNK as an extension over Song et al. (2016) is defined as follows.

Denote $P_{H_0 \rightarrow i|_s}(t)$ as the probability $l_{ij|ss'}$ can be reached from $i|_s$ at t given origin H_0 , formulated as

$$P_{H_0 \rightarrow i|_s}(t) = \begin{cases} 0 & t \in [t_{H_0}, t_{i|s}^-) \\ \int_{t_{i|s}^- - t_{H_0}}^{t - t_{H_0}} f_{ij|ss'}(\tau) d\tau & t \in [t_{i|s}^-, t_{j|s'}^+ - t_{ij|ss'}) \\ \int_{t_{i|s}^- - t_{H_0}}^{t_{j|s'}^+ - t_{ij|ss'} - t_{H_0}} f_{ij|ss'}(\tau) d\tau & t \in [t_{j|s'}^+ - t_{ij|ss'}, t_{H_1}] \end{cases} \quad (4)$$

Denote $P_{j|_{s'} \rightarrow H_1}(t)$ as the probability of reaching H_1 from $j|_{s'}$ based on available departure times, formulated as

$$P_{j|_{s'} \rightarrow H_1}(t) = \begin{cases} 0 & t \in [t_{j|_{s'}}^+, t_{H_1}] \\ \int_{t_{H_1} - t_{j|_{s'}}^+}^{t_{H_1} - t} f_{ij|ss'}(\tau) d\tau & t \in [t_{i|_s}^- + t_{ij|ss'}, t_{j|_{s'}}^+) \\ \int_{t_{H_1} - t_{j|_{s'}}^+}^{t_{H_1} - t_{i|_s}^- - t_{ij|ss'}} f_{ij|ss'}(\tau) d\tau & t \in [t_{H_0}, t_{i|_s}^- + t_{ij|ss'}) \end{cases} \quad (5)$$

The probability for visiting $l_{ij|ss'}$ at $t \in [t_{H_0}, t_{H_1}]$, denoted as $P(l_{ij|ss'}, t)$, is formulated as the joint probability of $l_{ij|ss'}$ being reached from $i|_s$ and arriving at H_1 within $(t_{H_1} - t)$. We normalize

probabilities among all accessible links at time t within the STP in SNK , for which all probabilities are added up to 1.

$$P(l_{ij|ss'}, t) = \frac{P_{H_0 \rightarrow i|s}(t) \times P_{j|s' \rightarrow H_1}(t)}{\sum_{l \in E} P(l, t)}, t_{i|s}^- + t_{ij|ss'} \leq t_{j|s'}^+ \quad (6)$$

Eq. (6) can be used to derive the visit probability of each link and the discrete distribution of the status space at a time moment in SNK .

Model estimation

We extract a transportation network and collect individuals' ATPs from GPS trajectories to estimate the holding time density functions for visit probability in the following steps.

Step 1: Calculate the extra available times for road segments and different types of activities. For travel links (road segments), the extra time of a sampled ATP is calculated as:

$$\tau = rec(t_{ATP}) - \min(t_{ATP}) \quad (7)$$

where $rec(t_{ATP})$ is the recorded ATP time from GPS data, $\min(t_{ATP})$ is the shortest ATP time with the same activity durations calculated using a path searching algorithm. Suppose $f_{ij|ss'}(\tau)$ of travel link $l_{ij|ss'}$ follows an exponential distribution, i.e.,

$$f_{ij|ss'}(\tau) = \begin{cases} \lambda e^{-\lambda\tau}, & \tau \geq 0 \\ 0, & \tau < 0 \end{cases} \quad (8)$$

$f^q(\tau|k)$ in Eq. (3) is substituted by Eq. (8) as $f_{ij|ss'}^q(\tau|\lambda_k)$, with parameter λ_k ($k = 1, 2, \dots, K$) to be estimated.

For transaction links of different types of activities, the extra activity duration of a sampled ATP is calculated as:

$$\tau_\alpha = rec(\tau_\alpha) - \min(\tau_\alpha) \quad (9)$$

where $rec(\tau_\alpha)$ is the recorded ATP's activity duration for conducting activity α , $\min(\tau_\alpha)$ is the minimum duration of α among all the extracted ATPs. We suppose $f_{ij|ss'}(\tau)$ (denoted as $f_\alpha(\tau)$) of conducting α follows the lognormal distribution, i.e.,

$$f_\alpha(\tau) = \frac{1}{\tau\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(\tau) - \mu)^2}{2\sigma^2}\right), \quad \tau > 0 \quad (10)$$

$f^q(\tau|k)$ in Eq. (3) is substituted by Eq. (10) as $f_\alpha^q(\tau|\mu_k, \sigma_k^2)$, with parameter μ_k, σ_k^2 ($k = 1, 2, \dots, K$) to be estimated.

Step 2: Estimate the parameters using maximum likelihood estimation (MLE). Given N the total number of extracted individuals' ATP, the likelihood for all individuals is:

$$L = \prod_{q=1}^N \left[\sum_{k=1}^K P_{qk} \cdot f^q(\tau|k) \right] = \prod_{q=1}^N \left[\sum_{k=1}^K \frac{\exp(\boldsymbol{\beta}_k \mathbf{x}_q)}{\sum_{k=1}^K \exp(\boldsymbol{\beta}_k \mathbf{x}_q)} \cdot f^q(\tau|\boldsymbol{\theta}_k) \right] \quad (11)$$

where θ_k are the parameters in Eq. (8) or Eq. (10) for travel and transaction links. The log-likelihood for the sampled ATPs is:

$$\ln L = \sum_{q=1}^N \left(\ln \sum_{k=1}^K P_{qk} \cdot f^q(\tau | \theta_k) \right) \quad (12)$$

We apply the gradient-descent method as a solution to solve the MLE for latent class models. The estimated parameters are used in simulating the visit probabilities in Eqs. (4-6).

3. RESULTS

The suggested model framework is implemented for an AP in a transportation network. STPs and visit probabilities are simulated for each activity state in *SNK* every 5 min in peak hours and 30 min in non-peak hours for travel links and activity locations.

Estimation results

In the numerical experiment, 2714 individuals' ATPs with one fixed workplace and one flexible activity (non-daily shopping) are extracted from GPS data. We select an AP in the Eindhoven area (the Netherlands) to simulate the STP construction and the visit probabilities within the STP (**Figure 1**). The settings are as follows:

- (1) The time budget is 742 min, 10% extra over the recorded ATP time. Given the recorded ATP departure time $t_{H_0} = 6:20$, the time window is [6:20, 18:42]. 816 flexible activity locations are selected as alternatives for non-daily shopping. The recorded minimum duration of working and flexible activity are 558 min and 36 min respectively.
- (2) The road network includes 47,901 nodes and 10,0581 directed road segments. The car speeds for non-peak hours [7:00, 9:00] and peak hours [16:30, 19:00] on 3 types of roads are <100, 80, 50> and <70, 50, 30> km/h.

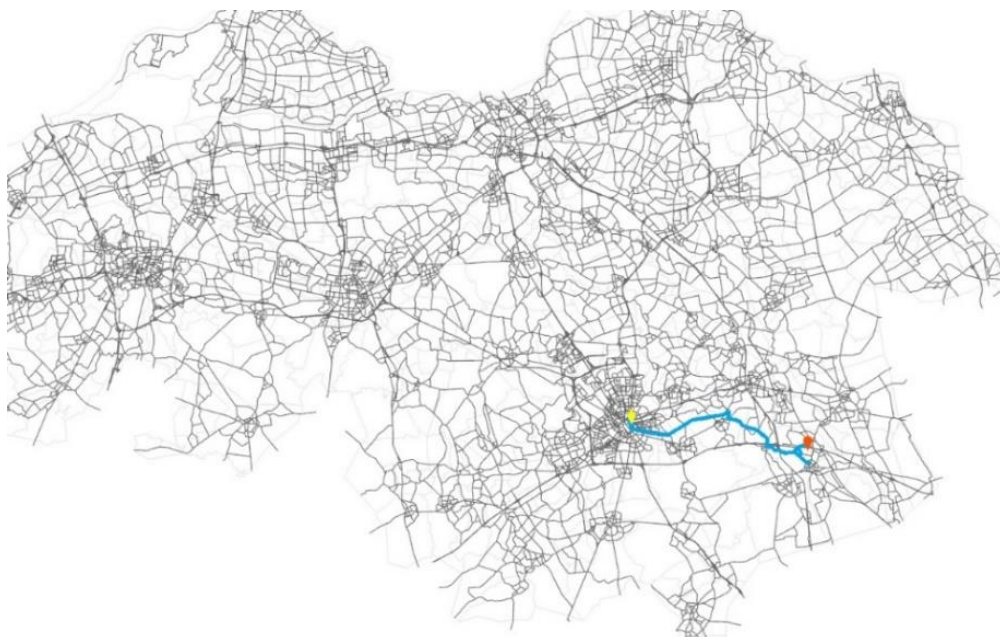


Figure 1. Selected ATP (anchors: red: home; yellow: workplace; blue line: GPS trajectory).

(3) The parameters of $f^q(\tau)$ for the traveling and non-daily shopping are estimated. Individual q 's gender and age are selected as the sociodemographic variables as x_{q1} and x_{q2} . Let $x_{q1} = 0$ if individual q is male, $x_{q1} = 1$ when q is female; $x_{q2} = 0$ if individual q 's age is between 20 to 59, $x_{q2} = 1$ when age ≥ 60 . The estimation outcomes of the latent class models are shown in Tables 1-6 (See Tables in the Appendix).

(4) For travel links, 4 activity states are simulated based on whether work and non-daily shopping are conducted in different sequences. For transaction links, 2 states are simulated, i.e, the flexible activity is conducted before and after work.

Illustration of visit probabilities

Travel links

The visit probabilities of travel links within the STPs reflect the heterogeneous interior for traveling with purpose, which is non-uniform and changing over activity states and time moments.

At activity state 0, the individual departs at 6:20 traveling to conduct the fixed and flexible activities. The number of travel links within STPs increases given available time during [6:20, 7:30], and decreases during [7:40, 8:30] since 558 min work has to be done (**Figure 2**). At activity state 2 where the 36-min non-daily shopping is conducted within [6:20, 7:00], the individual has more time for visiting more travel links at the same time point during [8:00, 8:55] at state 2 compared to the situation at state 0 (see in **Figure 3** (a) and (b) at 8:20).

There are no travel links within STP between (8:55, 16:00) since the individual is working. After 16:00 at activity state 1, the individual searches locations of non-daily shopping after work is conducted (**Figure 4**). At activity state 3 that all activities are done, around 30 to 40 min are available for the individual traveling to home during [16:40, 18:40], and thus more links can be accessed compared to the situation at state 1 (see **Figure 5** (c) and (d) at 18:00).

As shown in the results, the distribution of locations, activity sequences, available time, and the shortest path direction, together result in the changes of “higher visit probability travel links” (darker red).

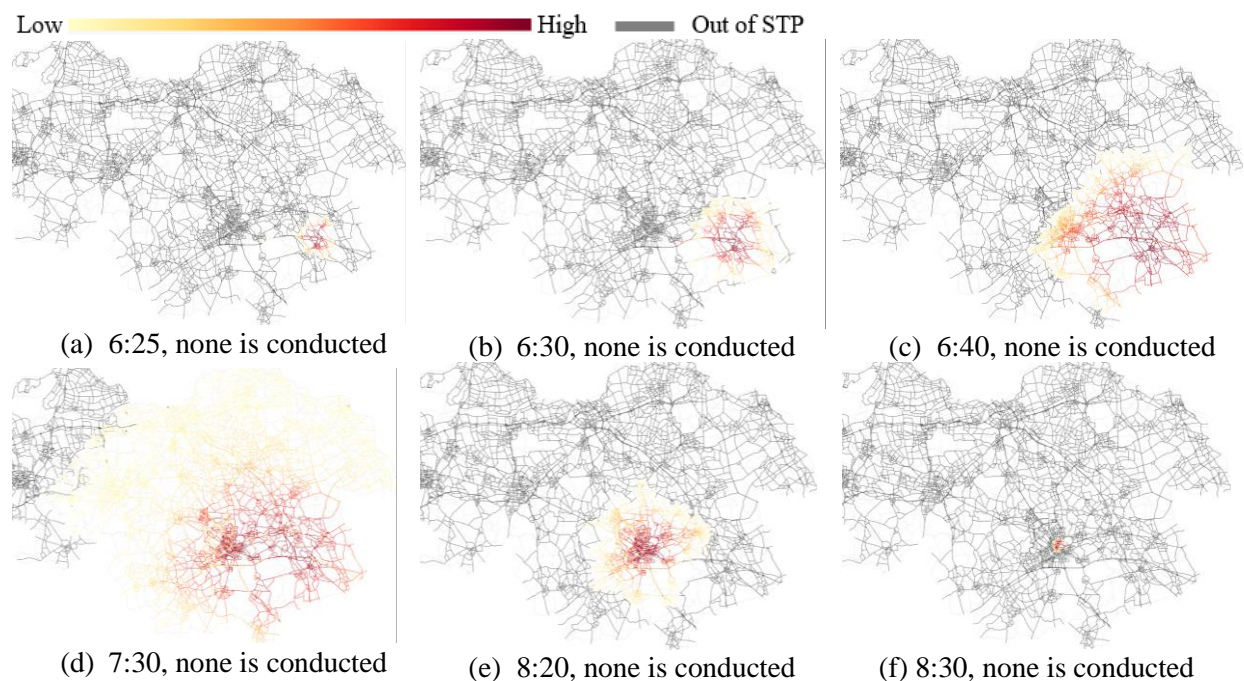


Figure 2. STPs interior of travel links at activity state 0.

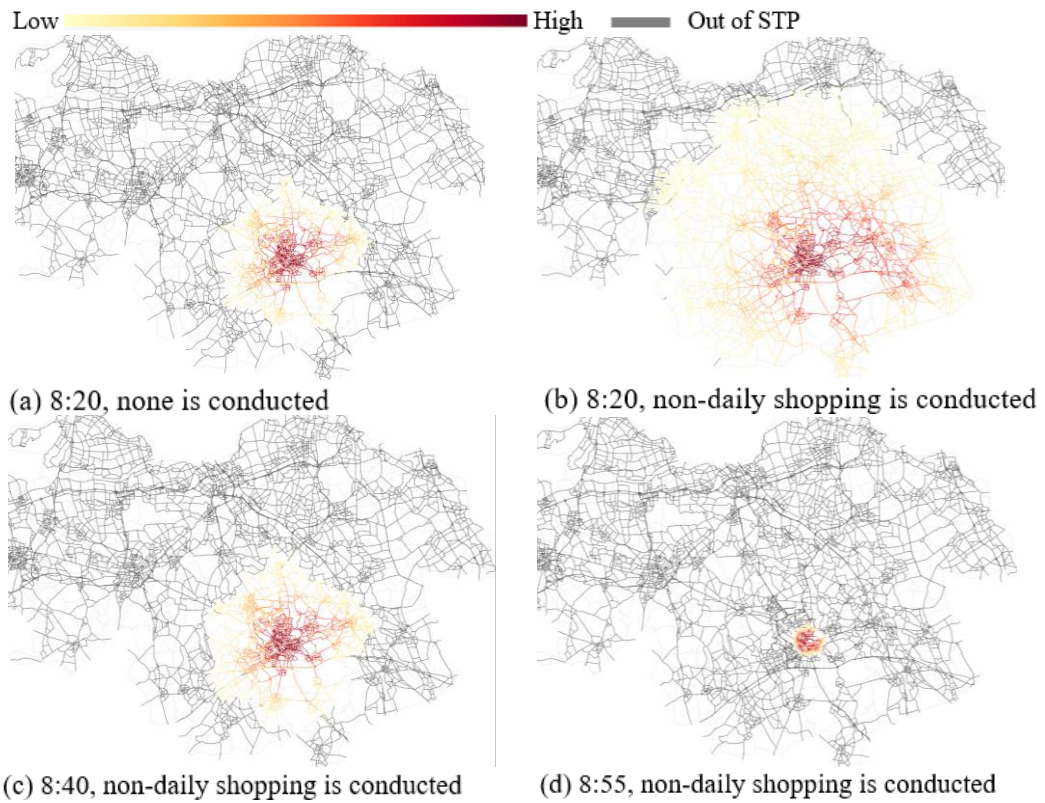


Figure 3. STPs interior of travel links at activity state 2.

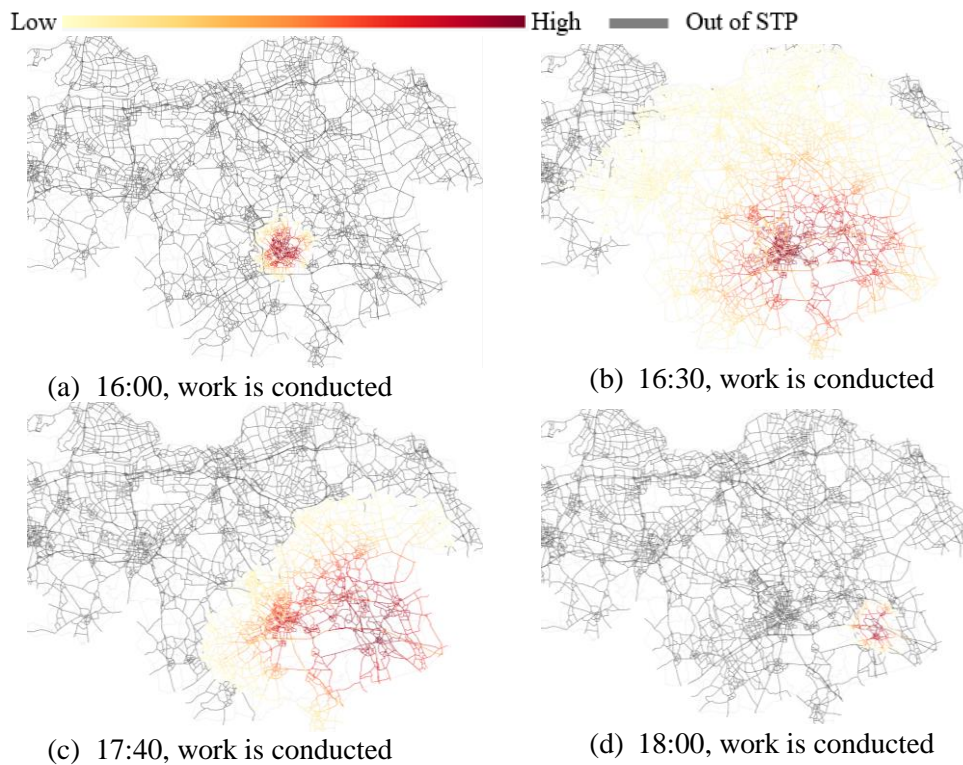


Figure 4. STPs interior of travel links at activity state 1.

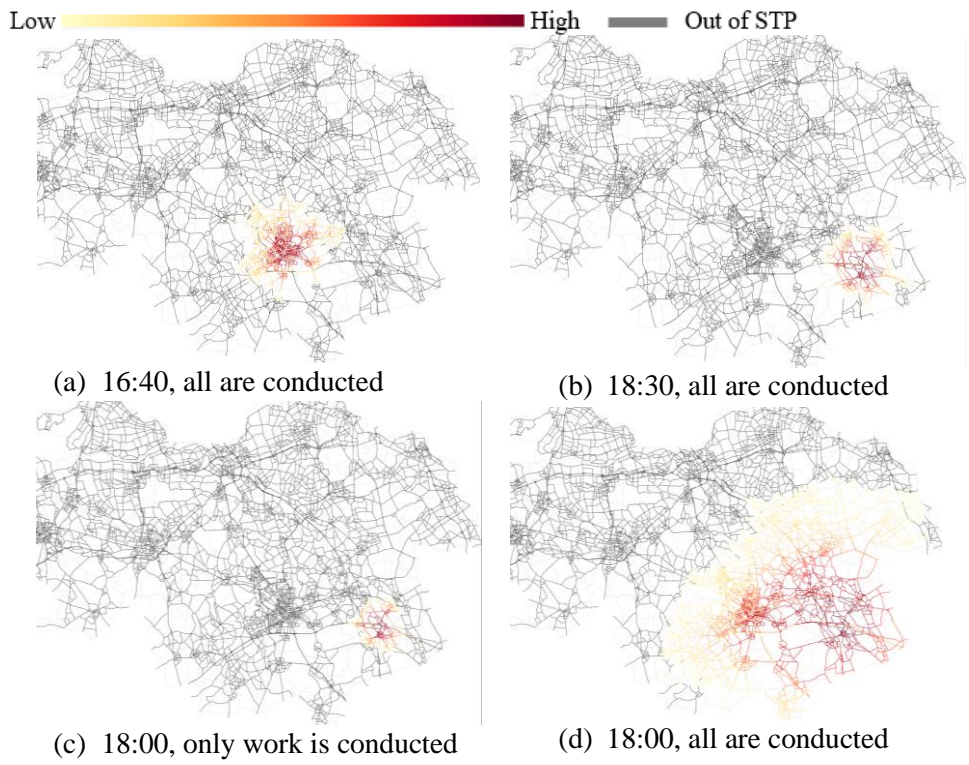


Figure 5. STPs interior of travel links at activity state 3.

Flexible activity locations

The level of accessibilities of the flexible activity locations within the STPs are reflected by the visit probabilities, which changes over activity sequence and time moments.

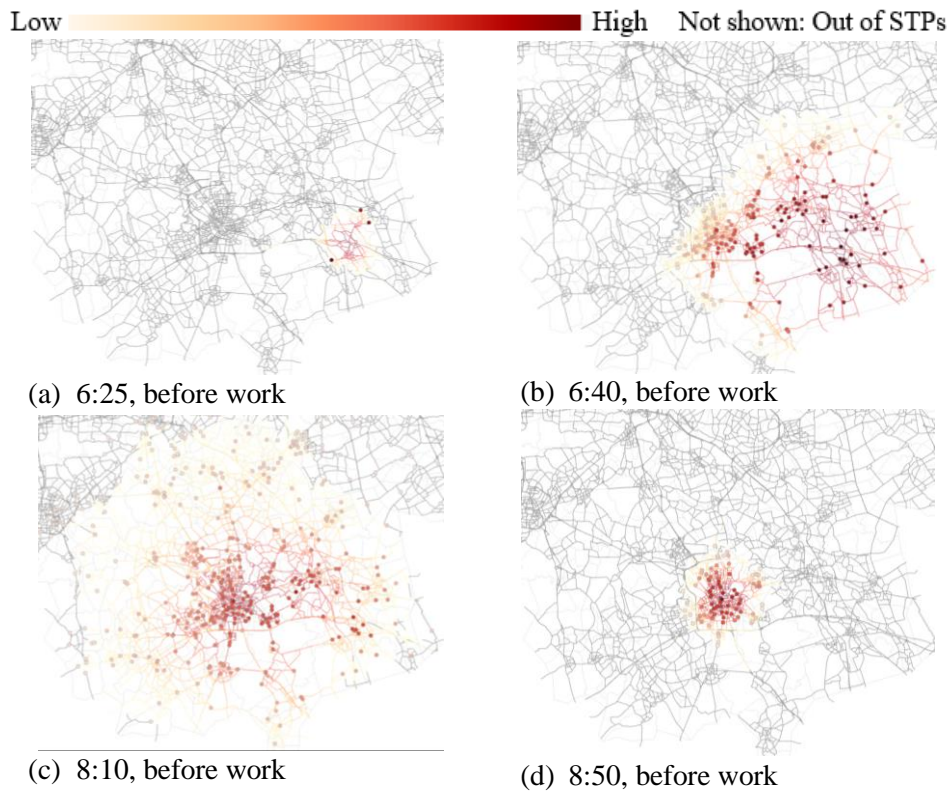


Figure 6. STPs interior of flexible activity locations before work.

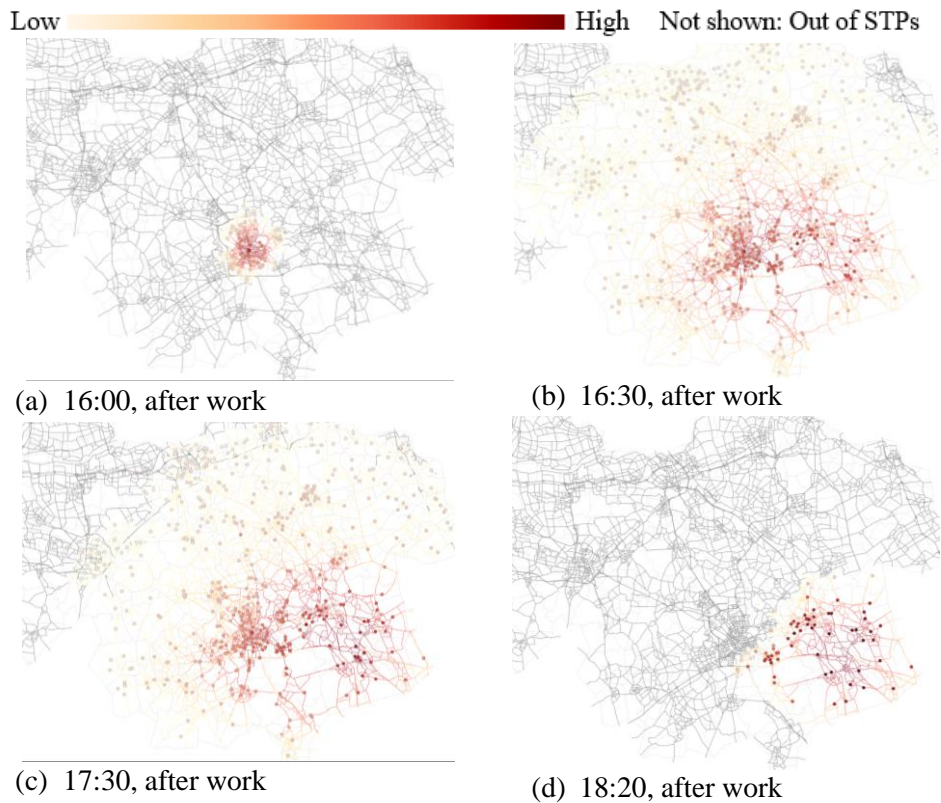


Figure 7. STPs interior of flexible activity locations after work.

The number of flexible activity locations within the STP increases during [6:20, 7:40] and the locations near home have relatively higher visit probabilities. After 8:00, the closer to the workplace, the higher visit probabilities of the locations are compared to the ones near home due to the available time (Figure 6). After work is conducted (Figure 7), the accessible locations first expand from the center of the workplace with higher probabilities and finally shrink to home at 18:20, given $t_{H_1} = 18:42$ and the 36-min flexible activity duration.

4. CONCLUSIONS

The state-dependent STPs of an AP delimit the potential mobility of an individual to access locations with limited time, but the deterministic characteristic of STP lacks the evaluation of how likely a location within the STP can be visited. This study proposes a model framework to construct and estimate the probabilistic STP of an AP based on the multi-state supernetwork representation. By assuming the distributions of holding time density functions of travel and transaction links in *SNK*, latent class models are further applied to capture and estimate the individuals' latent heterogeneity on extra time for traveling and conducting activities. The results illustrate that the visit probabilities of travel links and activity locations over different activity states and time points can describe the interior of the STPs. In the next step, we will develop proper accessibility measurements based on the quantifications of visit probabilities.

REFERENCES

Chen, X., Kwan, M.-P. 2012. Choice set formation with multiple flexible activities under space-time constraints. *International Journal of Geographical Information Science*, Vol. 26, No. 5, pp. 941-961.

- Hägerstrand, T. 1970. *What about people in regional science?* Paper presented at the Papers of the Regional Science Association.
- Howard, R. A. 1971. *Dynamic probabilistic systems, Volume II: Semi-Markov and Decision Processes* (Vol. 2). New York: Wiley.
- Kang, J. E., Chen, A. 2016. Constructing the feasible space-time region of the Household Activity Pattern Problem. *Transportmetrica A: Transport Science*, Vol. 12, No. 7, pp. 591-611.
- Liao, F. 2021. Exact space-time prism of an activity program: bidirectional searches in multi-state supernetwork. *International Journal of Geographical Information Science*, Vol. 35, No. 10, pp. 1975-2001.
- Liao, F., Arentze, T., Timmermans, H. 2010. Supernetwork approach for multimodal and multiactivity travel planning. *Transportation Research Record*, Vol. 2175, No. 1, pp. 38-46.
- Liao, F., Arentze, T., Timmermans, H. 2013. Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling. *Transportation Research Part B: Methodological*, Vol. 55, pp. 41-58.
- Miller, H. J. 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *International journal of geographical information systems*, Vol. 5, No. 3, pp. 287-301.
- Miller, H. J. 2017. Time Geography and Space-Time Prism. In *International Encyclopedia of Geography* (pp. 1-19).
- Song, Y., Miller, H. J. 2013. Simulating visit probability distributions within planar space-time prisms. *International Journal of Geographical Information Science*, Vol. 28, No. 1, pp. 104-125.
- Song, Y., Miller, H. J., Stempihar, J., Zhou, X. 2017. Green accessibility: Estimating the environmental costs of network-time prisms for sustainable transportation planning. *Journal of Transport Geography*, Vol. 64, pp. 109-119.
- Song, Y., Miller, H. J., Zhou, X., Proffitt, D. 2016. Modeling Visit Probabilities within Network-Time Prisms Using Markov Techniques. *Geographical Analysis*, Vol. 48, No. 1, pp. 18-42.
- Winter, S. 2009. Towards a probabilistic time geography. Paper presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on advances in geographic information systems.
- Winter, S., Yin, Z.-C. 2010a. Directed movements in probabilistic time geography. *International Journal of Geographical Information Science*, Vol. 24, No. 9, pp. 1349-1365.
- Winter, S., Yin, Z.-C. 2010b. The elements of probabilistic time geography. *GeoInformatica*, Vol. 15, No. 3, pp. 417-434.

APPENDIX

Tables 1-3 report the estimated parameters of $f^q(\tau)$ and the percentages that the sample belongs to each class with the highest membership probability for non-daily shopping, working, and traveling, respectively. To reflect the fit of the latent class models, Tables 4-6 shows the test results of Bayesian Information Criteria (BIC) and likelihood ratio (LR) given various numbers of classes, based on the assumptions of distributions and the hypotheses of the parameters.

The estimated parameters of latent class models:

- a. Flexible activity: non-daily shopping

Table 1. Parameters and highest P_{qk} percentages of selected class number

Class	β_k	μ_k, σ_k	Sample N_k / Highest P_{qk} %
1	const	0	μ_1 2.7109
	gender	0	σ_1 1.9683
	age	0	365 100%

b. Fixed activity: working

Table 2. Parameters and highest P_{qk} percentages of selected class number

Class	β_k		μ_k, σ_k		Sample N_k / Highest P_{qk} %	
1	const	0.536	μ_1	6.2672	1234	45.47%
	gender	-0.6813	σ_1	0.0941		
	age	-0.4178				
2	const	0	μ_2	5.2213	1480	54.53%
	gender	0	σ_2	0.9410		
	age	0				

c. Travel links

Table 3. Parameters and highest P_{qk} percentages of selected class number

Class	β_k		λ_k		Sample N_k / Highest P_{qk} %	
1	const	1.7586	λ_1	0.0271	380	24.58%
	gender	-2.9839				
	age	-2.7812				
2	const	-0.8908	λ_2	0.0346	842	54.46%
	gender	2.963				
	age	0.2311				
3	const	0	λ_3	0.0310	423	20.96%
	gender	0				
	age	0				

The results of the fit of the latent class models:

a. Flexible activity: non-daily shopping

Table 4. BIC and LR with latent classes number

Number of classes	Number of Parameters	Log-likelihood $\ln(L)$	BIC	LR (LR-val / p-val (df))
1	5	-1754.5719	3520.9438	0.5649 / 0.7539 (2)
2	10	-1697.8948	3454.7885	113.9192 / 8.6267e-20 (10)
3	15	-1681.2356	3450.9697	147.2378 / 8.5223e-24 (15)
4	20	-1675.6434	3469.2848	158.6689 / 1.3564e-23 (20)
5	25	-1665.4855	3478.4684	173.3506 / 3.6977e-24 (25)
10	50	-1650.1866	3595.3681	202.0330 / 3.6134e-20 (50)

$BIC = P \ln(N) - 2 \ln(L)$; P : number of β_k + number of μ_k, σ_k^2

LR test: The null hypothesis: $\mu_k = \frac{1}{N} \sum_{q=1}^N \ln \tau_q, \sigma_k^2 = \frac{1}{N} \sum_{q=1}^N \left(\ln \tau_q - \frac{1}{N} \sum_{q=1}^N \ln \tau_q \right)^2, \beta_k = 0$

b. Fixed activity: working

Table 5. BIC and LR with latent classes number

Number of classes	Number of Parameters	Log-likelihood $\ln(L)$	BIC	LR (LR-val / p-val (df))
1	5	-18973.5117	37962.8319	4046.5781 / 0 (2)
2	10	-17586.0957	35251.2531	6797.3008 / 0 (10)
3	15	-17469.0371	35056.6669	6919.1914 / 0 (15)
4	20	-17285.6230	34729.3696	7024.5898 / 0 (20)
5	25	-17259.5254	34716.7052	7470.7734 / 0 (25)
6	30	-17284.9824	34807.1502	7423.2188 / 0 (30)
8	40	-17287.2480	34890.7433	7415.6523 / 0 (40)
10	50	-17285.5410	34966.3909	7414.4844 / 0 (50)

c. Travel links

Table 6. BIC and LR with latent classes number

Number of classes	Number of Parameters	Log-likelihood $\ln(L)$	BIC	LR (LR-val / p-val (df))
1	3	-6882.1519	13771.6471	0. / 1.0 (1)
2	6	-6877.2329	13813.2132	9.7520 / 0.2828 (6)
3	9	-6875.3398	13838.8008	13.7549 / 0.3166 (9)
4	12	-6875.2188	13867.9323	13.6055 / 0.6281 (12)
5	15	-6875.0942	13897.0570	14.3369 / 0.8130 (15)
8	24	-6874.9126	13984.8148	14.4785 / 0.9966 (24)
10	30	-6874.9165	14043.5701	14.6758 / 0.9999 (30)

P : number of β_k + number of λ_k

LR test: The null hypothesis: $\lambda_k = \frac{N}{\sum_{q=1}^N \tau_q}, \beta_k = 0$

Analysing the Effects of Adding Shared Electric Bicycles as a New Mode on the Modal Split of Multimodal Trips between Delft and Rotterdam Using an Unlabelled Multimodal Supernetwork

Gijsbert Koen de Clercq*¹, Maaïke Snelder², Arjan van Binsbergen³, and Bart van Arem⁴

¹ PhD Researcher, Department of Transport and Planning, Delft University of Technology, Netherlands

^{2,3} Associate Professor, Department of Transport and Planning, Delft University of Technology, Netherlands

⁴ Professor, Department of Transport and Planning, Delft University of Technology, Netherlands

SHORT SUMMARY

Assessing to what extent new modes will change modal split is difficult, since revealed preference data is not available yet to estimate models. To address this, an unlabelled multimodal supernetwork is developed in which mode and route choice are simultaneously modelled. The model has been estimated based on data of existing modes and can be used to assess the impact of any new mode. We applied the model to analyse the effects of shared e-bicycles on one Origin-Destination pair between Delft and Rotterdam. The main scientific contribution of this paper is that it successfully demonstrates how an unlabelled multimodal supernetwork can be used to analyse the effects of shared e-bicycles on the modal split between Delft and Rotterdam. The results show that the modal share of shared e-bicycles is 35.3-40.5% for unimodal trips and occur in 36.2-46.3% of multimodal trips, indicating that shared e-bicycles can significantly change the modal split.

Keywords: agent-based modelling; multimodal; shared e-bicycles; supernetwork; unlabelled mode choice

Word count: 2496

1. INTRODUCTION

Several mobility systems, ranging from shared electric bicycles to autonomous vehicles, have been developed. These new mobility systems could change the way our societies function in terms of sustainability, equity, accessibility, and safety (Fagnant & Kockelman, 2015; Milakis et al., 2017; Shaheen et al., 2019).

Researching how new modes affect mode choice is difficult since revealed data of the potential users using these new mobility systems are not available yet, so mode specific parameters cannot be estimated. In our previous research, this challenge has been addressed by developing an abstract, or unlabelled, mode modelling approach to assess the modal share of any new mode and unimodal trips (De Clercq et al., 2022). The unlabelled mode modelling approach was first introduced by Quandt & Baumal and describes a method to formulate a discrete choice model by describing the utility of each mode with the same mode attributes for each mode and by leaving out mode-specific constants and parameters (Quandt & Baumal, 1966). In our previous research (De Clercq et al., 2022), it was shown that any new mode can be modelled using the unlabelled mode modelling approach as long as the new mode can be described as a (new) combination of existing attributes of which the relative importance can be estimated based on revealed preference

data. In those cases, the utility function of the new mode can be defined and thus the new mode can be added as an option in the choice set of a discrete choice model. However, a shortfall of this approach is that it does not yet cover multimodal trips.

New modes, such in our case shared electric bicycles, will be often used for the the first- and/or last-mile parts of a trip and will only available at certain locations (e.g., mobility hubs) (Van Eck et al., 2014). Therefore, to analyse new transport modes, it will be often necessary to analyse their use in a multimodal setting. This can be done by developing a multimodal supernetwork which can model multimodal trips without the need to predefine the combinations of mobility systems and where mode and route choice happens simultaneously (Liao, 2016; Van Eck et al., 2014; Vo et al., 2021). These models only need to know where people are allowed to switch modes (e.g., where the mobility hubs are located).

In our study (De Clercq et al., 2023), we developed such a supernetwork approach to assess the impact of new mobility systems. Following the same logic as with the previously mentioned unlabelled discrete choice model (De Clercq et al., 2022), in our supernetwork, one can define the new mobility system by describing it in terms of a broad set of mode attributes but without any mode-specific constants, and add that mode as a new ‘layer’ in the supernetwork. So far, the model has only been applied to theoretical simple networks and a fictive multimodal version of the Sioux Falls network to demonstrate and verify the method. In this paper, we applied the model in a more realistic case situation. We focus on the introduction of shared electric bicycles on one example Origin-Destination (OD)-pair between Delft and Rotterdam because cycling is a dominant mode of transport in the Netherlands and electric bicycles are becoming increasingly popular (Sun et al., 2020).

This paper contributes to the existing literature by demonstrating how an agent-based multimodal supernetwork-based traffic assignment model with unlabelled modes that takes into account mode and route choice simultaneously can be applied in a real use-case to gain insight into the influence of shared electric bicycles on a specific OD-pair between Delft and Rotterdam, including the use in the first and last-mile parts of a trip. Knowledge gaps and possible future research directions to research the effects of new modes on the modal split of urban areas are identified and discussed as well.

2. METHODOLOGY

Figure 1 summarizes the supernetwork approach that has been applied in this paper. Trip generation and distribution are considered exogenous. Mode and route choice are calculated simultaneously based on the resistance per edge. Subsequently, per timestep, trips are assigned to the supernetwork using a mesoscopic dynamic traffic assignment approach. The assignment model with the supernetwork is set up using Python 3.10.2 and the NetworkX package (Hagberg et al., 2008). Below, the main elements of the assignment model are described. For an extensive description of the network definition, the combined mode and route choice model and the network loading model, refer to (De Clercq et al., 2023).

The supernetwork consists of one layer for each mode. Edges represent aggregated road segments with generalized resistance (i.e., disutility), transit segments (representing aggregated transit lines) and dummy transition edges from and to a neutral layer (representing transfer resistance). The transport mode edges have a length and a number of attributes (representing the mode and link attributes), where for modes using the road network the ‘time’ attribute can change with the use of the edge (i.e., the flow).

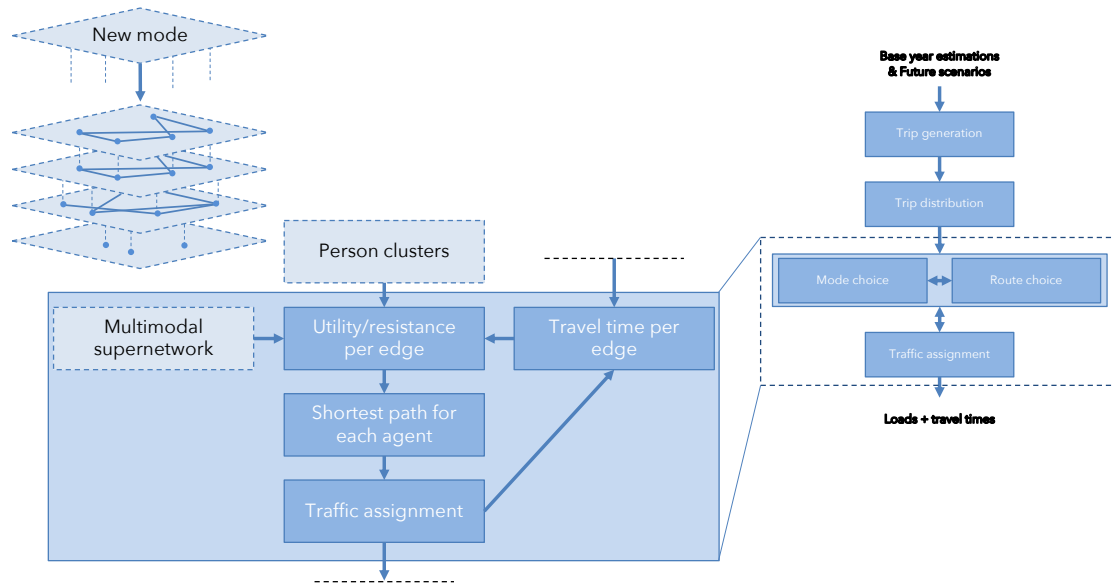


Figure 1: General layout of multimodal supernetwork (De Clercq et al., 2023)

Twelve mode attribute assumptions (times mode attribute parameters, see Table 3) have been used to define the resistance in the edges for all modes: Initial cost (€); Cost/trip (€); Time (min); Driving task (-); Skills (-) (i.e., driver’s license); Weather protection (-); Luggage (-); Shared (-); Availability (-); Reservation (-); Active (-); Accessible (-). The mode attribute assumptions per mode are shown in Table 1.

Table 1: Mode attribute assumptions

Mode attribute	Addable*	Source and determination
Initial cost (€)	Yes	Car, transit, cycle, walk = 0
Cost/trip (€)	Yes	Car = €0.19 per km; transit = 0.20 per km; walk = 0; bicycle = €100 purchase costs, with 4 trips per day for 5 years
Time (min)	Yes	Car, transit, bicycle and walk from Google Maps (<i>Travel times Google Maps</i> , n.d.)
Driving task (-)	No	Car, bicycle = 1; transit, walk = 0
Skills (-) (i.e., driver’s license)	No	Car = 1; transit, bicycle, walk = 0
Weather protection (-)	No	Car, transit = 1; bicycle, walk = 0
Luggage (-)	No	Car = 1; transit = 0.5; bicycle, walk = 0
Shared (-)	No	Car, bicycle, walk = 0; transit = 1
Availability (-)	No	Car = 1; transit = 0.5; bicycle = 1, walk = 1
Reservation (-)	No	Car, bicycle, walk = 1; transit = 0
Active (-)	No	Car, transit = 0; bicycle, walk = 1
Accessible (-)	No	Transit = 1; Car, bicycle, walk = 0

*Addable and non-addable attributes are implemented differently in the edge/route resistance calculations (see Eq. 1 and Eq. 2).

When agents want to switch modes, the resistance in the considered route also contains the time it takes to disembark from the mode, and embark to the mode (see Table 2) and the ‘initial cost’ of a mode (see Table 1). The times in these edges are multiplied by 3 to represent the extra mental effort it takes for users to switch modes and wait for the next mode (Wardman, 2004).

Table 2: Assumed time to switch modes, multiplied by 3 to model extra mental effort for switching mode, based on (Wardman, 2004)

Mode	Neutral to mode	Mode to neutral
Car	5 min (get in car)	2 min (parking)
Transit (BTM)	7.5 min (average waiting time, ass. freq.: 15 min)	5 min
Bicycle	1 min (get on bike)	1 min (parking)
Walk	0 min	0 min

The total resistance per route is defined as the sum of the edge resistances. Edge resistances are the sum of a series of products of mode attributes and the valuations of these attributes by users. For the modelling approach, homogenous groups of users are combined in clusters. There are two categories of mode attributes; one category which is addable over the route regardless of the length of each edge and the length of the route (e.g., travel time, costs) (see Eq. 1) and one category which is not addable over the route. The attributes within this last category need to be weighted with the length of each edge within the route to come to a weighted average value for those attributes (e.g., weather protection on 70% of the total length of the route) (see Eq. 2). These two edge resistance calculations are combined in Eq. 3, where the resistances of all edges in one route are summed up and divided by the length of that route to get the ratio of the mode attributes of each edge respective to their share in the total route. A multinomial logit model (MNL) (Ortuzar & Willumsen, 2011) with en-route route choice is used to determine the mode/route choice based on the route resistance. IIA (Independence of Irrelevant Alternatives) is assumed for all modes in this study.

$$E_{addable,i,t,n} = \sum_1^k \beta_n \chi_{k,i,t}, \quad \forall k \in K \quad \text{Eq. 1}$$

$$E_{non-addable,i,t,n} = len_i * \sum_1^m \beta_n \chi_{m,i,t}, \quad \forall m \in M \quad \text{Eq. 2}$$

$$U_{r,t,n} = \sum_{i \in I} E_{addable,i,t,n} + \frac{\sum_{i \in I} E_{non-addable,i,t,n}}{L} \quad \text{Eq. 3}$$

where;

- U = route resistance [-];
- $E_{addable}$ = edge resistance addable component [-];
- $E_{non-addable}$ = edge resistance non-addable component [-];
- β = cluster valuation of mode attribute [-];
- χ = value of mode attribute [-];
- len = length of edge [km];
- L = length of route [km];

- n = cluster index [-];
- i = edge index of edges within route [-];
- k = addable mode attribute index [-];
- m = non-addable mode attribute index [-];
- t = time index [-];
- r = route index [-];

- I = set of edges within mode layers in route r ;
- K = set of addable mode attributes (e.g., travel time); and
- M = set of non-addable mode attributes (e.g., weather protection).

The valuation (i.e., observed preferences for travellers for attributes) of the attributes have been estimated using a large-scale trip survey in the Netherlands (CBS, 2017) for the year 2017. The

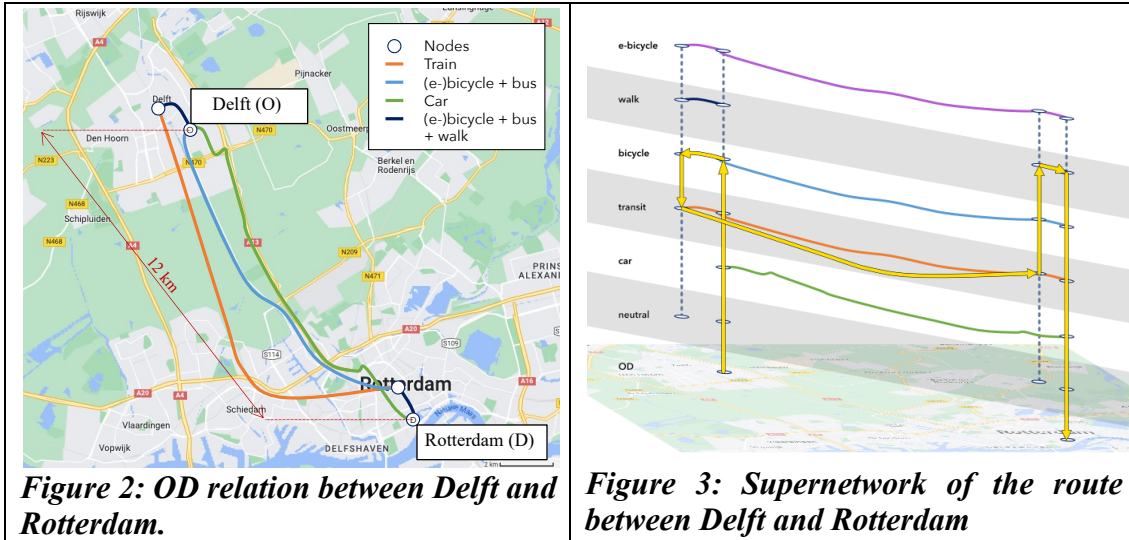
dataset contains personal, trip, and mode information. It contains five main transport modes: car, carpool, transit (BTM), bicycle, and walk. This dataset is further enriched with the mode attributes (see Table 1). Further more, clusters of travelers with similar characteristics have been identified, using k-means clustering and the elbow method to define the optimal number of clusters for the trips. The estimated valuations for the six clusters are shown in Table 3. Note that some parameters have different signs per cluster. This can be explained by considering that some clusters value certain traits positively (e.g., higher costs are a status symbol) and other clusters value certain traits negatively (e.g., higher costs make a mode less affordable). Because these attributes capture almost all aspects that determine mode and route choice, mode specific constants are no longer needed. The valuations of the attributes (i.e., betas) are also no longer mode-specific and therewith transferable to new modes.

Table 3: Mode attribute valuations per cluster (De Clercq et al., 2022)

<i>Cluster</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>
(Initial) Cost	0.0193	-0.0267	-0.0343	0.0229	-0.0123	-0.0806
Time	-0.0208	-0.0439	-0.0207	-0.0243	-0.036	-0.0218
Driving task	-0.571	-0.0884	-0.855	-1.21	-0.129	-1.2
Skills	-0.16	2.17	1.44	2.22	-2.81	1.52
Weather protection	-1.07	-0.402	-0.284	-0.781	-1.22	-0.0638
Luggage	-1.08	-1.39	-0.0653	-0.719	-0.952	-0.248
Shared	-0.611	-2.01	-1.25	-3.8	1.3	-1.43
Availability	-1.87	-3.96	-1.44	-8.28	-0.24	-2.46
Reservation	-0.733	-1.74	-0.292	0.313	1.79	0.672
Active	0.74	-0.00596	0.314	0.58	1.22	0.314
Accessible	-0.671	-2.58	-1.25	-3.49	1.41	-1.94

Case-study: Delft to Rotterdam

Figure 2 visualizes the network with first-mile, main, and last-mile edges for an example OD-relation between Delft and Rotterdam. Nodes depict centroids and mobility hubs (the network definition does not contain centroid feeders). First- and last-mile edges have multiple modes. The origin and the destinations are indicated using O and D respectively. Figure 3 gives a supernetwork representation of this network with five modes for which the model has been estimated (the reference case) and an added sixth mode shared e-bicycles (for the ‘new mode’ alternative). The network contains transition links between modes through a neutral layer. Note that when switching mode, an agent always needs to leave the mode, enter the neutral layer, and enter the new mode. The possible transfers on certain locations (circles) between modes are depicted with the vertical dashed lines. One example multimodal route is visualized in yellow, where the bicycle is used to cycle from the centroid (O, Delft) to the station, transit is used as main mode and walking is used for the last-mile to the destination (D, Rotterdam).



The implementation of shared electric bicycles can come with different variants for initial costs, speed, generalized transit times from the neutral layer to shared e-bicycles and vice versa. These variants are described in Table 4. The possible combinations (81) of these attribute values form scenarios. These scenarios are simulated to analyse the effects of a different level of service of shared electric bikes on the modal split and travel times.

Table 4: Attribute values

Nr.	Attribute	Pricing policy
1	Initial cost shared electric step [€]	3.2 – 4.0 – 4.8
2	Speed [km/h]	20 – 25 - 30
3	Neutral to e-bike [min]	2.4 – 3.0 – 3.6
4	E-bike to neutral [min]	1.6 – 2.0 – 2.4
5	Network [-]	With/without e-bicycles

Simulation

For the simulation, the earlier identified six clusters of travelers are used; 166 trips for each cluster are simulated. A timestep of 1/100 hours (=36 sec) is used. This time step is chosen such that an agent will spend at least two timesteps on the shortest link (1.8 km) considering the highest free flow speed in this model (90 km/h). The simulated time period in the model is 4 hours representing a morning peak hour from 6 am to 10 am. The same seed number is used in all simulations to be able to compare different scenarios. The number of transfers between modes in the multimodal networks is limited to 2 (first-mile → main → and last-mile). Since transit is modelled in one layer and contains all bus, tram, metro (BTM) transit, this is assumed to be realistic.

3. RESULTS AND DISCUSSION

The results of the simulations, including sensitivity analysis results between brackets, are shown in Table 5-7. It can be observed that the modal share of shared e-bicycles ranges between 35.3 and 40.5% in unimodal trips and occurs in 36.2-46.3% of the multimodal trips, indicating that shared e-bicycles can significantly change the modal split between Delft and Rotterdam, reducing mainly the modal share of cars, cycling, and, to a lesser extent, transit and walking. When looking at the total distance travelled per mode for all scenarios, it is interesting to point out that walking is used at least in 41.7% of the multimodal trips but amounts to only a maximum of 0.6% of all distance travelled. This indicates that walking occurs for first- and last-mile as would be expected. The average travel time changes depending on the configuration of shared e-bicycles, indicating

that travel time can be improved by introducing shared e-bicycles but depend on the level of service that shared e-bicycles have when they are introduced on the roads.

Table 5: Modal split (trips) with and without shared e-bicycle, including sensitivity analysis results between brackets

	Modal split [%]					
	Car	Transit	Bicycle	Walk	E-bicycle	Multimodal
No e-bicycles	45,4	0,8	45,8	0,0	-	8,0
E-bicycles	28,2 [25,5 - 30,4]	0,5 [0,2 - 0,5]	25,3 [23,4 - 25,9]	0,0 [0,0 - 0,0]	40,5 [35,3 - 40,5]	5,5 [5,1 - 5,8]
	Modal split of multimodal trips [%]*					
	Car	Transit	Bicycle	Walk	E-bicycle	
No e-bicycles	0,0	60,6	89,0	60,2	-	
E-bicycles	0,0 [0,0 - 0,0]	37,6 [28,6 - 37,6]	75,8 [67,9 - 75,8]	51,7 [41,7 - 51,7]	44,0 [36,2 - 46,3]	

*The modal split of mixed trips amounts to more than 100% since multiple modes can occur in one single trip. The numbers can be interpreted as the percentage of the mixed trips that contain a certain mode.

Table 6: Modal split (distance) with and without shared e-bicycles, including sensitivity results between brackets

	Modal split [% of distance]					
	Car	Transit	Bicycle	Walk	E-bicycle	Multimodal
No e-bicycles	43,6	0,9	46,6	0,0	-	8,8
E-bicycles	26,9 [24,2 - 28,9]	0,6 [0,3 - 0,7]	25,6 [23,6 - 29,6]	0,0 [0,0 - 0,0]	41,0 [35,7 - 45,2]	6,0 [5,5 - 7,7]
	Modal split of multimodal trips [% of distance]					
	Car	Transit	Bicycle	Walk	E-bicycle	
No e-bicycles	0,0	3,1	5,1	0,6	-	
E-bicycles	0,0 [0,0 - 0,0]	1,3 [0,7 - 1,6]	2,5 [2,1 - 3,8]	0,4 [0,3 - 0,5]	1,8 [1,5 - 2,4]	

Table 7: Average speed, distance, and duration with and without shared e-bicycles, including sensitivity analysis results between brackets

	Average speed [km/hr]	Average distance [km]	Average duration [min]
No e-bicycles	23,64	15,40	43
E-bicycles	24,01 [22,09 - 25,99]	15,56 [15,52 - 15,61]	41 [38 - 45]

The results in this study are plausible in the sense that the e-bicycle mainly replaces car and bike trips. This is in line with the findings in Sun et al. (2020). They used a longitudinal dataset from the Netherlands Mobility Panel to analyze the modal shift effects of people who bought an e-bicycle. However, the potential modal share of shared e-bicycles is slightly higher than the study of Sun et al (2020) shows. For trips of about 15 kilometres, they found a modal share of 33-36%. This might be explained by the relatively high share of normal bicycles in the reference situation for this specific origin-destination pair and the fact that a multinomial logit model is used to determine the next node for each agent, which does not account for overlap ('red/blue-bus paradox').

4. CONCLUSIONS AND RECOMMENDATIONS

This study successfully demonstrated how an agent-based unlabelled multimodal supernetwork-based traffic assignment model can be used to assess the effects of new modes such as shared electric bicycles on the modal split for an example origin-destination pair Delft- Rotterdam. This is done by using a supernetwork framework, where each available mode is modelled as a specific layer within this supernetwork with nodes and edges, where the edges' resistances are described by a set of attributes without an alternative-specific constant. The mode-specific layers are interconnected to a neutral layer with edges representing transfer resistances, also described with a set of attributes. The unlabelled approach, i.e., without mode specific parameters and constants, makes it possible to add any new mode.

The results show plausible modal-shift effects from cars and bicycles to shared electric bicycles. However, the absolute modal share seems slightly overestimated because overlap is not considered. To account for overlap in routes (and similar modes), it is recommended to explore grouping layers of similar modes and overlapping routes into one 'nest' by using a path-overlap factor. A path size correction logit model (PSCL) in combination with a multiplicative MNL is expected to work well on real networks (Bovy et al., 2008; Smits et al., 2018). PSCL models exist, for transit only, based on the shared number of transfer nodes, edges and travel times, which have higher accuracy for transit, but these models cannot be applied on car transport (Dixit et al., 2021). Since both transit and car modes are used, this method is not trivial to implement on a multimodal supernetwork.

The modular approach of this supernetwork allows for further explorations of other new modes, more scenarios, and other network configurations. Recommended options to explore further are to include other new modes, pricing policies, adding a time-based transit schedule, and changing the availabilities and stops of transit or other modes. Further, no disruptions (e.g., weather conditions, accidents) are analyzed in this study yet. Disruptions could be implemented by temporarily increasing the edge resistance in some places. This could be an approach to give insight into how a new mode changes the travel time reliability of a system. Finally, it is recommended to integrate this multimodal supernetwork in a land-use-transport interaction model to see how higher-order aspects, such as activity patterns, change over the years as a result of the introduction of new transport modes.

ACKNOWLEDGEMENTS

This research is part of the project SUMMALab which is supported by the research program Sustainable Living Labs from the Dutch Research Council (NWO) under Grant 439.18.460 B. The authors declare that there is no conflict of interest.

REFERENCES

- Bovy, P. H. L., Bekhor, S., & Prato, C. G. (2008). The Factor of Revisited Path Size: Alternative Derivation. *https://doi.org/10.3141/2076-15*, 2076, 132–140.
<https://doi.org/10.3141/2076-15>
- CBS. (2017). *Onderzoek Verplaatsingen in Nederland 2017*. July, 39.
<https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:61643>
- De Clercq, G. K., Van Binsbergen, A., Van Arem, B., & Snelder, M. (2022). Estimating the Potential Modal Split of Any Future Mode Using Revealed Preference Data. *Journal of Advanced Transportation*, 2022. <https://doi.org/10.1155/2022/6816851>
- De Clercq, G. K., Van Binsbergen, A., Van Arem, B., & Snelder, M. (2023). Estimating the Effects of Any Future Mode on the Travel Times of an Urban Area Using a Multimodal

- Supernetwork. *Preprint*. <https://doi.org/10.13140/RG.2.2.25907.91689>
- Dixit, M., Cats, O., Brands, T., Van Oort, N., & Hoogendoorn, S. (2021). Perception of overlap in multi-modal urban transit route choice Perception of overlap in multi-modal urban transit route choice Malvika Dixit, Oded Cats, Ties Brands, Niels van Oort & Serge Hoogendoorn Perception of overlap in multi-modal urban transit ro. *Citation*. <https://doi.org/10.1080/23249935.2021.2005180>
- Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77, 167–181. <https://doi.org/10.1016/j.tra.2015.04.003>
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference* (pp. 11–15). http://conference.scipy.org/proceedings/SciPy2008/paper_2/
- Liao, F. (2016). Modeling duration choice in space–time multi-state supernetworks for individual activity-travel scheduling. *Transportation Research Part C: Emerging Technologies*, 69, 16–35. <https://doi.org/10.1016/J.TRC.2016.05.011>
- Milakis, D., Van Arem, B., & Van Wee, B. (2017). Policy and society related implications of automated driving: A review of literature and directions for future research. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 21(4), 324–348. <https://doi.org/10.1080/15472450.2017.1291351>
- Ortuzar, D., & Willumsen, L. (2011). *Modelling Transport*.
- Quandt, R. E., & Baumal, W. J. (1966). The Abstract Mode Model: Theory and Measurement. *Northeast Corridor Transportation Project, Technical Paper No. 4*.
- Shaheen, S., Cohen, A., Chan, N., & Bansal, A. (2019). Sharing strategies: Carsharing, shared micromobility (bikesharing and scooter sharing), transportation network companies, microtransit, and other innovative mobility modes. In *Transportation, Land Use, and Environmental Planning* (pp. 237–262). Elsevier. <https://doi.org/10.1016/B978-0-12-815167-9.00013-X>
- Smits, E.-S., Pel, A. J., Bliemer, M. C. J., & van Arem, B. (2018). *Generalized Multivariate Extreme Value Models for Explicit Route Choice Sets*. 1–42. <http://arxiv.org/abs/1808.04280>
- Sun, Q., Feng, T., Kemperman, A., & Spahn, A. (2020). Modal shift implications of e-bike use in the Netherlands: Moving towards sustainability? *Transportation Research Part D: Transport and Environment*, 78, 102202. <https://doi.org/10.1016/J.TRD.2019.102202>
- Travel times Google Maps*. (n.d.). Retrieved February 2, 2023, from <https://www.google.com/maps>
- Van Eck, G., Brands, T., Wismans, L. J. J., Pel, A. J., & Van Nes, R. (2014). Model complexities and requirements for multimodal transport network design: Assessment of classical, state-of-the-practice, and state-of-the-research models. *Transportation Research Record*, 2429, 178–187. <https://doi.org/10.3141/2429-19>
- Vo, K. D., Lam, W. H. K., & Li, Z. C. (2021). A mixed-equilibrium model of individual and household activity–travel choices in multimodal transportation networks. *Transportation Research Part C: Emerging Technologies*, 131(August), 103337. <https://doi.org/10.1016/j.trc.2021.103337>
- Wardman, M. (2004). Public transport values of time. *Transport Policy*, 11(4), 363–377. <https://doi.org/10.1016/j.tranpol.2004.05.001>

MobilityCoins - an integrated multimodal Wardropian model for policy analysis

Allister Loder^{1,*} and Klaus Bogenberger¹

¹Chair of Traffic Engineering and Control, Technical University of Munich

SHORT SUMMARY

MobilityCoins are a tradable mobility credit (TMC) scheme, where all modes can have link-specific and origin-and-destination-specific charges and incentives. These schemes are alternatives to congestion pricing and fuel excise taxes. Their design as a cap-and-trade scheme means that a fixed market volume is defined based on a to-be-regulated quantity. MobilityCoins are distributed to all travelers, who use them to pay for mobility or sell them on a market. However, the question of how to select policy parameters of such schemes in real-world contexts remains unanswered. In this paper, we develop a multimodal Wardropian transport model with integrated MobilityCoins scheme for transport policy analysis. Travelers have the choice between cars, public transport, and bicycles, where only cars experiences congestion effects. Using a simple model, we illustrate how a MobilityCoins scheme impacts transport outcomes under different system designs, e.g., declining overall market volume of MobilityCoins.

Keywords: Transport policy; transportation network modeling; tradable mobility credits; road user charges

1 INTRODUCTION

It has been argued that “economists have had limited success in promoting economically efficient transportation and environmental externality policies” (Lindsey & Santos, 2020). The state-of-the-art policies, if one may called it, are fuel excise taxes. However, when considering the advent of electric vehicles that are not paying any fuel excise taxes at all, one realizes that not only the tax revenue will decline with the obvious consequences for the transportation system funding, but also does the ability to use this policy to manage demand and congestion vanishes. Thus, new policies and mixes of them are required for “for deep CO2 mitigation in road transport” (Axsen et al., 2020).

In economics, a long discussion on “price vs. quantities” exists for the regulation of an economic system, i.e., setting standards or limits or charging taxes (Weitzman, 1974). Here, Dales was one of the first proposing such a quantitative instrument to manage external costs using a cap-and-trade scheme (Dales, 1968). In transport, such policy instrument based on tradable mobility credits (TMC) has been put forward by Verhoef et al. to regulate externalities (Verhoef et al., 1997), but so far did not see any real-world implementation. Nevertheless, such a policy instrument did see already see implementation in energy in order to, e.g., manage carbon emissions (Perroni & Rutherford, 1993) and to promote green energy deployment (Frei et al., 2018). The general idea of taxation is to impose a tax on the market price which in case of elastic demand reduces demand. Contrary, in the cap-and-trade scheme, a regulator defines an upper limit to the to-be-regulated quantity, e.g., emissions or congestion delays, and issues credits or permits to use parts of this overall quantity. As market participants can negotiate and allocate the credits among themselves, greater market efficiency is aimed for.

TMC research already developed among others the fundamental mathematical mechanism (Yang & Wang, 2011) also in a multimodal context (Balzer & Leclercq, 2022), compared its effectiveness to common road pricing (de Palma et al., 2018), and studied user perceptions, the system’s acceptance and its feasibility (Krabbenborg et al., 2020, 2021; Kockelman & Kalmanje, 2005). Recently, it has been proposed to use TMCs not only as a charge, but to use them as an incentives too and integrate TMCs into the entire transportation system, hence using the term “MobilityCoins” to describe this integrated nature (Bogenberger et al., 2021; Blum et al., 2022). In the following, we build on this particular implementation of a TMC scheme.

Variables	Explanation
P	Tradable mobility credit market price
X_{odm}	Share of travelers using the car on origin-destination pair (o, d)
T_{ij}	Travel time on link (i, j)
C_{ij}	Travel cost on link (i, j)
Q_{ij}	Flow on link (i, j)
Y_{ijk}	Partial flow on link (i, j) towards k
W_{odm}	Minimum travel costs from origin-destination (o, d) using mode m
M_{ij}	Minimum travel costs between i and j
μ	Mode choice scale parameter
γ	Initially issued credits per traveler
λ_{odm}	Origin-destination-specific MobilityCoins charges
κ_{ij}	Link-specific MobilityCoins charges
τ_{odm}	Free-flow travel time between origin o and destination d using mode m
t_{ij}^0	Free-flow travel time on link i - j
c_{ij}	Capacity on link i - j
β	BPR function parameter
n	BPR function parameter

Table 1: Variables and parameters in the model

In this paper, we present an integrated multimodal Wardropian model for the MobilityCoin system. The model solves for the user equilibrium of car, public transport and bicycle travelers. Travel times for public transport and bicycles are fixed, while car travel times incorporate congestion effects. The overall travel demand is distributed to modes using a logit-based mode choice based on the origin-destination travel times. The MobilityCoin system has two kinds of charges: origin-destination-based charges for all modes and link-specific charges for cars to influence route choice and manage congestion. We illustrate the model using a simple network configuration and explore how the different design parameters impact transport outcomes.

2 MODEL

Consider a transport network with \mathcal{N} nodes, \mathcal{A} arcs, and \mathcal{M} modes of transport. Nodes are referenced by $i \in \mathcal{N}$ (and j or k), arcs are a distinct pair of nodes and are referenced by the link star-end pair $(i, j) \in \mathcal{A}$, modes are referenced by $m \in \mathcal{M}$. In this model, three modes are considered: $\mathcal{M} \in \{\text{car, public transport, bicycle}\}$. Travelers are distinguished by their origin-destination pair $(o, d) \in \mathcal{OD}$. The set of origins and destinations is a subset of the set of nodes, i.e., $\mathcal{OD} \subseteq \mathcal{N}$.

In this macroscopic model, travelers make up to two choices. First, they choose their mode m . Second, all users choosing the car also choose their route. The equilibrium condition follows the Wardropian user equilibrium (Wardrop, 1952). The presented multimodal extension is a generalization of the seminal mathematical formulation presented by Yang & Wang (2011). The model defined in the following is formulated as a mixed-complementarity problem (MCP) (Ferris et al., 1999) and is implemented in GAMS (GAMS Development Corporation, 2018). The model’s variables and parameters are summarized in Table 1.

The overall demand d_{od} between origin o and destination d is fixed and exogenous. This demand is distributed across modes using a logit based assignment. As shown in Eqn. 1, the choice of modes depends on the minimum travel costs W_{odm} between o and d using mode m and a scale parameter μ .

$$X_{odm} = \frac{\exp(-\mu W_{odm})}{\sum_{m' \in \mathcal{M}} \exp(-\mu W_{odm'})} \quad (1)$$

In the proposed model, cars experience congestion effects as a function of the flow of vehicles, while public transport and bicycles have fixed travel times. Thus, the minimum travel cost depends on the chosen mode as defined in Eqn. 2. The minimum travel cost for cars results from the network

assignment of all cars, where M_{od} is the resulting minimum origin-destination travel cost which includes all MobilityCoins link charges. For public transport and bicycles, the minimum travel costs comprises the fixed origin-destination travel times τ_{odm} and the origin-destination specific MobilityCoins charges λ_{odm} valued at the MobilityCoins market price P .

$$W_{odm} = \begin{cases} M_{od}, & m = \text{car} \\ \tau_{odm} + P \cdot \lambda_{odm}, & \text{otherwise} \end{cases} \quad (2)$$

The car travel costs C_{ij} on link i - j comprises two elements. First, the travel time T_{ij} and second the MobilityCoins link charges κ_{ij} valued at the MobilityCoins market price P . The link travel time is defined in Eqn. 3 and follows the Bureau-of-Public-Roads (BPR) function (Bureau of Public Roads, 1964) with the usual parameters and is a function of link flow Q_{ij} .

$$T_{ij} = t_{ij}^0 \left(1 + \beta \left(\frac{Q_{ij}}{c_{ij}} \right)^n \right) \quad (3)$$

This then leads to the link car travel costs C_{ij} being computed as defined in Eqn. 4.

$$C_{ij} = T_{ij} + P \cdot \kappa_{ij} \quad (4)$$

The arbitrage condition for car drivers to use link (i, j) follows the Wardropian user equilibrium (Wardrop, 1952). It is formulated in the model as shown in Eqn. 5, where Y_{ijk} are the partial flows on that link towards k . Only when the minimum travel costs from node i to k over j equal the minimum travel costs from node i to k , the link is used for car drivers towards k .

$$C_{ij} + M_{jk} \geq M_{ik} \quad \perp \quad Y_{ijk} \geq 0 \quad (5)$$

The partial link flows Y_{ijk} can then be aggregated to link flows Q_{ij} as the sum over all partial flows along that links as defined in Eqn. 6.

$$Q_{ij} = \sum_k Y_{ijk} \quad (6)$$

In the model, it must be ensured that the inflows and outflows at each node in the network are balanced. This is ensured by Eqn. 7.

$$d_{od} X_{odcar} = \sum_{(o,j) \in \mathcal{A}} Y_{ojd} - \sum_{(j,o) \in \mathcal{A}} Y_{jod} \quad (7)$$

Last, as the MobilityCoins scheme is a market-based system, Eqn. 8 resembles the market clearing condition. Here γ is amount of credits initially issued per traveler. In other words, the left-hand side of Eqn. 8 results into the total market volume of MobilityCoins. κ_{ij} is the MobilityCoins link charge for car travelers and λ_{odm} is a origin-destination mode-specific charge for all other travelers. The complementarity conditions ensures that the MobilityCoins market price P is only non-zero when supply and demand are balanced. If the market is over-supplied, the market price would be consequently zero.

$$\gamma \cdot \sum_{(o,d) \in \mathcal{OD}} d_{od} \geq \sum_{(i,j) \in \mathcal{A}} \kappa_{ij} Q_{ij} + \sum_{(o,d,m) \in \mathcal{OD}} \lambda_{odm} * d_{od} * X_{odm} \quad \perp \quad P \geq 0 \quad (8)$$

3 A CASE STUDY

To illustrative the primary transport and economic mechanisms of a MobilityCoins scheme, we apply the model developed in Section 2 to the simple network shown in Figure 1. The network has 17 nodes of which 13 are origin and destination nodes and four are through nodes, i.e., the demand entering or exiting the network at these nodes is 0. The network has directed arcs as shown in Figure 1.

This network is centered around node “9”, while having symmetry with the line from nodes “2”, “10”, “9”, “11”, “7”. The full list of network and demand parameters will be provided in the full paper, but network parameters and origin-destination matrix are similar to the values present in the familiar Sioux Falls network. We set the scale parameter in the mode choice to $\mu = 0.01$ and the origin-destination travel times τ_{odm} for public transport and bicycles as follows. First, we calculate the car free-flow travel times in the network shown in Figure 1. Second, we set the public

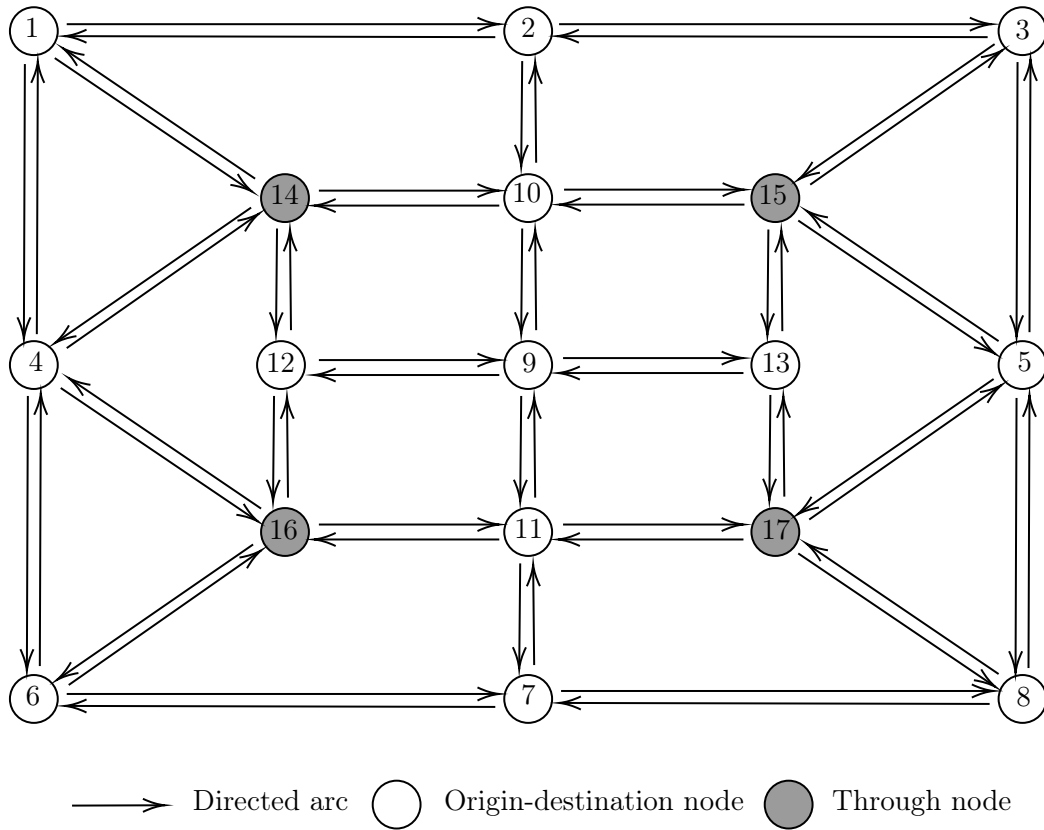


Figure 1: Case study network

transport travel times $\tau_{od,pt}$ on each origin-destination pair and the bicycle travel times $\tau_{od,bicycle}$ to a multiple of the car free-flow travel times. Third, we sample this multiplier for the public transport travel times from the uniform distribution in 1.35 to 1.45 and the bicycle travel times from the uniform distribution in 1.40 to 1.50.

In the case study, we define four scenarios. These are defined as follows.

- S1 Computes the status-quo scenario without any pricing, i.e., $\gamma = 0$ and $P \equiv 0$.
- S2 Imposes link charges for car travelers, i.e., $\kappa_{ij} \geq 0$, but no origin-destination specific charges for other modes of transport, i.e., $\lambda_{odm} = 0$. The charges κ_{ij} are set to 1.0 for $(i, j) \in \mathcal{N} \setminus \{9\}$ and 3.0 for $i \in \{9\} \vee j \in \{9\}$ to incentivize avoiding car travel in the inner zone of the network. The per-capita initial issue of MobilityCoins is evaluated at $\gamma \in \{1; 0.9; 0.8; 0.7; 0.6; 0.5\}$ to investigate transport outcomes when the overall budget of MobilityCoins is reduced. In other words, at the highest individual issue of MobilityCoins car travelers can travel one link outside the inner zone without acquiring additional MobilityCoins from the market.
- S3 Impose link charges for car travelers, i.e., $\kappa_{ij} \geq 0$, and allow as an incentive negative origin-destination specific charges for other modes of transport, i.e., $\lambda_{odm} \leq 0$. The values for κ_{ij} and γ are taken from S2. λ_{odm} is set for bicycles to -0.25 on all origin-destination pairs and to 0 for all public transport origin-destination pairs.
- S4 as a comparison we implement a congestion tax on all links from and to node “9”, i.e., fixing the product of $P \cdot \kappa_{ij}$. Considering the free-flow travel times of around 100 time units, we set $P \cdot \kappa_{ij} \in \{100; 200; 300; 400\}$ as these values increase the travel costs on these links considerably. All other links receive a charge one third of charges on the links from and to node “9”.

The results of the scenario analysis is presented in Section 4 along with their discussion.

Mode of transport	Trips		Travel time	
	Total	Share	Total (10e6)	Share
Car	84,409.6	27.5%	3.461	52.2%
Public transport	111,548.2	36.4%	1.624	24.0%
Bicycle	110,879.2	36.1%	1.676	24.8%
Total	306,837.0	100.0%	6.760	100.0%

Table 2: Trips and travel time in the status-quo scenario

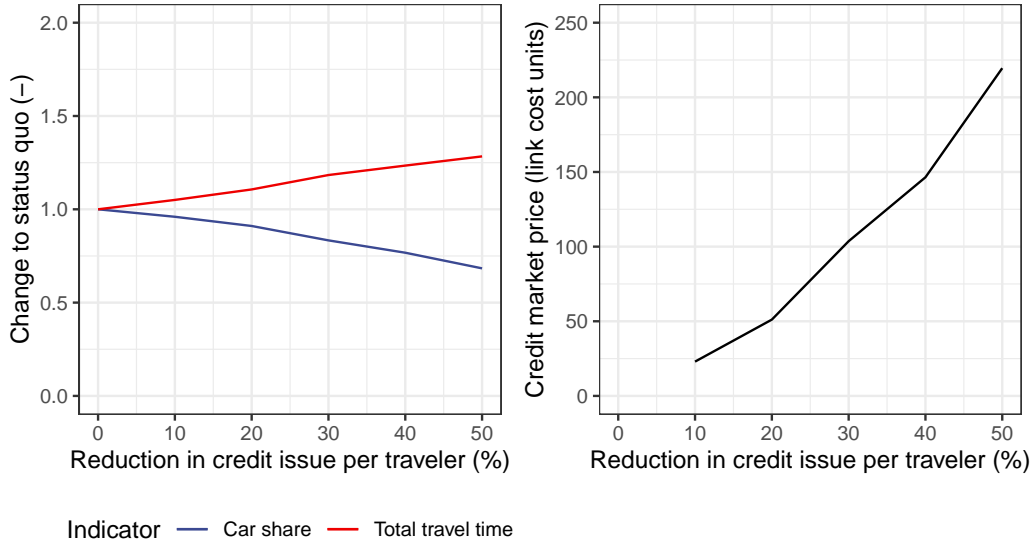


Figure 2: Changes of travel time, car share and credit price in a tradable credit scheme when the initial issue of credits is reduced.

4 FINDINGS

Scenario 1

The transport outcomes of status-quo scenario is summarized in Table 2. It can be seen that cars have the lowest modal share in terms of trips, but the highest in terms of travel time. This results from the congestion effects considered for this mode of transport. However, it is important to note here that the mode choice is only based on the origin-destination path costs, but does not consider trip length, number of transfers or any unobserved preferences that are usually impacting mode choice substantially (c.f. Ortúzar & Willumsen (2011); Train (2009)).

Scenario 2

In the second scenario, a conventional tradable credit scheme is implemented as described in Section 3. Figure 2 shows the results when the per-traveler issue of credits is gradually reduced from 1 credit to 0.5 credits. It can be clearly seen that when the overall market of credits is reduced by cutting down the initial issue, while leaving the parameter of the charging scheme $\kappa_{i,j}$ unaltered, car use declines. In this particular example with slower alternative modes, the total travel time in the system increases. This modal shift is achieved by an increasing credit market price resulting from a limited supply. In the particular example, the market price increases the travel costs on the arcs considerably when compared to the free-flow travel time.

Scenario 3

In the third scenario, a MobilityCoin system is implemented as described in Section 3. Note that the difference to the tradable credit scheme in the second scenario is that MobilityCoins are also

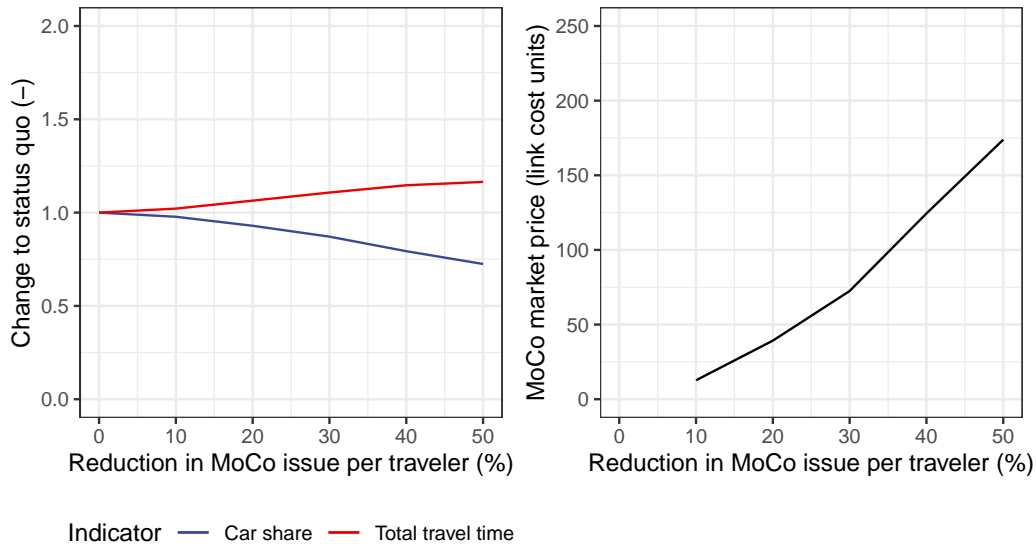


Figure 3: Changes of travel time, car share and the MobilityCoin (MoCo) market price in a MobilityCoins scheme when the initial issue of MobilityCoins is reduced.

used to incentive some mode choices, here the use of the bicycle. Figure 3 shows the results when the per-traveler issue of MobilityCoins is gradually reduced from 1 MobilityCoin to 0.5 MobilityCoins. A similar pattern is observed as for the common tradable credit scheme in Figure 2, but the changes to the transport outcomes compared to the status quo are not that strong. Arguably, using the bicycle generates additional MobilityCoins that are sold on the market; thus, the market volume is increased, leading to a lower market price compared to the second scenario and ultimately car use is not that strongly discouraged.

Scenario 4

In the fourth scenario, a conventional road user charging scheme with fixed charges is implemented as described in Section 3. Figure 4 shows the resulting impacts on the transportation system. Overall, a similar pattern as for the second (conventional TMC) and third scenario (MobilityCoins) is observed.

Comparison and discussion

The three different charging schemes presented in Scenario 2, 3 and 4 can be compared regarding their ability to reduce car trips (approximately related to the reduction in negative externalities) and the impact on travel times as a measure of impact on private costs. Figure 5 shows this comparison. First, it can be seen that the TMC scheme and conventional road user charging scheme (CC) with fixed charges perform similarly. This is perhaps surprising, but as shown in (de Palma et al., 2018) both schemes are equivalent when demand is fully adaptive as in this case study. Note that demand can fully adopt just based on travel costs and no multi-period constraints are considered. However, interestingly, we find that the MobilityCoin system achieves a similar reduction in car use (negative external costs) at lower travel time increases (private costs). This can be explained by the incentives provided to cyclists which adds more attractiveness to this mode in the mode choice. Nevertheless, this finding must be further corroborated with other system design parameter configurations and better behavioral parameters before making any generalization efforts.

5 DISCUSSION

The presented results underline the impacts a tradable mobility credit scheme, here MobilityCoins, has on transport outcomes. In particular, the scheme's benefit of reducing a desired quantity to

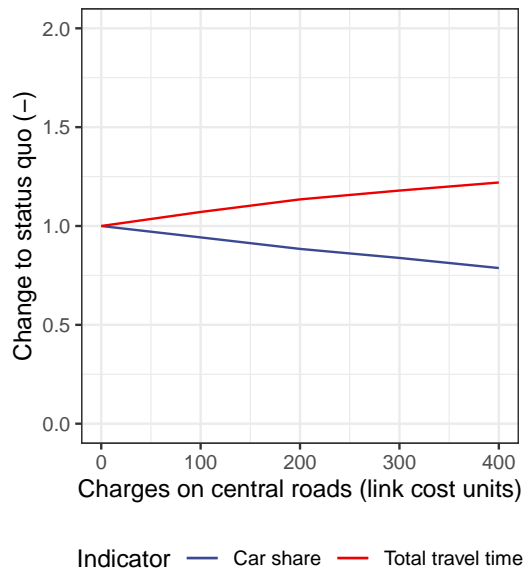


Figure 4: Changes of travel time and car share in a conventional road user charging scheme with fixed, but increasing charges.

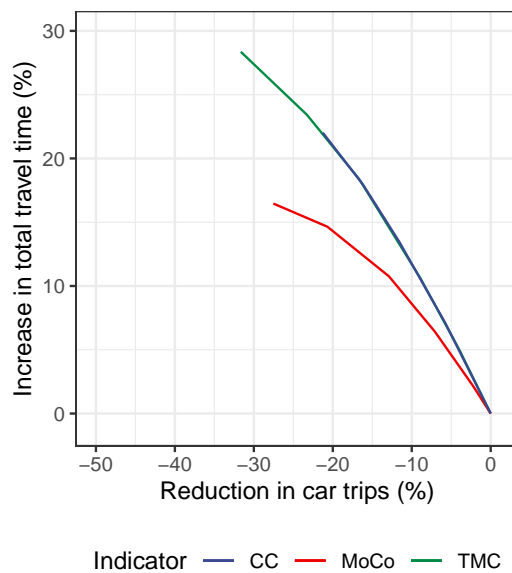


Figure 5: Comparing a conventional TMC scheme, a MobilityCoin scheme, and a conventional road user charging scheme (CC).

a target level, while providing direct financial incentives for travelers by direct transfers among themselves rather than redistributing tax revenue through a central organization. In addition, the provision of credits as an incentive, does not only increase trading activity and thus supports the market-based mechanism in general, but it seems to improve the economic allocation of resources by having more attractive alternatives.

Nevertheless, it is also apparent that a MobilityCoins or TMC scheme is not a simple system; it requires a careful policy design. Thus, for the identification of suitable policy designs future research has to start building models for real-world urban-scale cases for which appropriate choice parameters including unobserved preferences must be included (Train, 2009). In addition, the complex interactions of the key design parameters γ , κ_{ij} , λ_{odm} require the development of methods to identify those combinations - especially when considering the system's temporal evolution (Miralinaghi & Peeta, 2016) - that successfully and at little social costs lead to the desired targets. In addition, as will be an economic force in the decision making of individuals and firms, their impact on related fields like parking, housing and agglomeration should be focused on (Van Nieuwkoop et al., 2016; Loder et al., 2021; Venables, 2007).

ACKNOWLEDGEMENTS

Allister Loder acknowledges funding by the Bavarian State Ministry of Science and the Arts in the framework of the bidt Graduate Center for Postdocs. Further, both authors acknowledge support by the Free State of Bavaria funded research project “MobilityCoins – Anreiz statt Gebühr” and by the Federal Ministry of Education and Research (BMBF) funded project “MCube-SASIM” (Grant-no. 03ZU1105GA).

REFERENCES

- Axsen, J., Plötz, P., & Wolinetz, M. (2020). Crafting strong, integrated policy mixes for deep CO₂ mitigation in road transport. *Nature Climate Change*, *10*, 809–818. Retrieved from <http://dx.doi.org/10.1038/s41558-020-0877-y> (Publisher: Springer US) doi: 10.1038/s41558-020-0877-y
- Balzer, L., & Leclercq, L. (2022, June). Modal equilibrium of a tradable credit scheme with a trip-based MFD and logit-based decision-making. *Transportation Research Part C: Emerging Technologies*, *139*, 103642. Retrieved 2022-08-01, from <https://linkinghub.elsevier.com/retrieve/pii/S0968090X22000857> doi: 10.1016/j.trc.2022.103642
- Blum, P., Hamm, L., Loder, A., & Bogenberger, K. (2022, August). Conceptualizing an individual full-trip tradable credit scheme for multi-modal demand and supply management: The MobilityCoin System. *Frontiers in Future Transportation*, *3*, 914496. Retrieved 2022-08-10, from <https://www.frontiersin.org/articles/10.3389/ffutr.2022.914496/full> doi: 10.3389/ffutr.2022.914496
- Bogenberger, K., Blum, P., Dandl, F., Hamm, L.-S., Loder, A., Malcolm, P., ... Sautter, N. (2021, July). *MobilityCoins – A new currency for the multimodal urban transportation system*. arXiv. Retrieved 2022-06-24, from <http://arxiv.org/abs/2107.13441> (Number: arXiv:2107.13441 arXiv:2107.13441 [econ, q-fin])
- Bureau of Public Roads. (1964). *Traffic Assignment Manual* (Tech. Rep.). Washington D.C.: US Department of Commerce, Urban Planning Division.
- Dales, J. H. (1968). Land, Water, and Ownership. *The Canadian Journal of Economics*, *1*(4), 791–804.
- de Palma, A., Proost, S., Seshadri, R., & Ben-Akiva, M. (2018). Congestion tolling - dollars versus tokens: A comparative analysis. *Transportation Research Part B: Methodological*, *108*(July 2015), 261–280. doi: 10.1016/j.trb.2017.12.005
- Ferris, M. C., Meeraus, A., & Rutherford, T. F. (1999). Computing Wardropian equilibria in a complementarity framework. *Optimization Methods and Software*, *10*(5), 669–685. doi: 10.1080/10556789908805733

- Frei, F., Loder, A., & Bening, C. R. (2018, October). Liquidity in green power markets – An international review. *Renewable and Sustainable Energy Reviews*, *93*, 674–690. Retrieved 2018-11-24, from <https://www.sciencedirect.com/science/article/pii/S1364032118303691> doi: 10.1016/J.RSER.2018.05.034
- GAMS Development Corporation. (2018). *General Algebraic Modeling System (GAMS) Release 25.1*. (Place: Washington, DC, USA)
- Kockelman, K. M., & Kalmanje, S. (2005, August). Credit-based congestion pricing: A policy proposal and the public’s response. *Transportation Research Part A: Policy and Practice*, *39*(7-9), 671–690. doi: 10.1016/j.tra.2005.02.014
- Krabbenborg, L., Molin, E., Annema, J. A., Wee, B. V., Annema, J. A., & Wee, B. V. (2021). Exploring the feasibility of tradable credits for congestion management. *Transportation Planning and Technology*, *44*(3), 246–261. Retrieved from <https://doi.org/10.1080/03081060.2021.1883226> doi: 10.1080/03081060.2021.1883226
- Krabbenborg, L., Mouter, N., Molin, E., Annema, J. A., & van Wee, B. (2020, December). Exploring public perceptions of tradable credits for congestion management in urban areas. *Cities*, *107*, 102877. doi: 10.1016/j.cities.2020.102877
- Lindsey, R., & Santos, G. (2020, October). Addressing transportation and environmental externalities with economics: Are policy makers listening? *Research in Transportation Economics*, *82*, 100872. doi: 10.1016/j.retrec.2020.100872
- Loder, A., Schreiber, A., Rutherford, T. F., & Axhausen, K. W. (2021). Design of a multi-modal transportation system to support the urban agglomeration process. *Arbeitsberichte Verkehrs- und Raumplanung*, *1604*. Retrieved 2022-02-17, from <https://www.research-collection.ethz.ch/443/handle/20.500.11850/474590> (Publisher: IVT, ETH Zurich) doi: 10.3929/ETHZ-B-000474590
- Miralinaghi, M., & Peeta, S. (2016). Multi-period equilibrium modeling planning framework for tradable credit schemes. *Transportation Research Part E*, *93*, 177–198. Retrieved from <http://dx.doi.org/10.1016/j.tre.2016.05.013> doi: 10.1016/j.tre.2016.05.013
- Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling transport*. Chichester: Wiley-Blackwell.
- Perroni, C., & Rutherford, T. F. (1993, September). International Trade in Carbon Emission Rights and Basic Materials: General Equilibrium Calculations for 2020. *The Scandinavian Journal of Economics*, *95*(3), 257. Retrieved 2023-01-27, from <https://www.jstor.org/stable/3440355?origin=crossref> doi: 10.2307/3440355
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.
- Van Nieuwkoop, R., Axhausen, K. W., & Rutherford, T. F. (2016). A Traffic Equilibrium Model with Paid-Parking Search. *Journal of Transport Economics and Policy*, *50*(3), 262–286. Retrieved 2022-08-05, from <http://www.ssrn.com/abstract=2748539> doi: 10.2139/ssrn.2748539
- Venables, A. J. (2007). Evaluating urban transport improvements. *Journal of Transport Economics and Policy*, *41*, 173–188.
- Verhoef, E., Nijkamp, P., & Rietveld, P. (1997, August). Tradeable Permits: Their Potential in the Regulation of Road Transport Externalities. *Environment and Planning B: Planning and Design*, *24*(4), 527–548. Retrieved from <https://journals.sagepub.com/doi/abs/10.1068/b240527> doi: 10.1068/b240527
- Wardrop, J. G. (1952). Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers, Part II*, 325–378.
- Weitzman, M. L. (1974). Prices vs. Quantities. *The Review of Economic Studies*, *41*(4), 477–491.
- Yang, H., & Wang, X. (2011). Managing network mobility with tradable credits. *Transportation Research Part B: Methodological*, *45*(3), 580–594. Retrieved from <http://dx.doi.org/10.1016/j.trb.2010.10.002> doi: 10.1016/j.trb.2010.10.002

Random Utility Maximization model considering the information search process

Nova, G. *¹, Guevara, C.A.^{1,2}, Hess, S.^{3,4} and Hancock T.O.^{3,4}

¹ Department of Civil Engineering, University of Chile, Chile.

² Institute for Complex Engineering Systems (ISCI), Chile.

³ Choice Modelling Centre, University of Leeds, UK.

⁴ Institute for Transport Studies, University of Leeds, UK.

SHORT SUMMARY

Choice modelling has been dominated by static representations preferences due to their ease of implementation, transparent economic interpretability, and statistical coherency. Unlike, the Decision Field Theory (DFT) model explicitly includes the attribute scrutiny process within the choice decision, making it more closely related to the behavior that is observed in practice. However, the DFT model lacks the RUM model's microeconomic interpretability and has statistical limitations regarding the identification of the model parameters. This research introduces the "RUM-DFT" model, encompassing ideas from both approaches. Using Monte Carlo simulations and applying the proposed model to a database of real choices, it is first shown that the proposed approach can properly identify the parameters of the deliberation process, replicate the dynamic behavior of the utilities during the deliberation process; and retains full economic interpretability, since the estimated coefficients correspond to marginal indirect utilities when there is perfect knowledge of the information search process.

Keywords: RUM-DFT, Information Search Process, DFT, Cognitive Processing

1. INTRODUCTION

Discrete choice models have been widely proposed, used, and promoted in the literature for several decades. These correspond to a mathematical approach that allows estimations and predictions of the behaviors carried out by economic agents in different areas such as economics, health, marketing, and transportation, among others. Therefore, making a proper representation of the decision-making process, accurately accounting for the true behavioral mechanisms that are behind the choice dynamics, is crucial for making correct inference on the causal relations that are behind choices and for performing accurate forecasting to support informed public policy design. Neglecting the true dynamics that is behind the choice process will result in inconsistent estimators of the model parameters due to endogeneity, because of a model misspecification (Guevara, 2015).

Busemeyer and Townsend (1993) classify choice models according to whether utilities or preferences are dynamically or statically constructed. On the one hand, the static probabilistic models are discrete choice models that ignore the fact that choice probabilities are correlated with decision time and that deliberation time influences choice probability.

Although the different specifications of static choice models are extensive, simple to implement, have low computational cost and a high degree of economic interpretability of the parameters, they are conceptually unrealistic as they do not consider the cognitive process that individuals go

through when making a decision; for example, they do not include loss of information, filtration or information search cost. Under classical microeconomic and consumer behavior theory, most models assume that decision-makers evaluate and process information from alternatives in a perfect sense (Swait et al., 2001). However, it has been demonstrated that decision-makers may only focus on specific attributes, acquiring knowledge sequentially to make a final decision in simple and complex public transport route choice tasks (Nova, 2022). Likewise, Noguchi et al.,(2014), Stewart et al.,(2016), Stewart et al.,(2016), Sui et al.,(2020) evidenced, through eye tracking, that this behavior also occurs in simple, risky, and multi-attribute choices.

In contrast, dynamic models define choice probabilities are explicitly affected by the deliberation process, as the amount of time spent making a decision influences the final choice. Probabilities vary over time, as there is a constant acquisition and processing of information (attributes) that are incorporated to update the value of preferences or utilities before the choice is performed. Models that include a cognitive cost in the information search process, such as the Directed Cognition Model (Gabaix 2006) and the Adaptive Path Choice model have been shown to perform better than compensatory models in complex decision contexts (Gao 2011).

The dynamic processing of attributes has been represented by Decision Field Theory (DFT). It was initially designed as a cognitive model to capture the deliberation process in choice making (Busemeyer 1992, Busemeyer 1993). Then, DFT was then extended to a probabilistic-dynamic model that allowed for multiple attributes (Diederich, 1997) and was also generalized to multiple-alternative decision making (Roe et al.,2001). Recent contributions on DFT theory are from Hancock et al.,(2018) who improve the mechanisms that support the DFT model to make it more competitive with traditional discrete choice models. These advances allow for incorporating heterogeneity among and within decision-makers. Furthermore, Hancock et al.,(2021) introduce scale parameters in the basic mechanism of the DFT model, to avoid the requirement of conceptualizing a priori parameter values that may affect model estimation and identification. Finally, Hancock et al.,(2022) extend the model to include data from eye-tracking processes, to capture attribute attention weights more realistically during the deliberation process. All these recent studies show that the DFT model fits the data well and better than conventional static models. However, the DFT model has limitations, such as being based on ad-hoc matrix implementations of the model, identification problems, lack of a robust statistical theoretical framework and an approach compatible with the principle of random utility maximization that makes it impossible to interpret the parameters in a traditional way.

The combined limitations of current static and dynamic models motivate the need to create a new theoretical framework. This model should aim to keep the desirable properties of RUM model, whilst also overcoming its limitations with regards to its representation of the choice deliberation process. This could be achieved through the development of a RUM model that reflects cognitive dynamics, including the significant findings regarding the information search process, such as the process typically being breadth-first, that decision-makers revisit attributes more than once and that information is filtered. In this regard, this work introduces a new model 'RUM-DFT'. This new model also aims to rectify the DFT model's identification, inference, and parameter interpretation limitations. Likewise, the new model will include parameters that allow the deliberation process to be adequately modelled.

2. FORMULATION RUM-DFT

The RUM-DFT model proposes that, before choosing between alternatives, choice-makers perform a breadth-first information search process to update initially preconceived utilities. This means that individuals make comparisons of alternatives under a particular attribute at each preference updating step of the deliberation process until they make their choice. Specifically, the individual at each step t must choose whether to choose some alternative with the current information ($y^t = 1$) or whether to perform a new information search to update the utilities ($y^t = 0$). Suppose the individual decides on the option of choosing with the present utilities. In that case, the process of updating the underlying utilities is finished, and the one that provides the highest utility $[i|y^t]$ is chosen. By contrast, if the individual decides to continue with the information search process, the individual must determine which attribute will be attended to at step $t+1$ to update future utilities $[k(t+1)|y^t]$.

Hence, in this approach, we define a sequence of attributes attended to prior to the choice that ends with a choice. That is, $h^T = \{k(1), \dots, k(T)\}$ is defined as the attributes attended at each step t of the deliberation process up to the choice at T . Since, in traditional surveys, information such as eye-tracking data is not recorded, thus the sequence of attended attributes is not known, the set H^T is defined as all possible sequences that the individual can consider until the decision is made. Therefore, the probability of choosing an alternative i given the sequence of attended attributes and the probability of paying attention to an attribute k at step t ; given the decision to continue searching for information, can be represented by Eq. (1) and Eq. (2), respectively.

$$P(i|h^T) = P(i|y^T)\Omega(y^T|h^T) \quad [1]$$

$$P(k(t)|h^{T-1}) = P(k(t)|1 - y^T)\Omega(1 - y^{T-1}|h^{T-1}) \quad [2]$$

Where $P(i|y^T)$ is the probability of choosing alternative i conditional on the decision maker deciding to choose at step T , $P(k(t)|h^{T-1})$ is the probability of attending attribute k at deliberation step t conditional on the attributes attended up to step $t-1$, $\Omega(y^T|h^T)$ is the probability that the individual has decided to choose at step T and $\Omega(1 - y^{T-1}|h^{T-1})$ is that he/she decides to search for information.

The iterative process of searching for information and updating the utilities can be stopped for two main reasons. The first may be due to an external limitation that forces the person to choose an alternative in a maximum deliberation time. The second reason corresponds to the individual reaching their internal limit of preference. Thus, without an external limitation, the individual performs this process until future utilities (U_s^T) do not present a significant change compared to current utilities (U_c^T), modelled through an internal tolerance. Therefore, the probability that the individual has decided to choose, as shown mathematically in Eq. (3), allows the model to capture the difference between the expected value of choosing one of the alternatives at step T and the expected value of continuing to observe an attribute at step $T+1$.

$$\Omega(y^T|h^T) = P(|U_c^T - U_s^T| \leq \delta^T) \quad [3]$$

So far, no assumptions have been made about the functional form of utilities. The following section details the dynamics of the utilities and how the information search process is incorporated in breadth-first during the choice deliberation process up to the point at which a decision is made.

Utility Functions

The functional form of the utilities in the RUM-DFT model are proposed to capture the evolution of the individuals' preferences in accordance with the information search process. Therefore, the utilities explicitly represent the dynamic aspect, which depends on the attribute $k(t)$ attended to in step t of the deliberation process and on past information already considered, that is the utility of the previous instant U_{jn}^{t-1} , given by:

$$U_{jn}^t = \alpha \cdot U_{jn}^{t-1} + (1 - \alpha) \cdot \beta_{k(t)} \cdot X_{jn k(t)} + \varepsilon_{jn}^t \quad [4]$$

where α is the memory parameter representing the influence of time on past utilities (i.e., forgetting), $\beta_{k(t)}$ is the parameter of the attribute k attended in t , $X_{jn k(t)}$ corresponds to the value of attribute k of alternative i for individual n observed in step t , and ε_{ni}^t is the error that distributes extreme value I. It is necessary to point out that the random utility U_{jn}^t decomposes into two parts. The systematic part containing the information search process of each step t and the random part of the current step.

Calculating probabilities

The conditional probability $\Omega(y^T | h^T)$ that the person decides to choose in step T , given the sequence of attributes attended h^T , will depend on the difference between the current and future utilities.

$$\omega(y^T | h^T) = P(|V_c^T(h^T) + \varepsilon_c^T - V_s^T(h^T) - \varepsilon_s^T| \leq \delta^T) \quad [5]$$

Where δ is the threshold and in this dynamic, it is assumed that individuals become more intolerant over time, requiring more considerable expected changes to decide to continue searching for information.

Now, the probability of choosing an alternative i conditional on what the individual has decided to choose is like the one from the MNL model, but only considering the attributes attended up to step T :

$$P(i | y^T) = \frac{e^{V_i^T}}{\sum_j e^{V_j^T}} \quad [6]$$

Finally, the probability of attending to an attribute k is defined. Two possible formulations were considered for the analysis. On the one hand, a Logit model of constants indicating the weight of attention on each attribute in the deliberation process was considered (Φ_k), as shown in Eq. (7). On the other hand, a Logit model was constructed that considers the expected value of the change in overall utilities if the k^{th} attribute is observed in the next step. This aligns with the assumption that people perform breadth-first information search, as shown in Eq. (8).

$$\phi(k | h^t) = \frac{e^{\Phi_k}}{\sum_{\kappa} e^{\Phi_{\kappa}}} \quad [7]$$

$$\phi(k | h^t) = \frac{e^{\psi_{k(t)}^{t+1}}}{\sum_{\kappa} e^{\psi_{\kappa}^{t+1}}}, \quad \psi_{k(t)}^{t+1} = \ln \left(\sum_J |\phi_{k(t)} \cdot x_{jk(t)}| \right) \quad [8]$$

Thus, Eq. (5) and Eq.(6) allow for the construction of the probability of choosing an alternative i if the choice-maker decided to choose at step T under a particular sequence of attended attributes (h^T), as shown in Eq.(9). Similarly, Eqs.(7) and (8) construct the probability that he/she decides to search for information or attend to the k^{th} attribute at step t , given the sequence of attributes up to that step (h^T), as shown in Eq.(10). However, they only model a particular sequence of attended attributes within which the individual could have decided. In general, without knowing the deliberation process, a modeler must integrate or consider all possible sequences of attended attributes H^T , which results in the probability of choosing alternative i being as shown in Eq.(11).

This specification, which we name RUM-DFT-SC (A RUM-DFT model explicitly for stated choice scenarios) considers all possible sequences of attended attributes without some maximum deliberation time. Therefore, to estimate this model specification, the maximum deliberation time T_{max} must be fixed. However, if there is some knowledge about the information search process or decision-makers deliberation, the estimation process can be reduced.

$$P(i|h^T) = \left(\frac{e^{V_i^T}}{\sum_j e^{V_j^T}} \right) \left(\frac{e^{V_s^T + \delta^T}}{e^{V_c^T} + e^{V_s^T + \delta^T}} + \frac{e^{V_c^T + \delta^T}}{e^{V_s^T} + e^{V_c^T + \delta^T}} \right) \quad [10]$$

$$P(k(t)|h^{T-1}) = \left(\frac{e^{\phi_k}}{\sum_\kappa e^{\phi_\kappa}} \right) \left(\frac{e^{V_c^{T-1}}}{e^{V_c^{T-1}} + e^{V_s^{T-1} + \delta^{T-1}}} + \frac{e^{V_s^{T-1}}}{e^{V_s^{T-1}} + e^{V_c^{T-1} + \delta^{T-1}}} \right) \quad [11]$$

$$P(i) = \sum_{h^T}^{H^T} P(i|h^T) \cdot P(h^T) \quad [12]$$

First, it is possible to have information on the time or the number of steps that individuals perform in the choice process since they can be considered valid proxies on cognitive processes (Horstmann, 2009). From this approach, the likelihood is reduced to the possible sequences of attributes attended up to the maximum time of deliberation found in the database, obtaining the specification RUM-DFT-DT (with deliberation time):

$$P(i) = \sum_{h_i^T \in H^T} P(i|h_i^T) P(h_i^T) \quad [13]$$

Secondly, the actual sequence of the attributes attended by the respondents can be uncovered using instruments of process data tracking, such as a mouse-tracker, click-tracker or eye-tracker (for example Nova and Guevara (2022)). This data reduces the probability of choice to only one sequence of attributes, resulting in the specification RUM-DFT-IS (with information search process):

$$P(i) = \sum_{h^T}^{H^T} P(i|h^T) \cdot P(h^T) = P(i|\tilde{h}^{T_n}) \quad [14]$$

Where \tilde{h}^{T_n} is the sequence of attended attributes specified for decision maker n with a deliberation time T_n .

3. RESULTS AND DISCUSSION

We analyze the performance of the new model when compared to conventional models in three simulated case studies and one stated preference study. The aim is to test the explanatory power

of the RUM-DFT model and to verify whether the parameters are recovered correctly compared to other conventional models.

RUM-DFT-IS

First, 20 simulations and estimates of database A were generated considering the RUM-DFT-IS model, in which the sequence of attributes attended is known. Given this knowledge regarding attribute attendance, it is not necessary to calculate the probabilities of deciding to choose or continue with the information search process at each step t.

Table 1: Average of RUM-DFT model estimates considering the attributes attended.

20 Iterations (* p<0.05)	Value	Estimate	S.E	$t(\hat{\beta})$
β_c	-0.5	-0.492	0.020	0.888
β_{tt}	-2.0	-1.986	0.128	0.771
U_A^0	-0.5	-0.486	0.271	0.655
U_B^0	-0.5	-0.512	0.316	0.757
U_D^0	-1.0	-1.031	0.318	1.237
α	0.6	0.598	0.056	0.757
LL		-4,046.047		
$\bar{\rho}^2$		0.269		
AIC		8,102.1		
BIC		8,134.6		
Time [min]		2.6 (0.2)		

Table 1 shows the average of these estimates and from these results it is possible to see that statistically significant parameters are obtained in most of the cases. Moreover, in all the iterations β_t , β_c and α are statistically significant, the latter being the estimate with the highest efficiency. Therefore, the RUM-DFT-IS model can be applied to process data independent of the generated attributes that define the route choice situation in each iteration. It should be highlighted that this approach makes it possible to incorporate the attended attributes explicitly at each step of the deliberation process into the modelling. It will also be able to deliver estimates of the coefficients of the attributes plus the deliberation process correctly in magnitude, with expected signs and relative importance, allowing the decision-makers to represent the information search process adequately.

RUM-DFT-DT

The second specification considers that the deliberation time of the respondents is known, but not the sequence of attributes attended. This analysis compares the proposed model with a DFT model that includes the total number of fixations in the attention weights. At this level, it is reasonable to compare these approaches as they both use fixations in an aggregated form to represent deliberation time. To estimate these models, we simulated a number of fixations for each choice task for all decision-makers.

Table 2 shows the estimation results of the RUM-DFT-DT model. It is worth mentioning that, like the previous specification, it is not feasible to estimate the parameters of the deliberation process, but the memory factor can be known. Based on the results, the estimated parameters are

close to the real value with which they were generated, being the memory factor (α) and the coefficient associated to the travel time (β_t) the most efficient ones.

Table 2: Estimation results of RUM-DFT and scaled DFT model including the aggregate fixations of the deliberation process.

1 Iteration (* p < 0.05)	Value	RUM-DFT-DT	DFT-Scaled-DT1	DFT-Scaled-DT2
β_c	-0.5	-0.510*		0.411*
β_t	-2.0	-2.046*		1.000
γ_c			-1.069*	
γ_t			0.000	
U^0/P^0_A	-0.5	-0.772*	1.429	0.930
U^0/P^0_B	-0.5	-0.211*	2.337*	1.418
U^0/P^0_D	-1.0	1.258*	-1.514	-0.885
α	0.6	0.670*		
ϕ_1			0.781	0.105
ϕ_2			-0.016	-0.106
σ_ε			6.311*	4.560*
τ			1.875*	0.876
α_f			0.418*	0.496*
LL		-4,153.369	-4,205.220	-4,206.560
$\bar{\rho}^2$		0.250	0.240	0.240
AIC		8,318.738	8,428.440	8,431.120
BIC		8,356.502	8,485.086	8,487.766
Time [min]		21.3	4.1	3.3

RUM-DFT-SC

The third application corresponds to a case in which neither the sequence of attributes attended, nor the deliberation time, is known by the researcher. This is the case with most SP and RP data sources.

Table 3 shows the estimation results of the RUM-DFT-SC, DFT-scaled-1, DFT-scaled-2 (not including fixations) and MNL model. It is observed that the proposed model presents the best log-likelihood and lower values in the AIC and BIC information criteria than the rest of the approaches. The cost and travel time parameters (attention weights in DFT) differ significantly from zero in all models. However, only the RUM-DFT-SC model delivers close values, in magnitude and sign, to the true ones. The alternative-specific constants cannot be recovered correctly in most cases. Therefore, testing the proposed model with the methodological improvement mentioned in the previous section is necessary.

Table 3: Estimation results of RUM-DFT-SC and scaled DFT model.

Parameters (* p<0.05)	Value	RUM-DFT-SC $T_{max}=8$ steps	DFT-Scaled-1	DFT-Scaled-2	MNL
β_c	-0.5	-0.476*		0.000	-0.295*
β_t	-2.0	-1.959*		0.844*	-0.732*
γ_c			1.000		
γ_t			0.838*		
U_A^0/P_A^0	-0.5	-1.779	-0.469	-0.746	0.000
U_B^0/P_B^0	-0.5	0.444	-0.596	-2.904	-0.132*
U_D^0/P_D^0	-1.0	-3.223	0.000	-1.283	-0.047
δ	0.001	0.093			
ϕ_1			25.545*	2.557	
ϕ_2			-0.114*	-0.016	
σ_ε			6.881*	7.887*	
τ			2.780*	2.969*	
LL		-4,221.788	-4,385.14	-4,386.42	-4,379.93
$\bar{\rho}^2$		0.238	0.209	0.208	0.210
AIC		8,455.58	8,786.27	8,788.84	8,769.87
BIC		8,493.34	8,836.62	8,839.2	8,801.34
Time [min]		41.2	4.2	3.5	0.2

Empirical application to SwissMetro Data

This section compares the RUM-DFT-SC against the DFT-Scaled, RUM, C RRM, mu RRM and P RRM using the SwissMetro database (Bierlaire et al.,2001). The DFT model shown corresponds to the one that estimates the attention weights (Hancock et al.,2018). However, the memory value is fixed to 0, the sensitivity to 0, and the error term to 1 since these values are generally insignificant when there is no information on the deliberation of the respondents. The RUM approach corresponds to a Multinomial Logit, and the last three are variants of the RRM model in which the depth of regret is incorporated (van cranenburgh et al., 2015).

The results shown in Table 4 demonstrate that the models that include parameters that model the deliberation process, both the RUM-DFT model and the DFT approach, have a better performance than the rest, in terms of log-likelihood, AIC and BIC information criteria. However, only from the proposed model is it possible to make an economic interpretation of the coefficients of the attributes, with the tolerance, memory and attention weight also allowing for interpretation of the information search process.

Table 4: Model estimates applied to real SwissMetro data.

Parameters (* p 0.05)	RUM-DFT-SC $T_{max}=7$ steps	RUM-DFT-SC $T_{max}=10$ steps	RUM MNL	DFT	C RMM ¹	μ RRM ¹	P RRM ¹
β_c	-3.745*	-3.657*	-1.150*		0.010*	0.010*	0.010*
β_t	-4.665*	-4.649*	-1.270*		0.010*	0.010*	0.010*
γ_c				-0.867*			
γ_t				-1.251*			
U_{Train}^0	-0.902*	-0.904*	-1.168	-1.120*	N.R.	N.R.	N.R.
U_{Car}^0	-0.010	-0.013	0.250	-0.076*	N.R.	N.R.	N.R.
δ	0.273*	0.259*					
α	0.959 ²	0.931 ²					
ϕ_{time}	0.467*	0.482*					
ϕ_1				0.000			
ϕ_2				0.000			
σ_ε				1.000			
τ				0.564*			
LL	-4,233.941	-4,233.918	-4,382.500	-4,277.170	-4,539.672	-4,373.356	-4,418.252
$\bar{\rho}^2$	0.312	0.312	0.305	0.305	0.263	0.290	0.283
AIC	8,479.882	8,479.820	8,772.712	8,564.340	9,087.265	8756.452	8844.482
BIC	8,519.673	8,519.611	8,799.626	8,597.499	9,113.618	8,790.094	8,870.835
Time [min]	59.7 ² + 8.3 ³	301.9 ² + 3.4 ³	0.03	7.03	N.R.	N.R.	N.R.

¹Estimates from (Belgiawan et al., 2017).

²Estimated using DEoptim.

³Estimated using OptimParallel setting α .

4. CONCLUSIONS

This work meets the initially stated general objective since it was possible to incorporate the characteristics of the deliberation process implicit in the public transport route choices as methodological improvements in the development of the Random Utility Maximization model considering the information search process (RUM-DFT)

The results of the RUM-DFT model are promising for the first simulated case study, as this specification recovers the parameters, and the utilities behave as expected. This is mainly because this approach avoids the integration in the space of all possible sequences of attributes, which considerably increases its dimension at each additional step (K^t). For the other simulated cases, the model provides attribute parameters close to the real values. However, this is not the case for the initially preconceived utilities.

On the other hand, the RUM-DFT model was estimated using the SwissMetro database. Signs, magnitudes and significance of parameters, goodness-of-fit indicators and estimation time were compared with the classical models. In general, the models incorporating the assumptions supported in this paper obtain the best fit indicators. Moreover, the RUM-DFT model specification, which in this case does not include information about the deliberation process, can still significantly estimate the attribute, tolerance and recall coefficients with reasonable values, successfully outperforming the DFT and MNL approaches.

REFERENCES

- Bierlaire, M., Axhausen, K., Abay, G., 2001. The acceptance of modal innovation: The case of swissmetro, in: Swiss transport research conference.
- Busemeyer, J.R., Townsend, J.T., 1992. Fundamental derivations from decision field theory. *Mathematical Social Sciences* 23, 255–282.
- Busemeyer, J.R., Townsend, J.T., 1993. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review* 100, 432.
- Chorus, C., 2012. Random regret minimization: an overview of model properties and empirical evidence. *Transport reviews* 32, 75–92.
- Chorus, C.G., 2010. A new model of random regret minimization. *European Journal of Transport and Infrastructure Research* 10.
- Chorus, C.G., Arentze, T.A., Timmermans, H.J., 2008. A random regret-minimization model of travel choice. *Transportation Research Part B: Methodological* 42, 1–18.
- Diederich, A., 1997. Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology* 41, 260–274.
- Gabaix, X., Laibson, D., Moloche, G., Weinberg, S., 2006. Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review* 96, 1043–1068.
- Gao, S., Frejinger, E., Ben-Akiva, M., 2011. Cognitive cost in route choice with real-time information: An exploratory analysis. *Procedia-Social and Behavioral Sciences* 17, 136–149.
- Guevara, C.A., 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice* 82, 240–254.
- Hancock, T.O., Hess, S., Choudhury, C.F., 2018. Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. *Transportation Research Part B: Methodological* 107, 18–40.
- Hancock, T.O., Hess, S., Choudhury, C.F., 2022. Secret in their eyes’: Incorporating eye-tracking and stress indicator data into travel behaviour models Working paper University of Leeds.
- Hancock, T.O., Hess, S., Marley, A.A., Choudhury, C.F., 2021. An accumulation of preference: two alternative dynamic models for understanding transport choices. *Transportation Research Part B: Methodological* 149, 250–282.
- Horstmann, N., Ahlgrimm, A., Glöckner, A., 2009. How distinct are intuition and deliberation? an eye-tracking analysis of instruction-induced decision modes. *An Eye-Tracking Analysis of Instruction-Induced Decision Modes* (April 1, 2009). MPI Collective Goods Preprint.
- Noguchi, T., Stewart, N., 2014. In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition* 132, 44–56.

Nova, G., Guevara, C.A., 2022. In depth, breadth-first or both? characterising the information search process in a public transport sp experiment Working paper University of Chile.

Roe, R.M., Busemeyer, J.R., Townsend, J.T., 2001. Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological review* 108, 370.

Stewart, N., Gächter, S., Noguchi, T., Mullett, T.L., 2016a. Eye movements in strategic choice. *Journal of behavioral decision making* 29, 137–156.

Stewart, N., Hermens, F., Matthews, W.J., 2016b. Eye movements in risky choice. *Journal of Behavioral Decision Making* 29, 116–136.

Sui, X.Y., Liu, H.Z., Rao, L.L., 2020. The timing of gaze-contingent decision prompts influences risky choice. *Cognition* 195, 104077.

Swait, J., Adamowicz, W., 2001. The influence of task complexity on consumer choice: a latent class model of decision strategy switching. *Journal of Consumer Research* 28, 135–148.

van Cranenburgh, S., Guevara, C.A., Chorus, C.G., 2015. New insights on random regret minimization models. *Transportation Research Part A: Policy and Practice* 74, 91–109.

The influence of the built environment on real world car energy efficiency

Chris Djie ten Dam*¹, Francisco Bahamonde-Birke², Dick Ettema³, Gert Jan Kramer,
Vinzenz Koning

¹ PhD candidate, Copernicus institute, Utrecht University, the Netherlands

² Assistant professor, Department of Methodology and Statistics, Tilburg University, the Netherlands

³ Professor, Department of Human Geography and Spatial Planning, Utrecht University, the Netherlands

⁴ Professor, Copernicus institute, Utrecht University, the Netherlands

⁵ Assistant professor, Copernicus institute, Utrecht University, the Netherlands

SHORT SUMMARY

To reduce CO₂ emissions and safeguard our energy security, we need to electrify our car-fleet, increase its efficiency, and limit car-dependence. This paper therefore answers the following research question: How and to what extent are features of residential neighborhoods and their residents currently related to energy relevant car type choice? For this purpose, it analyzes Dutch vehicles' real-world energy consumption with a multilevel discrete choice model of fuel- and weight-preferences in one- and multicar households. Small, lower-income, female households in non-green (urban) environments tended to own light, efficient vehicles. Households with a private parking spot tended to own heavy, electric vehicles. Lastly, households with multiple cars tended to live in non-urban areas and to prefer heavier vehicles. These correlations imply that studies that omit vehicle energy efficiency underestimate the environmental impact of urban planning interventions. However, improving vehicle testing procedures may be a more effective energy-saving strategy.

Keywords: built environment, car ownership, discrete choice modelling, energy efficiency, electric vehicles

1. INTRODUCTION

Vehicle gasoline consumption causes climate change and threatens energy security. Electric vehicles can reduce emissions. Yet, these EVs come with energy security concerns of their own due to their rare mineral requirements (International Energy Agency, 2022). Moreover, they can cause a prohibitive increase in electricity use (Galvin, 2022). Especially heavy EVs consume a lot of electricity (Galvin, 2022; Weiss et al., 2020). It is thus important to limit the deployment of heavy, energy inefficient vehicles as well as of cars in general.

A large body of earlier research has shown that a sustainable urban environment with high densities and limited distances to city centers can enable green-minded citizens to live a car-free life (Banister, 2011; Ewing and Cervero, 2010; Næss, 2012; Newman and Kenworthy, 1989; Silva et al., 2017; Stevens, 2017). Other studies have shown that residents of dense urban areas are less likely to own large (inefficient) vehicle designs like vans, trucks, and SUVs (Bhat et al., 2009; Brownstone and Fang, 2014; Cao et al., 2006; Chen et al., 2021; Eluru et al., 2010; Garikapati et al., 2014; Li et al., 2015; Liu and Shen, 2011; McCarthy and Tay, 1998; Potoglou, 2008; Prieto and Caemmerer, 2013; Song et al., 2016).

Yet, most of these studies ignored efficient compact cars. Most also did not actually compute car energy efficiency whereas an SUV does not necessarily consume more energy than a sedan (Li et al., 2015; Timmons and Perumal, 2016). Moreover, no articles could be found that directly analyze the correlation between the built environment and car weight: the major determinant of both conventional and electric vehicles' energy efficiency. A clear link between the built environment and the (future) energy consumption of cars has thus not been established.

The present paper will help fill this gap by answering the following research question: How and to what extent are features of residential neighborhoods and their residents currently related to energy relevant car type choice? For this purpose, it made use of real-world specific energy consumption data in megajoules per vehicle kilometer (MJ/vkm) of cars in the Netherlands. The energy-relevant car type choices were analyzed with a multilevel discrete choice modeling framework of fuel type and weight preferences in one- and multicar households.

2. METHODOLOGY

Travel, sociodemographic and built environment data

Travel and sociodemographic data were obtained from the Netherlands Mobility Panel (MPN) from KiM Netherlands Institute for Transport Policy Analysis (Hoogendoorn-Lanser, 2019). This panel consists of multiple surveys and a three-day travel diary. Household members complete these on a predetermined moment in September, October, or November.

This study analyzed data from households who participated in 2019, 2018, or 2017. Preference was given to the latest year available. The response rate is likely similar to the 64% in 2013 (Hoogendoorn-Lanser et al., 2015). In the end, 4316 households were included who together owned 3498 cars of which the energy use could be accurately determined. All vehicles were analyzed. The sample weight was used to avoid overrepresentation of the multicar households.

KiM provided us with the respondents' residential addresses on the postcode-6 level (1234 AB, representing part of a street). Specifically, we used the postcode-6 addresses to couple the local address density and distances to stations and big (transfer) stations from Statistics Netherlands (CBS; Statistics Netherlands, 2018, 2019). Moreover, this data was analyzed with Geopy to compute distances to city center proxies: destination-rich postcode-6 areas. We also included the NDVI green-space and land-use mix indexes from the Vitality Data Center (VDC) Project (Ren et al., 2019; Wang, 2020; 202). A number of other built-environment and sociodemographic variables were included from the MPN-survey itself. The variables are described in Table 1.

Car energy data

The discrete choice model assessed the effect of the above-described variables on direct (consumer) energy. This can be easily converted to tailpipe CO₂ emissions.

The Netherlands Vehicle Authority (RDW) registers fuel use (Team Open Data RDW, 2021) based on the standardized New European Driving Cycle (NEDC). However, the Netherlands Organization for Applied Scientific Research (TNO) has shown this data to be biased. The gap with the real-world fuel use varies systematically with vehicle building year and can be expected to depend on other vehicle characteristics as well. It was therefore decided to instead use real-world data from fuel-cards from Travelcard Nederland BV. The data were scraped from Praktijkverbruik.nl and coupled based on the MPN vehicles' fuel type, building year, and model.

Table 1: The explanatory variables included in the data analysis

Included variables	Resolution	Unit or categories	Source
Density_PC5	Postcode-5	The average number of addresses per km ² in a 1 km buffer around the addresses of the inhabitants of the postcode-5 area	CBS
log_km_station	Postcode-6	The natural logarithm of the average travel distance of the inhabitants of the postcode-6 area to the nearest train station	CBS
log_km_bigstation	Postcode-6	The natural logarithm of the average travel distance of the inhabitants of the postcode-6 area to the nearest transfer train station	CBS
km_hugecenter	Postcode-6	Distance of center postcode area to the nearest center of a postcode-6 area containing ≥ 500 known destinations on average in 1 km travel distance of inhabitants: Amsterdam, Rotterdam, the Hague, and Utrecht.	CBS
km_largecenter	Postcode-6	Distance of center postcode area to the nearest center of a postcode-6 area containing ≥ 200 known destinations on average in 1 km travel distance of inhabitants. These are the centers of 23 cities, including the huge centers	CBS
NDVI	Postcode-6	Average Normalized Difference Vegetation Index for a 1 km buffer around each postcode-6 area	VDC
land-use	Postcode-6	Average entropy index of the mix between five land-use categories for a 1 km buffer around each postcode-6 area	VDC
km_bus	Household	Travel distance of the household to a bus stop with a frequency of at least 1 bus per hour	MPN
Parkingspot	Household	Whether the household has a private parkingspot	MPN
HH_under18	Household	Number of household members of less than 18 years old	MPN
HH_18to39	Household	Number of household members of 18 to 39 years old	MPN
HH_40to59	Household	Number of household members of 40 to 59 years old	MPN
HH_60plus	Household	Number of household members of at least 60 years old	MPN
inc<20k	Household	A household income of less than 20 thousand euros per year	MPN
inc40to60k	Household	A household income of 40 to 60 thousand euros per year	MPN
inc60to120k	Household	A household income of 60 to 120 thousand euros per year	MPN
inc \geq 120k	Household	A household income of ≥ 120 thousand euros per year	MPN
Males	Household	Fraction of adult household members who are male	MPN
Higheducation	Household	Fraction of adult household members who have a university degree	MPN
Workers	Household	Fraction of adult household members who work at least 24 hours per week	MPN

Travelcard data was available for 60% of the valid MPN vehicles. The energy used by remaining gasoline, diesel, and gasoline hybrid vehicles was instead computed using TNO-models calibrated with Travelcard-data (de Ruiter et al., 2021). These estimate emissions based on car weight, building year, and engine power. The precise energy efficiency of plug-in hybrid (PHEV) and battery electric (BEV) car-models was taken from TNO Travelcard-based research (van Gijlswijk et al., 2020; de Ruiter et al., 2021).

Non-electric cars with missing fuel type, building year, or weight data or with a registered weight under 500 kilograms were excluded from the analysis.

Data processing

Data cleaning and standardization was done using Pandas and Sklearn (pandas development team, 2020; Wes McKinney, 2010; Pedregosa et al., 2011). Variables that were insignificant at the 20% level or that were insignificant at the 10% level for all car ownership classes, fuel types, and weight classes were excluded. The sample-weights were scaled to avoid in- or deflated P-values.

The cars were lastly categorized into types based on fuel type and weight: the main determinants of energy consumption. The (hybrid) electric vehicles (HEVs) were given their own category

because of their importance for future energy consumption. Diesel vehicles were given their own fuel type class too since these efficient vehicles constitute a major fraction of the sample. The standard (mostly gasoline) and diesel cars were subdivided into the following weight categories: light (<1000 kg), midlight (1000-1250 kg), midheavy (1250-1500 kg), and heavy (>1500 kg).

Car energy exploration

The real-world energy consumption of the vehicles per car type category is shown in the boxplots of Figure 3. The BEVs are visible as a group of HEVs (green) consuming less than 1 MJ/vkm. Diesel vehicles (red) are also efficient, with the exception of heavy diesel vans. The old standard-fueled cars (blue) consume around 3-4 MJ/vkm, but these often have a low sample-weight. As expected, heavy standard cars also use considerably more energy than their lighter counterparts.

The energy consumption quantiles according to the NEDC test cycles are shown in gray for reference purposes. As expected, this official data consistently underestimates real-world energy consumption. The gap seems somewhat larger for light vehicles.

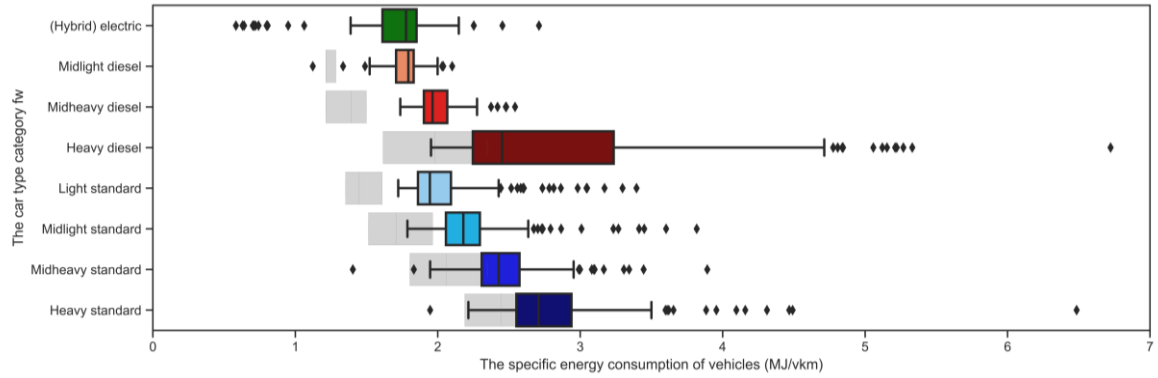


Figure 1: The real-world energy consumption of the eight car type categories fw

The discrete choice model

It was decided to analyze cars' real-world specific energy consumption using a multilevel discrete choice modeling framework. Given that decisions on car ownership cannot be disentangled from household preferences for fuel- and weight-based car types fw , both choices were modeled jointly, taking the multilevel characteristics of the decision-making process into account: the number of vehicles available to households influences the types of vehicles being purchased. Both decisions are fundamental to understanding households' travel energy use. Moreover, both decisions depend on the household's sociodemographic characteristics and the built environment.

At a first stage, car ownership classes were considered using a discrete choice model, whereby the utility U_{cn} of each of the three classes c (no car, one car, or two or more cars) for a household n was the sum of the utilities associated with the v sociodemographic and built environment variables x_{in} as determined by the estimated coefficients β_{ic} , the aspecific constant ASC_c and the (EV1 iid) unobserved utility term ε_{cn} :

$$U_{cn} = ASC_c + \sum_{i=1}^v \beta_{ic} x_{in} + \varepsilon_{cn} \quad (1)$$

Then, a car type model was specified, which considered the number of cars owned as discrete latent attributes. This car type model estimated household choices for car fuel types f and weight

categories w explicitly by defining the utility of each of the eight fuel- and weight-based car types $U_{fwn|c}$ as the sum of the utility of the fuel type $U_{fn|c}$, the utility of the weight category $U_{wn|c}$, an aspecific constant ASC_{fw} , and the (EV1 iid) unobserved utility term ε_{fwn} . The utility of the fuel type and weight category were adjusted by a fixed amount β_{2car} in households that were estimated to own two or more vehicles.

$$U_{fwn|c} = ASC_{fw} + U_{fn|c} + U_{wn|c} + \varepsilon_{fwn} \quad (2)$$

$$U_{fn|c} = \sum_{i=1}^v \beta_{if} x_{in} + \beta_{2car,f} \quad (3)$$

$$U_{wn|c} = \sum_{i=1}^v \beta_{iw} x_{in} + \beta_{2car,w} \quad (4)$$

Attempted nested logit and mixed logit models collapsed into the multinomial model. It was thus decided to analyze car types using the above-explained Multinomial Logit specification. This allowed us to use the independence of irrelevant alternatives property such that a household's probability of choosing a light gasoline over a heavy gasoline vehicle was only determined by the weight coefficients β_{iw} . These weight category coefficients should therefore remain valid in a future sample with more electric vehicles.

The joint model was estimated in Biogeme (Bierlaire, 2020) by maximizing the loglikelihood LL function below. P_{cn} and $P_{fwn|c}$ are the respective probability of belonging to the car ownership class and having a car of a certain fuel- and weight-based type as given by the well-known multinomial logit equation. The dummy q is 1 if the household owns a valid car and 0 otherwise.

$$LL(\beta_{ic}, \beta_{if}, \beta_{iw}) = \ln(\prod_n \prod_c P_{cn}(c|x_i; \beta_{ic}, \varepsilon) \prod_n \prod_{fw} [\sum_c P_{fwn|c}(fw|x_i, c; \beta_{if}, \beta_{iw}, \varepsilon) \times P_{cn}(c|x_i; \beta_{ic}, \varepsilon)]^q) \quad (5)$$

3. RESULTS AND DISCUSSION

The study results are provided in Table 2 and described below.

As expected, large families with many employed individuals, a middle- to high-income, and a non-urban residential location tended to possess a car. Households owning multiple cars were especially likely to have many adult and working members, a high income, and a low-density residential environment. They had a preference for heavy, non-diesel vehicles.

Owning heavy rather than light vehicles was directly correlated with a large number of (older) family members, a higher household income, and a high fraction of males. Living in a green area and having a private parking spot increased (mid)heavy over light vehicle ownership as well, which is logical as compact cars are easier to park and maneuver in densely built areas. Yet, electric vehicles were also owned by households with a private parking spot. One explanation is that EV owners prefer to charge their vehicles at home (Westin et al., 2018).

Interestingly, the local address density, street connectivity, and distances to public transport had no significant direct effect on vehicle fuel type or weight. Previously found effects of these variables on vehicle choice may be due to correlations with open (green) space and parking possibilities. Earlier studies may also have captured indirect effects of the built environment through ownership of two or more cars and associated heavy car preferences.

The combined built environment effect was a noticeably higher vehicle energy consumption in non-urban areas. Building new residences in existing cities could therefore have a stronger effect on future energy consumption and CO₂ emissions than indicated by earlier studies that did not take vehicle energy efficiency into account.

The capturing of this indirect effect was made possible by the multilevel discrete choice modeling framework. In addition, this design allowed analysis of households' car fuel and weight preferences separately. This improved the model accuracy and increased its future relevance. The model was fed with built environment data from multiple sources on the fine-grained postcode-6 level. This helped achieve a high degree of accuracy and allowed testing of a wide variety of built environment variables. Moreover, the real-world car energy efficiency could be precisely determined by coupling data from the Netherlands Vehicle Authority, TNO, and Travelcard BV.

This also illuminated the large gap between vehicles' real world energy consumption and the official data based on the NEDC test. Further analysis showed this gap to be greater than the built environment effect on vehicle energy use. Potential consequences include undermining of CO₂ -emission standards, flawed estimates of (technologies') emission reduction potentials, and the misleading of consumers. It is therefore important that a new WLTP test-cycle has recently been introduced, which should reduce - but not eliminate - the real-world gap (Ligterink et al., 2016).

4. CONCLUSIONS

This paper investigated how and to what extent features of residential neighborhoods and their residents are currently related to energy relevant car type choice by analyzing real-world energy use with a multilevel model of fuel- and weight-preferences in one- and multicar households.

Small, female households with few older members, and a lower income in non-green (urban) environments were most likely to own light, efficient vehicles. Households with a private parking spot tended to own both heavy and electric vehicles. Small, lower-income, urban households were lastly less likely to own one or multiple cars, whereby the ownership of multiple cars was associated with the choice of heavier vehicles.

The combined effect was a mild preference for efficient, low-energy vehicles in urban environments. Earlier studies that focused on vehicle kilometers thus underestimated the environmental impact of urban planning interventions. However, the easiest way to reduce vehicles' energy consumption and CO₂ -emissions seems to further improve the testing procedures in order to tighten policies, stimulate innovation, and better inform consumers.

Table 2: The estimated coefficients for each of the utility functions. Variables with a P-value of 5% or less have been made bold. Coefficients give the change in utility when increasing the variable by 1 standard deviation (std).

Car ownership classification									
Onecar class utility	Coef β_c	Std err	t-score	P-value	Twocar class utility	Coef β_c	Std err	t-score	P-value
Aspecific Constant	2.004	0.060	33.2	0.000	Aspecific Constant	0.763	0.076	10.1	0.000
HH_under18	0.435	0.066	6.6	0.000	HH_under18	0.533	0.071	7.5	0.000
HH_18to39	0.571	0.072	8.0	0.000	HH_18to39	1.571	0.084	18.8	0.000
HH_40to59	0.598	0.072	8.3	0.000	HH_40to59	1.579	0.084	18.8	0.000
HH_60plus	1.160	0.078	15.0	0.000	HH_60plus	2.032	0.107	19.1	0.000
inc<20k	-0.292	0.036	-8.2	0.000	inc<20k	-0.384	0.082	-4.7	0.000
inc40to60k	0.196	0.045	4.3	0.000	inc40to60k	0.304	0.061	5.0	0.000
inc60to120k	0.313	0.062	5.0	0.000	inc60to120k	0.613	0.070	8.7	0.000
inc≥120k	0.155	0.078	2.0	0.047	inc≥120k	0.252	0.080	3.1	0.002
Workers	0.375	0.050	7.5	0.000	Males	0.164	0.057	2.9	0.004
Density_PC5	-0.307	0.041	-7.4	0.000	Workers	1.052	0.076	13.8	0.000
log_km_station	0.165	0.049	3.4	0.001	Density_PC5	-0.705	0.075	-9.4	0.000
log_km_bigstation	0.125	0.050	2.5	0.012	log_km_station	0.272	0.063	4.3	0.000
km_hugecenter	0.162	0.043	3.8	0.000	log_km_bigstation	0.177	0.069	2.6	0.010
km_bus	0.246	0.085	2.9	0.004	km_hugecenter	0.142	0.058	2.4	0.015
Parkingspot	0.252	0.046	5.5	0.000	km_bus	0.350	0.090	3.9	0.000
					Parkingspot	0.398	0.059	6.8	0.000

Fuel- and weight-based car types									
Aspecific Constants	ASC_t	Std err	t-score	P-value	Standard fuel type utility	Coef β_f	Std err	t-score	P-value
Standard_light	0.344	0.166	2.1	0.039	2car Coefficient (β_{2car})	-0.174	0.267	-0.7	0.514
Standard_midlight	1.116	0.135	8.3	0.000	HH_18to39	0.108	0.067	1.6	0.105
Standard_heavy	-1.961	0.302	-6.5	0.000	inc<20k	0.221	0.077	2.9	0.004
Diesel_midlight	-1.230	0.168	-7.3	0.000	FracAdultHighedu	-0.071	0.049	-1.5	0.147
Diesel_midheavy	-0.951	0.170	-5.6	0.000	FracAdultMales	-0.166	0.057	-2.9	0.004
Diesel_heavy	-1.428	0.227	-6.3	0.000	km_hugecenter	-0.228	0.073	-3.1	0.002
HEV	-1.390	0.227	-6.1	0.000	km_largecenter	0.149	0.053	2.8	0.005
Diesel fuel type utility	Coef β_f	Std err	t-score	P-value	HEV fuel type utility	Coef β_f	Std err	t-score	P-value
2car Coefficient (β_{2car})	-2.656	0.444	-6.0	0.000	2car Coefficient (β_{2car})	-0.328	0.383	-0.9	0.391
HH_18to39	0.699	0.111	6.3	0.000	HH_60plus	-0.216	0.116	-1.9	0.064
HH_40to59	0.463	0.097	4.8	0.000	inc≥120k	0.067	0.046	1.5	0.146
Workers	0.396	0.099	4.0	0.000	Higheducated	0.229	0.097	2.4	0.018
km_largecenter	0.224	0.074	3.0	0.003	Parkingspot	0.234	0.096	2.4	0.015
Landuse	-0.121	0.068	-1.8	0.076					
Parkingspot	0.155	0.081	1.9	0.055					
Light weight utility	Coef β_w	Std err	t-score	P-value	Midlight weight utility	Coef β_w	Std err	t-score	P-value
2car Coefficient (β_{2car})	-0.609	0.324	-1.9	0.060	2car Coefficient (β_{2car})	-1.552	0.258	-6.0	0.000
HH_under18	-0.372	0.064	-5.8	0.000	HH_under18	-0.220	0.052	-4.3	0.000
HH_40to59	-0.169	0.066	-2.5	0.011	HH_60plus	-0.249	0.073	-3.4	0.001
HH_60plus	-0.661	0.085	-7.8	0.000	inc≥120k	-0.092	0.045	-2.1	0.040
inc40to60k	-0.116	0.058	-2.0	0.044	Workers	0.236	0.068	3.5	0.001
inc60to120k	-0.149	0.064	-2.3	0.020	km_hugecenter	0.123	0.076	1.6	0.109
inc≥120k	-0.198	0.078	-2.5	0.011					
Males	-0.355	0.065	-5.5	0.000					
km_hugecenter	0.201	0.088	2.3	0.022					
NDVI	-0.125	0.061	-2.1	0.039					
Parkingspot	-0.096	0.058	-1.7	0.097					
Midheavy weight utility	Coef β_w	Std err	t-score	P-value	Heavy weight utility	Coef β_w	Std err	t-score	P-value
2car Coefficient (β_{2car})	-1.864	0.567	-3.3	0.001	2car Coefficient (β_{2car})	1.194	0.280	4.3	0.000
HH_40to59	0.185	0.112	1.7	0.098	HH_under18	0.141	0.057	2.5	0.014
inc60to120k	0.356	0.083	4.3	0.000	HH_18to39	-0.309	0.085	-3.6	0.000
Workers	0.267	0.146	1.8	0.067	HH_60plus	0.190	0.090	2.1	0.035
NDVI	0.297	0.110	2.7	0.007	Higheducated	0.113	0.070	1.6	0.103
					Males	0.331	0.098	3.4	0.001
					Parkingspot	0.212	0.075	2.8	0.005

ACKNOWLEDGEMENTS

For this paper, we made use of data from the Netherlands Mobility Panel administered by the KiM. We would especially like to thank KiM researcher Mathijs de Haas for his continuous support. Moreover, we used recent research by the TNO to compute the specific energy consumption of the vehicles in the dataset. We were thereby assisted by TNO-researcher van Gijlswijk. We would like to thank Job Hoenderdos for helping us scrape the Travelcard data from Praktijkverbruik.nl. Furthermore, we want to thank Matthieu Brinkhuis for his input and ideas. We lastly want to thank Bogdan Kapatsila and Santiago Cardona Urrea for their constructive feedback.

REFERENCES

- Banister, D. (2011). Cities, mobility and climate change. *Journal of Transport Geography*, 19:1538–1546.
- Bhat, C. R., Sen, S., and Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B*, 43:1–18.
- Bierlaire, M. (2020). A short introduction to pandasbiogeme. technical report transp-or 200605. Technical report, Transport and Mobility Laboratory, ENAC, EPFL.
- Brownstone, D. and Fang, H. A. (2014). A vehicle ownership and utilization choice model with endogenous residential density. *The Journal of Transport and Land Use*, 7(2):135–151.
- Cao, X., Mokhtarian, P. L., and Handy, S. L. (2006). Neighborhood design and vehicle type choice: Evidence from northern California. *Transportation Research Part D*, 11:133–145.
- Chen, N., Akar, G., Gordon, S. I., and Chen, S. (2021). Where do you live and what do you drive: Built-environmental and spatial effects on vehicle type choice and vehicle use. *International Journal of Sustainable Transportation*, 15(6):444–455.
- de Ruiter, J., van Gijlswijk, R., and Ligterink, N. E. (2021). Model m1: estimation of the real-world co2 emissions of passenger cars based on vehicle properties - version 2. Technical report, Netherlands Organisation for Applied Scientific Research, The Hague.
- Eluru, N., Bhat, C., Pendyala, R., and Konduri, K. (2010). A joint flexible econometric model system of household residential location and vehicle fleet composition/usage choices. *Transportation*, 37:603–626.
- Ewing, R. and Cervero, R. (2010). Travel and the built environment - A meta-analysis. *Journal of the American Planning Association*, 76:265–294.
- Galvin, R. (2022). Are electric vehicles getting too big and heavy? modelling future vehicle journeying demand on a decarbonized us electricity grid. *Energy Policy*, 161(112746).
- Garikapati, V. M., Sidharthan, R., Pendyala, R. M., and Bhat, C. R. (2014). Characterizing household vehicle fleet composition and count by type in integrated modeling framework. *Transportation Research Record*, 2429.

Hoogendoorn-Lanser, S. (2019). Mobiliteitspanel Nederland (MPN2019). KiM Netherlands Institute for Transport Policy Analysis. DOI: 10.17026/dans-zyc-7qfv [dataset].

Hoogendoorn-Lanser, S., Schaap, N. T. W., and OldeKalter, M.-J. (2015). The netherlands mobility panel: An innovative design approach for web-based longitudinal travel data collection. 10th International Conference on Transport Survey Methods, *Transportation Research Procedia*, 11:311–329.

International Energy Agency (2022). The role of critical world energy outlook special report minerals in clean energy transitions - executive summary. Technical report, 9 rue de la Fédération 75739 Paris Cedex 15 France.

Ligterink, N. E., van Mensch, P., and Cuelenaere, R. F. A. (2016). NEDC - WLTP comparative testing. Technical report, Netherlands Organisation for Applied Scientific Research (TNO).

Liu, C. and Shen, Q. (2011). An empirical analysis of the influence of urban form on household travel and energy consumption. *Computers, Environment and Urban Systems*, 35(5):347–357.

McCarthy, P. S. and Tay, R. S. (1998). New vehicle consumption and fuel efficiency: a nested logit approach. *Transport Research E: Logistics and Transportation Review*, 34:39–51.

Næss, P. (2012). Urban form and travel behavior: Experience from a Nordic context. *Journal of Transport and Land Use*, 5(2):21–45.

Newman, P. W. G. and Kenworthy, J. R. (1989). Gasoline consumption and cities: A comparison of U.S. cities with a global survey. *Journal of the American Planning Association*, 55(1):24–37.

Otten, M., 't Hoen, M., and den Boer, E. (2015). STREAM personenvervoer 2014 - Studie naar TRansportEmissies van Alle Modaliteiten Emissiekentallen 2011. Technical report, CE Delft, Delft.

Pandas development team, T. (2020). *pandas-dev/pandas: Pandas*.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Potoglou, D. (2008). Vehicle-type choice and neighbourhood characteristics: An empirical study of hamilton, canada. *Transportation Research Part D*, 13:177–186.

Prieto, M. and Caemmerer, B. (2013). An exploration of factors influencing car purchasing decisions. *International Journal of Retail & Distribution Management*, 41(10).

Ren, X., Wang, Z., Nast, C., Ettema, D., and Brombacher, A. (2019). Integrating industrial design and geoscience: A survey on data-driven research to promote public health and vitality. *ACM International Conference Proceeding Series*, pages 91–95.

Silva, M., Oliveira, V., and Leal, V. (2017). Urban form and energy demand: A review of energy-relevant urban attributes. *Journal of Planning Literature*, 32(4):346–365.

Song, S., Diao, M., and Feng, C.-C. (2016). Individual transport emissions and the built environment: A structural equation modelling approach.

Statistics Netherlands (2018). Kerncijfers per postcode - Naar volledige postcode (PC6), 2018, v3. Retrieved from: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode>. [Dataset].

Statistics Netherlands (2019). Kerncijfers per postcode - Naar postcode-5 (PC5), 2019, v3. Retrieved from: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode>. [Dataset].

Stevens, M. R. (2017). Does compact development make people drive less? *Journal of the American Planning Association*, 83(1):7–18. 19

Team Open Data RDW (2021). Open data rdw: Gekentekende_voertuigen_brandstof. URL: https://opendata.rdw.nl/Voertuigen/Open-Data-RDW-Gekentekende_voertuigen_brandstof/8ys7-d773 [dataset]. Downloaded 16 August 2021.

Timmons, D. and Perumal, A. (2016). Us vehicle fuel-efficiency choices: demographic, behavioral, and cultural factors. *Journal of Environmental Planning and Management*, 59:2179–2197.

van Gijlswijk, R., Paalvast, M., Ligterink, N. E., and Smokers, R. (2020). Real-world fuel consumption of passenger cars and light commercial vehicles. Technical report, Netherlands Organisation for Applied Scientific Research, The Hague.

Wang, Z. (2020). Vitality Data Center - Environmental factors. Retrieved from: https://github.com/vitality-data-center/environmental_factors. [Dataset].

Wang, Z., Ettema, D., and Helbich, M. (2021). Objective environmental exposures correlate differently with recreational and transportation walking: A cross-sectional national study in the Netherlands. *Environmental Research*, 194(110591).

Weiss, M., Cloos, C. K., and Helmers, E. (2020). Energy efficiency trade-offs in small to large electric vehicles. *Environmental Sciences Europe*, 32(46).

Wes McKinney (2010). Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61.

Westin, K., Jansson, J., and Nordlund, A. (2018). The importance of socio-demographic characteristics, geographic setting, and attitudes for adoption of electric vehicles in Sweden. *Travel Behaviour and Society*, 13:118–127.

Zijlstra, J. and Rietkerk, M. (2020). Methodiek berekening CO₂-emissiefactoren personenauto's. Technical report, Milieu Centraal, Utrecht

Who is Responsible for the Externalities Produced by Freight Carriers? Hint: The Answer is Not as Simple as it Seems...

José Holguín-Veras, Diana Ramírez-Ríos, and Trilce Encarnación

William H. Hart Professor, Civil and Environmental Engineering, Rensselaer Polytechnic Institute,
110 8th St, Room JEC 4030. Troy, NY 12180, USA, Email: jhv@rpi.edu

Department of Industrial and Systems Engineering, University at Buffalo, The State University of
New York, Bell Hall North Campus, Buffalo, NY 14260, USA Email: dgramire@buffalo.edu

Supply Chain Management and Analytics, College of Business Administration, University of
Missouri, 206 ESH, St. Louis, MO 63121, USA, Email: tencarnacion@umsl.edu

Abstract

This research sheds light into an important and overlooked aspect of urban freight management and planning: the impacts of the decisions made by shippers, receivers, transportation and land use agencies, the real estate sector and other agents—referred to as non-carrier agents, or NCAs—on the generation of freight externalities. The paper is based on the insight that, since freight carriers must meet the constraints set by these agents, NCAs' decisions could force the carriers to create externalities above and beyond those that the carriers would generate if they had complete control over their operations. As part of the research, the authors: identify a number of NCAs' decisions that could negatively affect the operational performance and the externalities produced by freight carriers; and compute the corresponding Shapley Values to allocate the responsibilities for the freight externalities among carriers and NCAs using numerical experiments based on real-life supply chains. The insights gained are used to identify policy implications related to electrification of the trucking sector.

Keywords: Urban Freight Management, Externalities, City Logistics, Shapley Value

1. Introduction and Background

The mitigation, or elimination, of the externalities produced by the various forms of transportation activity is one of the most important objectives of transportation planning and policy. Central to this quest's success is the correct identification of the root decisions that create externalities, as such an understanding provides crucial insight into how to address these effects. This is important because NCAs make decisions that directly influence carrier operations and could create externalities. The role played by NCAs in the generation of freight transportation externalities is frequently overlooked. At first glance, it seems logical to conclude that, since freight vehicles are the ones that physically produce the externalities, the carriers are solely responsible for the generation of the externalities. However logical this perception may be, it flies on the face of the complex web of decisions that influence supply chains and the resulting freight carrier activity.

To start, it is important to highlight that the agents involved in supply chains do not operate in isolation of each other, and freight operations do not take place in a vacuum. Freight carrier operations they are the result of the interactions between the economic agents that participate in the numerous production and consumption stages in supply chains. At each of these production-consumption links, an agent produces and/or sends supplies (the shipper) that are then consumed by a different agent (the receiver), after they are transported by the carrier. Because of these tight interconnections, no single agent can make unilateral changes without impacting in one way or another the other participants in the supply chain. As a result, in cases of conflicting preferences, the agent with more power tends to impose its will on the others. This cold logic of power explains why the carriers—frequently the weakest agent in supply chains (Holguín-Veras et al. 2015)—have to abide by the decisions and preferences of the more powerful NCAs. If the carriers do not follow the NCAs' instructions, they run the risk of being fired or fined.

Supply chains do not take place in a vacuum either. They are embedded in the fabric of rural/urban/suburban areas and have to abide by the regulations enacted by transportation, environmental, and land use agencies. These agencies typically control: (1) where manufacturers, distribution centers, truck terminals, retail stores, and other participants in supply chains are located; (2) the size and nature of the activities performed at these locations; (3) access to transportation networks by time of day and facility type, and to public spaces such as the curbsides and sidewalks; and (4) the environmental standards to be met by vehicles; among other impactful constraints. Complicating matters, with the advent of e-commerce, households became another active agent in supply chains. As a result, untold numbers of deliveries are made to buildings unprepared to receive supplies and parcels, forcing carriers to park on local streets aggravating urban congestion. Throughout the paper, the term “receivers” denotes both commercial establishments and households. The real estate sector establishes constraints—setting delivery time windows and building hours, determining whether or not the building design is conducive to efficient freight operations—that impact carrier operations. In addition, the markets in which the various segments of the freight industry operate determine the profitability of carriers’ operations and, consequently, the ability of carriers to purchase environmentally friendly vehicles.

At first sight, in game theory terms, this situation seems to be similar to a multi-layer, multi-player Stackelberg game where the leading players make decisions that maximize their returns, while the follower agents are forced to make decisions under the constraints established by the leaders. However, close inspection reveals that these interactions are not a Stackelberg game, because the agents involved may have the option to cooperate with the others. In a separate publication, (Holguín Veras et al. 2023) defined the “Supply Chain Game” as one where the agents are interested in participating in the supply chain (if they are not, the supply chains would not exist). However, the conditions of their participation depends on the balance of power with the participating agents. For each interaction with another agent, they are only three possibilities: submit to the wish of a more powerful agent, cooperate with the agent if this outcome is the best for both agents, or impose its will on a weaker agent. (Holguín Veras et al. 2023) established that the Supply Chain Game leads to two outcomes: a state of natural cooperation, and the “Battle of the Sexes” game.

These insights have major implications. Most significantly, since the actions by NCAs to influence carrier operations could increase the externalities produced by the carrier—above and beyond the externalities that would be generated without the NCAs’ interventions—NCAs would be responsible for the incremental externalities produced. To eliminate freight externalities, transportation planners must target all the agents—carriers and NCAs—involved in the decisions that contribute to the creation of the externalities. It is crucial to correctly identify which agents are responsible for the externalities in question, to be able to identify the public-sector initiatives needed to induce NCAs to modify the behaviors that create the externalities in question. These considerations take an existential importance in the era of climate change, where all policy weapons ought to be brought to bear to reduce global warming gases. Solely focusing on carriers is bound to be ineffective at best, and counterproductive at worst.

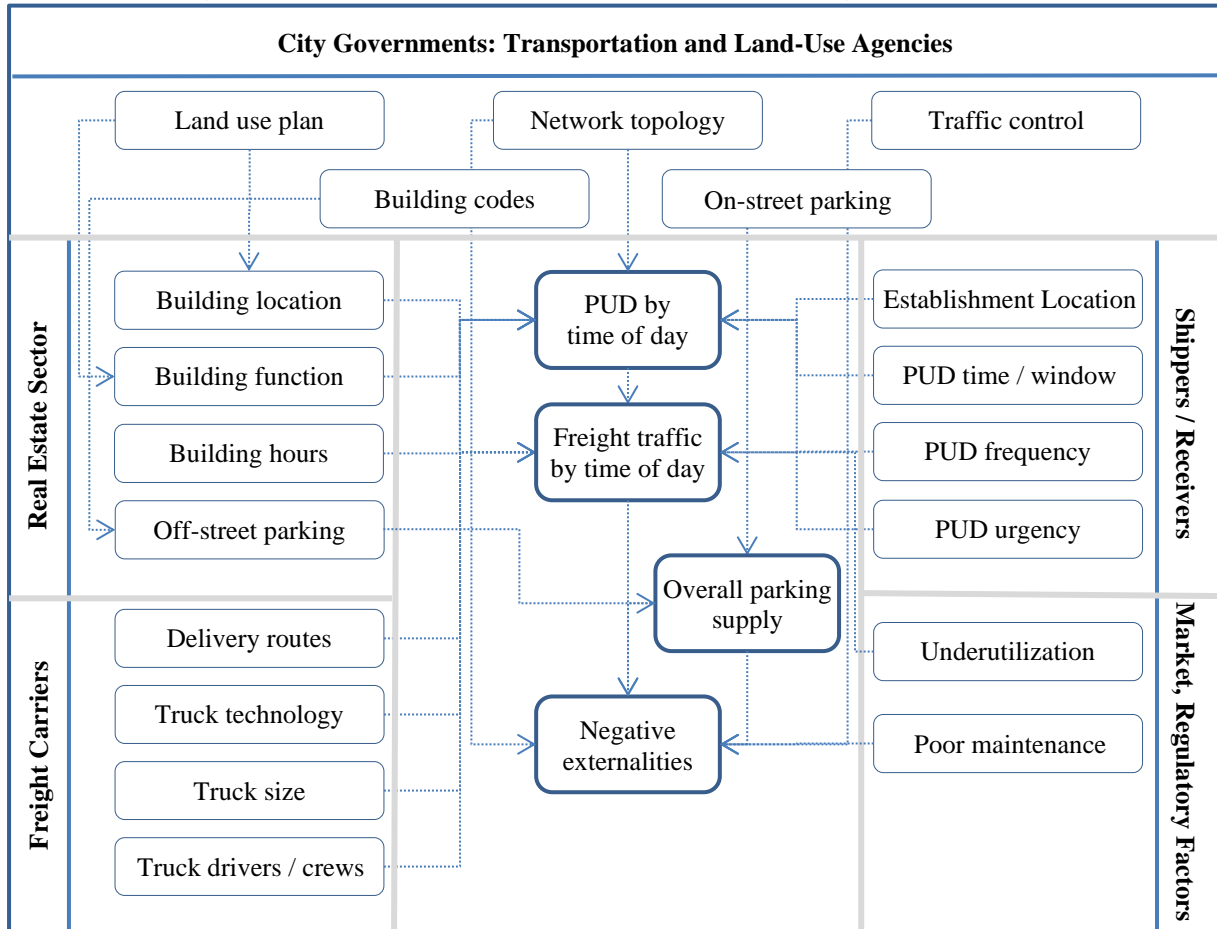
The main objective of this paper is to help fill an important gap in the research literature regarding the identification and quantification of the responsibilities of carriers and NCAs in the generation of urban freight externalities. In doing so, the paper significantly expands the literature on urban freight management and planning by considering the role played by NCAs, and develops a methodology to quantify the responsibilities of the NCAs for the externalities they help create. To this effect, the authors establish how NCAs influence carrier operations, discuss NCAs’ roles in the generation of freight externalities; and conducts numerical experiments based on real-life supply chains to provide insight into the relative importance of the roles played by various agents. The paper concludes with a summary of the key insights developed during the research.

2. Contributors of Freight Externalities

This section seeks to illustrate the NCAs’ decisions that influence carrier operations and externalities. As a result, the section does not comprehensively discuss the array of factors these agents consider when making other business decisions. Figure 1 shows some key ways in which the different agents influence freight activity and create urban freight externalities. As shown, five groups

of influencing factors are considered. Hinting at the complexity of the processes that create freight externalities, only three of them—freight carriers, shippers, and receivers (both commercial and households)—are directly involved in the supply chain. The rest of the agents establish the geographic and economic environments where supply chains take place and, in doing so, have a direct effect on the externalities produced.

Figure 1: Contributors to Traffic Related Urban Freight Externalities



Note: FTG refers to “freight trip generation”, and PUD to “pick-up and deliveries”.

Moreover, to exemplify the impacts of NCAs’ decisions, the authors analyzed the impacts of a number of decisions that have major impacts on freight externalities: (1) shipment size and frequency; (2) time of delivery; (3) allocation of curbside space for loading and unloading of deliveries (both off-street and on-street); and (4) location of logistical facilities. Strikingly, Table 1 makes clear that these decisions involve tradeoffs between the efficiency of freight activity and the benefits of some kind to the private-sector NCAs; and/or to influential stakeholders, such as passenger car drivers, in the case of public-sector NCAs. As hinted by Figure 1 and Table 1, using public policy levers to remove the most impactful constraints imposed on the carriers by the NCAs is bound to be effective in reducing freight externalities. As demonstrated by the Off-Hour-Deliveries (OHD) program, inducing receivers to accept deliveries during the 7PM to 6AM period led to emission reductions in the range of 50% to 67% (Holguín-Veras et al. 2018b). Inducing other NCAs to change their behaviors could be equally impactful.

Table 1: Summary of Impacts for Typical Decisions

Decision	Who Pays?	Who Benefits?
Shippers / receivers: Reduce shipment size, increase frequency of PUDs	Carrier: Increased costs / VKT/ parking	Receiver: Reduced inventory costs and the size of storage areas
	City: Increase infrastructure maintenance	
	Society: Increased congestion / health	
	Environment: Increased emissions	
Receivers: Require PUDs at the most congested times of day	Carrier: Increased costs / VKT / parking fines	Receiver: Continue following traditional practices, convenience
	City: Increase traffic control costs	
	Society: Increased congestion, health	
	Environment: Increased emissions	
Real Estate: Not to provide sufficient off-street loading docks for PUDs	Carrier: Increased operational costs / parking fines / VKT looking for parking	Real Estate: Increased amounts of space to commercialize
	City: Have to provide PUD on-street	
	Society: Increased congestion / health	
	Environment: Increased emissions	
City agencies: Not to provide sufficient on-street parking for PUDs	Carrier: Increased costs / parking fines / VKT looking for parking	City: Increased curbside space available for buses, passenger cars, and
	Society: Increased congestion, health	
	Environment: Increased in emissions	
Land use agencies: Not to provide space for urban/near-urban logistical facilities	Carrier: Increased costs / VKT	City: More land available for other uses
	City: Increase infrastructure maintenance	
	Society: Increased congestion, health	
	Environment: Increased emissions	

Note: “City” refers to “city agencies,” “health impacts” refer to those associated with emissions; VKT denotes “vehicle-kilometers-traveled,” and “PUD” refers to “pick-up and deliveries”.

3. Quantification of Responsibilities and Experimental Setup

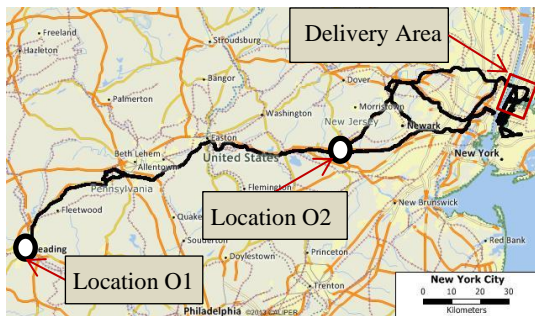
To quantify the responsibilities among carriers and NCAs, the authors used cooperative game theory to allocate the social costs of making deliveries. Cooperative game theory study the formation of coalitions to determine the payoffs that should be given to the partners to ensure the coalition is stable. Although conceived as a mechanism to optimally distribute earnings in profit making ventures, cooperative game theory is well-suited to allocate the responsibilities for freight externalities. Among the analytical formulations developed to compute these payoffs—*Stable Set* (Von Neumann and Morgenstern 1944), *Core of the Game* (Gillies 1959), *Aumann-Maschler Value* (Aumann and Maschler 1961), the *Kernel Value* (Davis and Maschler 1963), and others—the Shapley Value (Shapley 1952) stands out on account of its axioms that, once defined in terms of the economic value of externalities (instead of total profits) become: (1) the summation of the individual allocations equal to the total economic value of the externalities (efficiency principle); and (2) the individual allocations to each agent are proportional to their contributions to the creation of externalities (individual rationality principle). Mathematically speaking, the Shapley Value for agent i , ϕ_i , shown in equation (1), is the expected value of the value of the externalities, $v(S_i)$, produced by the potential coalitions where agent i is involved. See equation (1):

$$\phi_i = \sum_{\{S \in N: i \in S\}} \frac{(S-1)!(n-S)!}{n!} (v(S) - v(S - \{i\})) \quad (1)$$

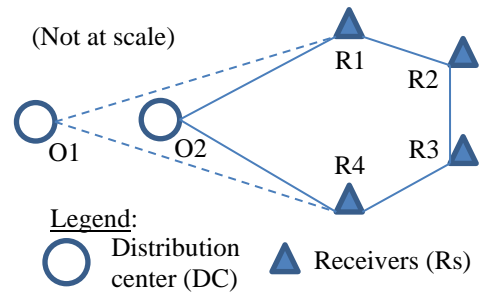
The experiments focus on the impacts of three important decisions made by NCAs: (1) location of a distribution center (DC), typically made by shippers under the constraints set by the real estate sector

and city agencies; (2) availability of parking at the receivers' locations, a decision of the real estate sector (off-street) and city agencies (on-street); and, (3) the time of delivery, a decision of the receivers. In this context, a shipper S operates a distribution center (DC) with two possible locations ($O1, O2$), a carrier C that makes deliveries to four receivers (R_1, R_2, R_3, R_4). There is a city government and real estate sector, $CG-RE$, that control the parking allocation to carriers. Thus, the game has four players ($S, C, R, CG-RE$). The carrier chooses the delivery route, under the operational constraints set by the NCAs. The shipper decides on the location of the DC, the $CG-RE$ decides on the availability of parking, while the receivers (assumed to make the same decision) decide on time of the delivery. To ground the numerical experiments in reality, the authors selected a delivery route (conducted in both regular and off-hours) from the New York City Off-Hour Delivery project (Holguín-Veras et al. 2018a). The actual delivery route (pane a in Figure 2) was the base for the idealized delivery routes (pane b in Figure 2) used in the experiments. The parameters used in the numerical experiments come from the GPS data collected in New York City.

Figure 2: Idealized Delivery Routes Considered



Pane (a): Actual Delivery Route



Pane (b): Idealized Delivery Routes

The marginal social costs—the sum of private and external costs—were estimated as the difference between the preferred scenarios for each of the NCAs and the carrier preferred scenario, (i.e., shortest travel distance to the delivery area, off-hour deliveries, and ready access to parking). More often than not, carriers and NCAs have different definitions of their preferred conditions. In equality of conditions: the shipper is interested in low land costs, frequently far from city centers; receivers tend to prefer regular hour deliveries; while city governments / real estate sector, for various reasons, do not provide the parking needed by carriers, and are not inclined to allow DCs close to urban cores.

The private costs account for the carrier's travel time, the distance travelled to make the delivery, and the expected value of parking fines. The external costs consider the costs of emission of various pollutants ($CO_2, NO_x, PM_{2.5}, PM_{10}$). See Tables 2 and 3. An important effect not considered is the increase in congestion and emissions produced by freight vehicles on the rest of the traffic stream and the blocking effects in urban streets. Thus, the marginal social costs created by receivers, and city governments / real estate sector are underestimated.

Table 2: Private Transportation Costs

Transportation costs	Time of Day	
	Regular Hours	Off Hours
Highway average travel speed (km/h)	53.24	57.75
City street average travel speed (km/h)	11.30	19.38
Cost per unit distance (\$/km)	\$0.61	\$0.61
Cost per unit time (\$/hour)	\$45.13	\$45.13
Average time between stops (min/stop)	66.83	49.00
Time cruising for parking (minutes/PUD)	5-10	0.00
Parking fine (\$/PUD)	\$68.91	0

Table 3: Emission Rates and Valuation of Pollutants

Pollutant	Emission rates		Economic Valuation of Pollutants		
	Regular Hours	Off Hours	Rate (original)	Rate (US\$ in 2020)	Source
CO2 (g/km)	1548.16	694.45	\$0.49/kg	\$0.58/kg	Stern Review (UK, 2006)
NOX (g/km)	0.581	0.236	£6,199/ton	\$5.61/kg	Ricardo Energy and Environment (2017)
PM10 (g/km)	0.086346	0.029762	£105,836/ton	\$95.79/kg	Ricardo Energy and Environment (2017)
PM25 (g/km)	0.083	0.028	\$12.85/kg	\$22.16/kg	Small and Kazimi (1995)

4. Discussion of Results and Conclusions

Table 4 shows the breakdown of the social cost for the carrier preferred scenario—i.e., DC close to the delivery area, off-hour deliveries, and parking available—and the Shapley Values quantifying the increases in social costs of making deliveries attributed to NCAs’ decisions. The contributions of shippers and carriers are reported together because their contributions are interwoven (Friesz and Morlok 1980, Friesz et al. 1986). A similar situation is that of city governments / real estate sector that jointly determine freight parking availability. The experiments consider two scenarios of cruising times for parking. See Table 4.

Table 4: Social Costs and Shapley Values

A) Carrier Preferred Scenario						
Agent:	Social Costs		Private Costs		Externalities	
Shipper-Carrier	\$426.42		\$367.61		\$58.81	
B) Shapley Values (Expected Values of the Increases in Social Costs)						
Case 1: Cruising for Parking = 5 min						
Agent:	Social Costs		Private Costs		Externalities	
Shipper-Carrier	\$309.37	54.32%	\$212.26	62.13%	\$97.11	42.62%
Receivers	\$201.45	35.37%	\$90.98	26.63%	\$110.47	48.48%
City Gov. /Real Estate Sector	\$58.70	10.31%	\$38.41	11.24%	\$20.29	8.90%
Total	\$569.51	100%	\$341.65	100%	\$227.86	100%
Increase = B/A	133.56%		92.94%		387.46%	
Case 2: Cruising for Parking = 10 min						
Agent:	Social Costs		Private Costs		Externalities	
Shipper-Carrier	\$309.37	49.83%	\$212.26	57.06%	\$97.11	39.02%
Receivers	\$200.71	32.33%	\$89.54	24.07%	\$111.16	44.67%
City Gov. /Real Estate Sector	\$110.77	17.84%	\$70.19	18.87%	\$40.58	16.31%
Total	\$620.84	100%	\$371.99	100%	\$248.84	100%
Increase = B/A	145.59%		101.19%		423.13%	

These results compellingly demonstrate the significant effects of NCAs’ decisions, found to increase emission externalities by 387%-423%, while increasing private costs by 92%-101%. As shown, the majority of the freight externalities, 57-61% of the total, are the result of the decisions made by receivers (45-48%), and city governments / real estate sector (9-16%). The duplet shipper-carrier contribute the rest (39-43%). It should be noted that the contribution of the city governments / real estate sector to the freight externalities is underestimated as these estimates do not include the increase in emissions from the traffic stream resulting from the trucks traveling in congested networks, cruising for parking in narrow streets, and double-parked trucks.

These results, based on real-life data, have major implications on urban freight management on account to the practical impossibility of rapidly electrification of trucking (at best a medium-term possibility). In the short- and medium-terms, freight demand management—fostering off-hour deliveries, staggered deliveries, and the like—remain the best short-term option. Fostering freight-efficient land-uses (Holguin-Veras et al. 2021) is bound to play a key role in the medium and long-terms complementing electrification. Without doubt, using all these tools provide the best chance of mitigating the worst effects of climate change.

5. References

- Aumann, R. J. and M. Maschler (1961). The Bargaining Set for Cooperative Games. Research Memorandum No. 34. Princeton, NJ, Princeton University Press: 1-35.
- Davis, M. and M. Maschler (1963). "Existence of Stable Payoff Configurations for Cooperative Games: Abstract." Bull. Amer. Math. Soc. **69**: 106–108.
- Friesz, T. L., J. A. Gottfried and E. K. Morlok (1986). "A Sequential Shipper-Carrier Network Model for Predicting Freight Flows." Transportation Science **20**(2): 80-91.
- Friesz, T. L. and E. K. Morlok (1980). Recent Advances in Network Modeling and their Implications for Freight Systems Planning. Proceedings of Transportation Research Forum 21.
- Gillies, D. B. (1959). Solutions to General Non-Zero Sum Games. Contributions to the Theory of Games (AM-40). A. M. Tucker and R. D. Luce. Cambridge, MA, Princeton University Press. **4**: 47-86.
- Holguín-Veras, J., F. Aros-Vera and M. Browne (2015). "Agent Interactions and the Response of Supply Chains to Pricing and Incentives." Economics of Transportation.
- Holguín-Veras, J., T. Encarnación, C. A. González-Calderón, J. Winebrake, S. Kyle, N. Herazo-Padilla, L. Kalahasthi, W. Adarme, V. Cantillo, H. Yoshizaki and R. A. Garrido (2018a). "Direct Impacts of Off-Hour Deliveries on Urban Freight Emissions." Transportation Research Part D: Transport and Environment **61**(1): 84-103.
- Holguín-Veras, J., S. Hodge, J. Wojtowicz, C. Singh, M. Jaller, F. Aros-Vera, K. Ozbay, A. Weeks, M. Repogle, C. Ukegbu, J. Ban, M. Brom, S. Campbell, I. Sanchez-Díaz, C. González-Calderón, A. Kornhauser, M. Simon, S. McSherry, A. Rahman, T. Encarnación, X. Yang, D. Ramírez-Ríos, L. Kalahasthi, J. Amaya, M. Silas, B. Allen and B. Cruz (2018b). "The New York City Off-Hour Delivery Program: A Business and Community-Friendly Sustainability Program." INFORMS Journal on Applied Analytics **48**(1): 70-86.
- Holguín-Veras, J., D. Ramirez-Rios, J. Ng, J. Wojtowicz, D. Haake, C. T. Lawson, O. Calderón, B. Caron and C. Wang (2021). "Freight-Efficient Land Uses: Methodology, Strategies, and Tools." Sustainability **13**(6).
- Holguín Veras, J., T. E. Encarnación-Rivera and D. Ramírez-Rios (2023). "Urban Freight Externalities: Contributing Factors and Agent Responsibilities " (in review).
- Shapley, L. S. (1952). "A Value for n-Person Games." **295**: 1-13.
- Von Neumann, J. and O. Morgenstern (1944). Theory of Games and Economic Behavior. Princeton, NJ, Princeton University Press.

A new flexible and interpretable choice model with monotonicity constraints, non-linearity, and taste heterogeneity

Eui-Jin Kim*¹, Prateek Bansal²

¹ Assistant Professor, Department of Transportation System Engineering, Ajou University, South Korea

² Assistant Professor, Department of Civil and Environmental Engineering, National University of Singapore, Singapore (prateekb@nus.edu.sg)

SHORT SUMMARY

This study proposes a flexible and interpretable discrete choice model (DCM) capturing key behavioural mechanisms simultaneously: (i) interactions between alternative-specific and individual-specific attributes (e.g., taste heterogeneity), (ii) interactions between alternative-specific attributes, (iii) inherent non-linear utility of alternative-specific attributes (e.g., diminishing marginal utility of travel cost). Deep neural networks (DNNs) have been considered as candidates to flexibly capture these mechanisms, but they fail to provide trustworthy and explainable economic information (i.e., interpretability) obeying domain-specific knowledge (e.g., decrease in utility of travel mode due to an increase in its travel cost). We propose a DCM based on a lattice network (LN) that efficiently imposes attribute-specific monotonicity constraints in the utility specification while ensuring the trustworthy interpretation of DNNs. The proposed LN-based DCM is benchmarked against DNN in a Monte Carlo study. The results show that it outperforms even the parametric DCM in terms of interpretability while slightly underperforming the DNN in terms of predictability.

Keywords: discrete choice model; monotonicity; deep neural network; lattice network; interpretability.

1. INTRODUCTION

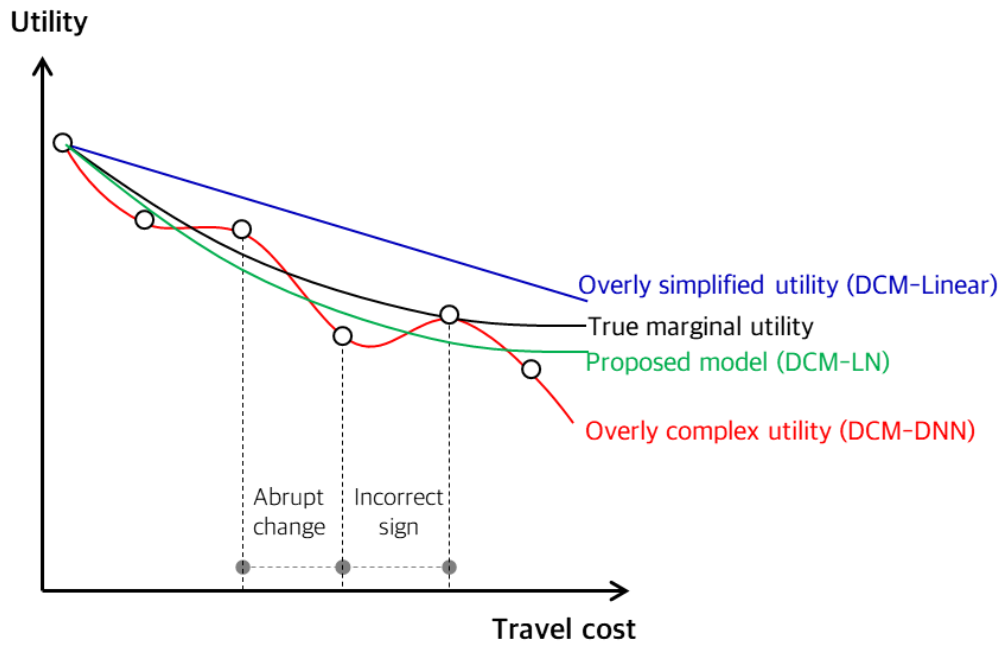
In discrete choice models (DCMs), correctly specifying the systematic utility is critical to achieving good predictability and interpretability. Interpretability indicates the extent to which it provides trustworthy and explainable economic information at an individual level. Ensuring the monotonicity of the utility relative to a subset of alternative-specific attributes is crucial to maintain interpretability. For instance, the utility should monotonically decrease with the increase in cost in most situations.

Traditional parametric DCMs rely on linear-in-parameter utility specifications, with hand-crafted interactions between attributes. Such models are appealing due to the ease of associating the meaning with parameter estimates. However, misspecifications of parametric utility not only result in poor prediction accuracy, but also biased parameter estimates for interaction effects, leading to counterintuitive willingness to pay (WTP) estimates (e.g., negative WTP to save travel time).

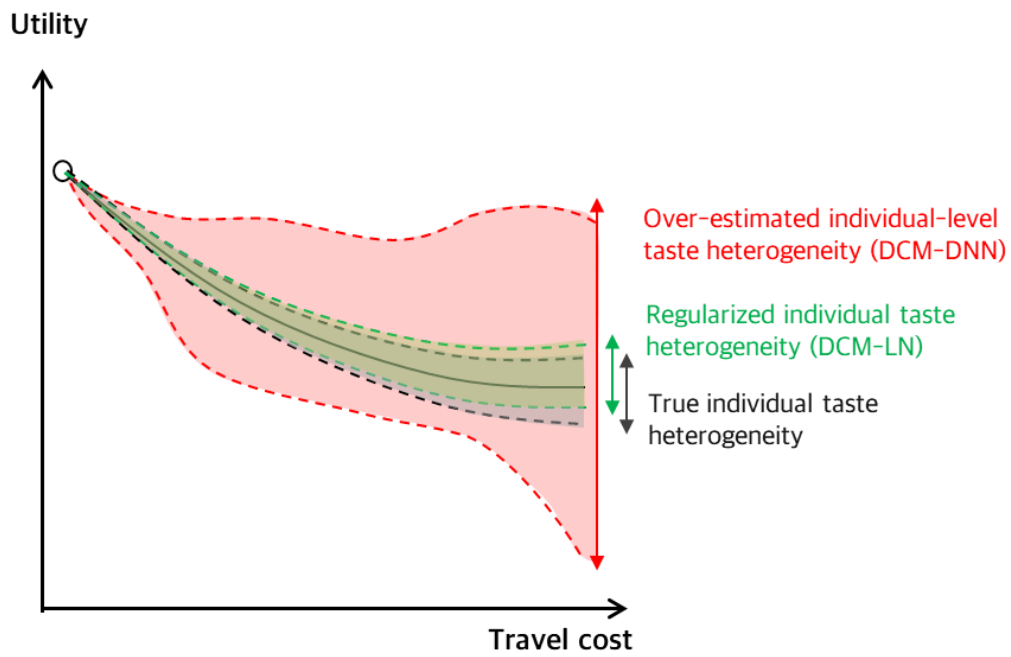
To address the issues of parametric DCMs, researchers have adopted deep neural networks (DNNs) (Cranenburgh et al., 2022). The DNNs improve the DCM's predictability by considering complex non-linear and interaction effects of attributes, and WTP and elasticity estimates can also be extracted (Wang, Wang, et al., 2020). While ensuring monotonicity requirements is challenging in DCM-DNN, ignoring it might lead to counterintuitive interpretations in certain attribute domains (Wang, Wang, et al., 2020).

This study contributes with a theory-constrained DCM where the systematic utility is specified using a lattice network (DCM-LN henceforth). First, the lattice network (LN) segments the input space into grids or cells. The input attribute vector is transformed into a vector of interpolation weights (i.e., model parameters to be estimated) over the vertices of the cell that represents the input space. Second, the function value is obtained as a linear transformation of the first step’s interpolation weights (Gupta et al., 2016). While the linear transformation in the second step leads to easy-to-implement theoretical conditions for monotonicity, the transformation of the input attribute vector into the first step captures non-linear effects and interactions of attributes. We also add a calibration layer before and after the lattice network to improve the ability of DCM-LN to capture non-linearities in attribute-specific effects, obviating the need to create a fine-grained lattice that requires much more model parameters. DCM-LN can thus simultaneously infer underlying non-linear effects of all alternative- and individual-specific attributes and interactions between them while achieving the monotonic effect of a subset of attributes for every individual.

Figure 1 symbolically benchmarks the systematic utility of the DCM-LN against the traditional DCM with linear utility specification without interactions (DCM-linear) and DCM-DNN at a population and an individual level. Figure 1(a) shows that the overly complex DCM-DNN model represents abrupt changes in marginal utility and even incorrect signs. Such behavioural irrationalities are even worse at an individual level, as indicated by potentially higher heterogeneity (see Figure 1b). On the other hand, the overly simplified DCM-linear model causes serious bias in the marginal utility estimates. In contrast, the DCM-LN can recover true marginal utilities over the domain of input space at an individual level since the monotonicity constraints prevent the incorrect sign of the attribute effects at the individual level; thus, the population-level effect is naturally corrected.



(a)



(b)

Figure 1. Symbolic benchmarking of the proposed model against existing DCM models

2. METHODOLOGY

Formulation

The indirect utility for individual $n \in \{1, \dots, N\}$ for alternative $o \in \{1, \dots, O\}$ is a sum of systematic utility (V_{on}) and error term (ε_{on}) as denoted in **Equation (1)**, and the V_{on} can be represented as a function \mathbf{X}_{on} , with parameters $\boldsymbol{\theta}$.

$$U_{on} = V_{on} + \varepsilon_{on} = F_{on}(\mathbf{X}_{on}; \boldsymbol{\theta}) + \varepsilon_{on} \quad (1)$$

where $\mathbf{X}_{on} = (\mathbf{x}_{on}, \mathbf{z}_n)$ is a vector of input attributes, such that $\mathbf{x}_{on} \in R^{M-Q}$ is a vector of alternative-specific attributes and the $\mathbf{z}_n \in R^Q$ is a vector of individual-specific attributes. If the ε_{on} is assumed to be identically and independently Gumbel-distributed, the choice probability for the alternative o takes the form of the *Softmax* function as in **Equation (2)**.

$$P_{on} = \frac{e^{F_{on}(\mathbf{X}_{on}; \boldsymbol{\theta})}}{\sum_{j=1}^O e^{F_{jn}(\mathbf{X}_{jn}; \boldsymbol{\theta})}} \quad (2)$$

Based on the *Softmax* form of choice probability, the DCMs can be estimated by standard empirical risk minimization as in **Equation (3)**.

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \sum_{n=1}^N \mathcal{L}(\mathbf{y}_n, \mathbf{P}_{on}) \quad (3)$$

where \mathcal{L} is the standard cross-entropy loss function and \mathbf{y}_n is the choice made by individual n . The DCM-Linear assume the F_{on} as a linear function F_{on}^{Li} with $\boldsymbol{\theta}^{Li}$ that is a vector of coefficients for \mathbf{X}_{on} . The $\boldsymbol{\theta}^{Li}$ directly relates to the main effects $\boldsymbol{\beta}$ as in **Equation (4)**.

$$F_{on}^{Li}(\mathbf{X}_{on}; \boldsymbol{\theta}^{Li}) = \boldsymbol{\beta}^T \mathbf{X}_{on} \quad (4)$$

The DNN represent the F_{on}^{DNN} by multiple neurons in the multiple hidden layers (\mathbf{h}) as in **Equation (5)**.

$$F_{on}^{DNN}(\mathbf{X}_{on}; \boldsymbol{\theta}^{DNN}) = \mathbf{A}_{on}^T(\mathbf{h}_H \circ \dots \circ \mathbf{h}_2 \circ \mathbf{h}_1)(\mathbf{X}_{on}) \quad (5)$$

where H is the number of layers in the DNN, and \mathbf{A}_{on}^T is the last layer before the Softmax function to make the alternative-specific utility. $\boldsymbol{\theta}^{DNN}$ are the weights (i.e., parameters) connecting neurons and hidden layers.

Empirical studies have shown that F_{on}^{DNN} minimizes the empirical risk with overly complex models (i.e., over-estimated interactions) (Wang, Mo, et al., 2020). To address this issue, this study proposes a flexible but constrained form of LN-based utility function as in **Equation (6)**.

$$\begin{aligned}
F_{on}^{LN}(\mathbf{X}_{on}; \boldsymbol{\theta}^{LN}) &= \sum_{m=1}^M \mathbf{g}_m(\mathbf{X}_{on}[m]) \\
&+ \sum_{m' \neq m} \mathbf{r}_{m,m'}^1(\mathbf{g}_m(\mathbf{X}_{on}[m]), \mathbf{g}_{m'}(\mathbf{X}_{on}[m'])) + \dots \\
&+ \sum_{m' \neq m} \mathbf{r}_{m,\dots,m'}^M(\mathbf{g}_m(\mathbf{X}_{on}[m]), \dots, \mathbf{g}_{m'}(\mathbf{X}_{on}[m']))
\end{aligned} \tag{6}$$

where \mathbf{g}_m denotes an attribute-specific utility function that capture the inherent non-linear effect and $\mathbf{X}_{on}[m]$ is m -th attribute of \mathbf{X}_{on} . The $\mathbf{r}_{m,m'}^1$ indicates the first order interaction between the non-linear effects of m -th and m' -th attributes, and the $\mathbf{r}_{m,\dots,m'}^M$ captures the M -th order interactions. To ensure the trustworthy attribute-specific effect, we need to impose partial monotonicity constraints on \mathbf{g}_m and $\mathbf{r}_{m,\dots,m'}^M$ at individual level.

Lattice network

The key requirement for trustworthy attribute-specific effect is the partial monotonicity of utility function relative to a subset of attributes. For example, increase in travel cost never increases the utility of travel mode if all other attributes are unchanged, regardless of the level of travel cost and the individual attributes. The monotonicity constraints can be implemented by restricting a sum of interaction effects, which requires considering several inequality constraints during training. The LN captures the attribute-specific non-linear effect as segmented effects for each cell (i.e., piecewise linear effect) in the lattice and the interactions of these non-linear effects using multilinear-interpolation. Such combination of piecewise linear functions and multi-linear interpolation enable LN to drastically reduce the number of inequality constraints to be evaluated for the monotonicity constraints (Gupta et al., 2016). Figure 2 shows the LN framework consisting of the calibrator layer and lattice layer.

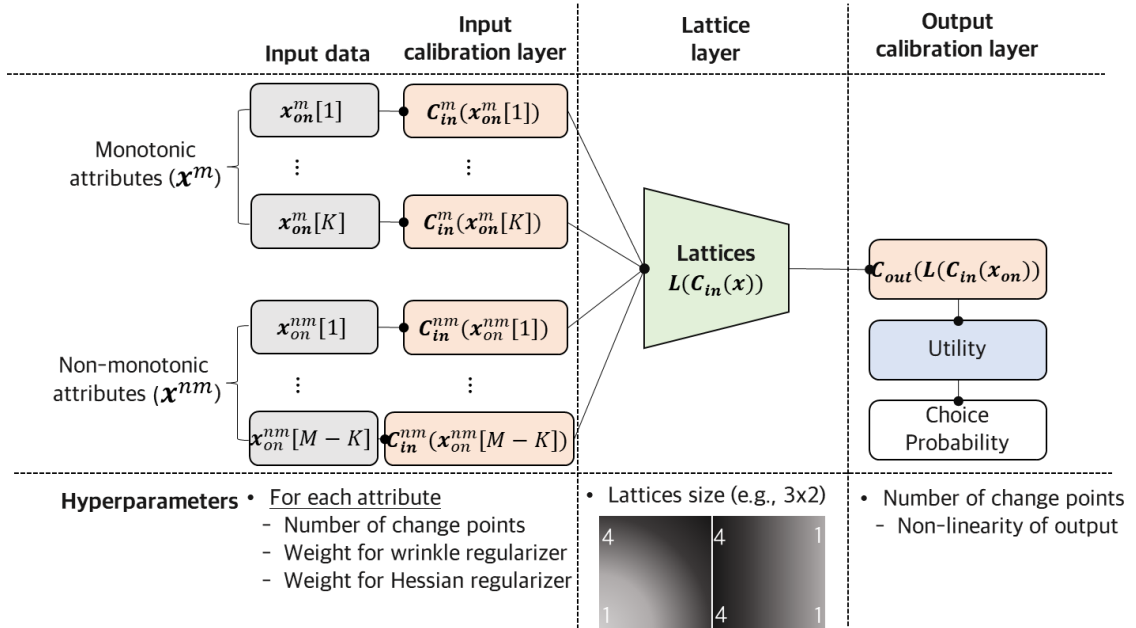


Figure 2. Lattice network framework

The input calibration layers \mathbf{C}_{in} in **Figure 2** implements the K attribute-specific transformation to capture the non-linear effect before the lattice layer, using one-dimensional monotonic piecewise linear function. These K transformation functions are estimated jointly with the lattice in the training. **Figure 3** illustrates the examples of transformation function in the calibration layer. The only hyperparameter for k -th attribute in the calibration layer is the number of changing points CP_k if we set the equally distanced cells.

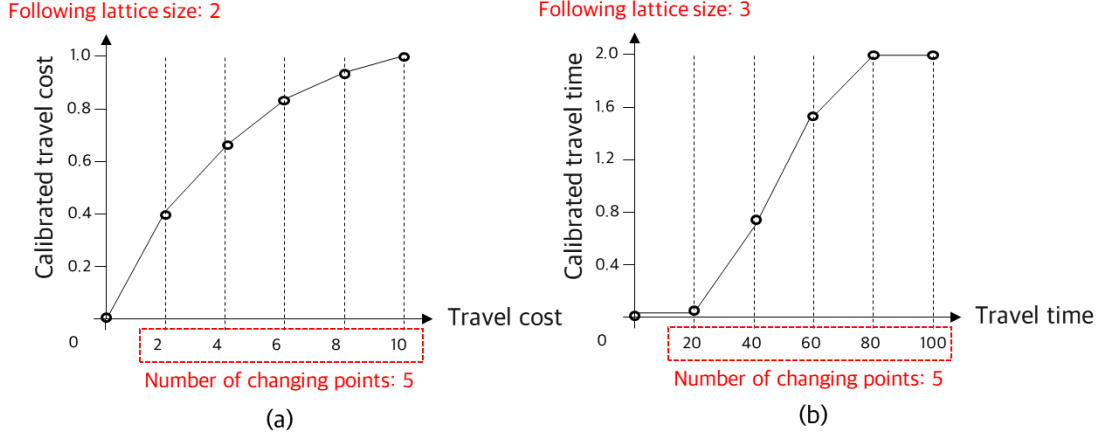


Figure 3. Examples of attribute-wise non-linear transformation through the calibration layer.

For the input attributes, $\mathbf{X}_{on} \in R^M$, we define the lattice size S_k for each attribute dimension, which is the number of vertices along the k -th attribute dimension. Then, the lattice can be represented by $S = S_1 \times S_2 \times \dots \times S_K$ parameters and spans the $(S_1 - 1) \times (S_2 - 2) \times \dots \times (S_K - 1)$ hyper-rectangle. The lattice estimates the value of function $L(\mathbf{C}_{in}(\mathbf{x}))$ by S parameters that is the value of function at each vertex. The larger lattice size can represent more flexible utility function. However, even if the lattice size is two for an attribute, the non-linear effect can be captured by an input calibration layer before the lattice. $\mathbf{C}_{out}(L(\mathbf{C}_{in}(\mathbf{x}_{on})))$ in **Figure 2** estimates the $F_{on}^{LN}(\mathbf{X}_{on}; \boldsymbol{\theta}^{LN})$ in **Equation (6)**, and $\boldsymbol{\theta}^{LN}$ consists of (i) the slopes of piecewise linear function in the attribute-specific calibration layers $\boldsymbol{\theta}_{inCal}^{LN}$ (ii) S parameters representing the value of function in the vertices $\boldsymbol{\theta}_{Lat}^{LN}$, and (iii) the slope of piecewise linear function in the output calibration layer $\boldsymbol{\theta}_{outCal}^{LN}$.

For the discrete choice datasets consisting of input attributes \mathbf{X}_{on} and output choices \mathbf{y}_n , the objective of the training LN is to estimate the $\boldsymbol{\theta}_{Lat}^{LN}$ while ensuring the monotonicity constraints. For the k -th attributes ($x_{on}^m[k]$), the increasing monotonicity is ensured if $\theta_{Lat,s}^{LN} > \theta_{Lat,r}^{LN}$ for all adjacent vertices s and r along the k -th attribute dimension. Similarly, the monotonicity constraints are also imposed on the attribute-specific input calibration layer. For the parameters of calibration layer for k -th monotonic attributes, $\theta_{inCal}^{LN}[k]$ (i.e., the slopes of piecewise linear function in each segment), $\theta_{inCal,u}^{LN}[k] > \theta_{inCal,v}^{LN}[k]$ should be maintained for all adjacent u and v , to make the $C_{in}(x_{on}^m)$ be a piecewise monotonic linear function. With these two levels of inequality constraints, the LN is estimated using a structural risk minimization. Then, the updated parameters are projected to ensure their monotonic constraints. The estimation of the $\boldsymbol{\theta}^{LN}$ is formulated as in **Equations (7-9)**.

$$F_{on}^{LN}(\mathbf{X}_{on}; \boldsymbol{\theta}^{LN}) = \mathbf{C}_{out}(\mathbf{L}(\mathbf{C}_{in}(\mathbf{x}_{on}; \boldsymbol{\theta}_{inCal}^{LN}); \boldsymbol{\theta}_{Lat}^{LN}); \boldsymbol{\theta}_{outCal}^{LN}) \quad (7)$$

$$P_{on}^{LN} = e^{F_{on}^{LN}(\mathbf{x}_{on}; \boldsymbol{\theta}^{LN})} / \sum_{o=1}^O e^{F_{on}^{LN}(\mathbf{x}_{on}; \boldsymbol{\theta}^{LN})} \quad (8)$$

$$\begin{aligned} & \arg \min_{\boldsymbol{\theta}^{LN}} \sum_{n=1}^N \mathcal{L}(y_n, P_{on}^{LN}) + R(\boldsymbol{\theta}_{inCal}^{LN}) \\ & s. t \mathbf{A}\boldsymbol{\theta}_{Lat}^{LN} \leq 0, \mathbf{B}\boldsymbol{\theta}_{inCal}^{LN} \leq 0, \text{ and } \mathbf{C}\boldsymbol{\theta}_{outCal}^{LN} \leq 0 \end{aligned} \quad (9)$$

$R(\boldsymbol{\theta}_{inCal}^{LN})$ is the regularization for input calibration layers (i.e., the wrinkle and Hessian regularizer). The matrix \mathbf{A} represents the inequality constraints for $S(= 2^K)$ parameters, and partial monotonicity is considered through the matrix \mathbf{A} . The matrix \mathbf{B} and \mathbf{C} play a similar role in implementing the partial monotonicity constraints in the input and output calibration layers, respectively. Details on the efficient optimization strategies for **Equation (9)** can be referred to Gupta et al. (2016).

This study adopts the individual conditional expectation (ICE)(Goldstein et al., 2015) as post-analysis tools for DNN and LN to explain the attribute-specific effect (i.e., utility function) at individual level. Readers can refer to more details of ICE and its pros and cons in the Molnar (2018).

3. RESULTS AND DISCUSSION

Simulation study

The data generating process (DGP) for binary choice (Han et al., 2022) are defined as follows. The three individual attributes – income (IN), full-time job (FUL), and flexible commuting (FLX) create systematic taste heterogeneity for the effects of two alternative-specific attributes – travel time (TT) and waiting time (WT). The IN is a categorical variable with 10 intervals, while the FUL and FLX are dummy variables. We also define the crowding (CR) and its interaction with TT, and inherent non-linear utility of TC. **Equation (10)** denotes the true systematic utility of individual n for alternative j .

$$\begin{aligned} V_{nj} = & -0.1 - \mathbf{8} \cdot \sqrt{TC_{nj}} - \mathbf{2.0} \times CR_{nj} + \\ & \left(\begin{array}{l} -0.1 - 0.5 \times IN_n - 0.1 \times FUL_n + 0.05 \times FLX_n - \mathbf{0.02} \times CR_{nj} \\ -0.2 \times IN_n \times FUL_n + 0.05 \times IN_n \times FLX_n + 0.1 \times FUL_n \times FLX_n \end{array} \right) \times TT_{nj} + \\ & \left(\begin{array}{l} -0.2 - 0.8 \times IN_n - 0.3 \times FUL_n + 0.1 \times FLX_n \\ -0.3 \times IN_n \times FUL_n + 0.08 \times IN_n \times FLX_n + 0.3 \times FUL_n \times FLX_n \end{array} \right) \times WT_{nj} \end{aligned} \quad (10)$$

Benchmark models

The DCM-DNN and DCM-Linear models are used as benchmark models. The DCM-Linear considers the true first-order interactions between the alternative and individual attributes (e.g., $FLX_n \times TT_{nj}$, $FUL_n \times TT_{nj}$) while ignoring second-order interactions (e.g., $FUL_n \times FLX_n \times TT_{nj}$), interactions between alternative attributes (e.g., $TT_{nj} \times CR_{nj}$), and inherent non-linearity (e.g., $8\sqrt{TC_{nj}}$). For the DCM-DNN and DCM-LN, we do not provide any information for the

DGP and input all the alternative and individual attributes. The DCM-LN only uses the prior knowledge in which TT, WT, CR, and TC monotonically decrease the utility at individual level.

Evaluation metrics for interpretability and predictability

This study evaluates the interpretability by comparing the true and estimated WTP. We compare the true and estimated distribution of value-of-time (VOT) and value-of-waiting-time (VOWT) to evaluate the capability of capturing individual taste heterogeneity. We define 40 (=10×2×2) individual groups by IN, FUL, and FLX. For DCM-DNN and DCM-LN, VOT for each individual group is calculated by aggregating ICE along all levels of the attribute value and each individual group. We examine the distribution of estimated VOT and VOWT using five quantile values: 1%, 25%, 50% (median), 75%, and 99%. Then, the accuracy of the estimated VOT and VOWT for 40 individual groups is evaluated by the root mean squared error (RMSE) and mean absolute percentage error (MAPE). The predictability is evaluated by accuracy for binary choice.

Evaluation results

Table 1 summarizes the evaluation results for interpretability and predictability. All the estimates are obtained from 50 synthetic datasets, and their mean and standard deviations are calculated. We examine the VOT and VOWT distributions using five quantiles value. The evaluation results provide four interesting findings. First, the predictability of the DCM-Linear is significantly worse than the DCM-DNN and DCM-LN, and its interpretability is worse than the DCM-LN. These results clearly show how the misspecification of utility function dramatically reduces the predictive performance and the trustworthiness of behavioural interpretation. The main cause may be the large discrepancy between the linear and non-linear utility functions for TC, which dramatically impacts WTP estimates. Second, DCM-DNN shows the best predictability and the worst interpretability, indicating that the trade-off relationship still holds for more complex functions. The interpretability of DCM-DNN is even lower than those of DCM-Linear. These results imply that the overly complex function fitted by DCM-DNN does not consider trustworthiness during training. Third, DCM-LN highly outperforms the DCM-Linear and DCM-DNN in terms of interpretability. It shows the best performance for all distribution and individual group values. In terms of predictive performance, DCM-LN highly outperforms the DCM-Linear but it is slightly outperformed by the DCM-DNN. Considering the balanced interpretability and predictability performance, the DCM-LN is the best model. Forth, both DCM-Linear and DCM-DNN estimate the negative VOT and VOWT for some individual groups, which substantially decreases the trustworthiness of the model’s interpretation. In comparison, the DCM-LN that ensures the individual-level monotonicity does not suffer such misidentification and provides slightly low but consistent WTP estimates.

Table 1. Interpretability and predictability evaluation.

Parameter		True		MNL		DCM-DNN		DCM-LN	
		Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
<i>Interpretability:</i> recovery of distribution	<i>VOT (Median)</i>	0.284	0.014	0.126	0.019	0.075	0.105	0.188	0.080
	<i>VOT (1%)</i>	0.142	0.010	-0.026	0.029	-0.012	0.281	0.093	0.063
	<i>VOT (25%)</i>	0.216	0.013	0.066	0.021	0.040	0.085	0.135	0.072

	<i>VOT (75%)</i>	0.404	0.024	0.200	0.013	0.123	0.172	0.257	0.089
	<i>VOT (99%)</i>	0.675	0.027	0.372	0.024	0.222	0.214	0.456	0.105
	<i>VOWT (Median)</i>	0.480	0.019	0.258	0.148	0.146	0.210	0.322	0.134
	<i>VOWT (1%)</i>	0.252	0.011	-0.068	0.124	-0.114	0.797	0.153	0.082
	<i>VOWT (25%)</i>	0.372	0.017	0.118	0.141	0.086	0.159	0.244	0.127
	<i>VOWT (75%)</i>	0.779	0.033	0.414	0.175	0.233	0.273	0.472	0.181
	<i>VOWT (99%)</i>	1.236	0.049	0.719	0.288	0.402	0.560	0.892	0.263
<i>Interpretability:</i> recovery of individual groups' value	<i>VOT (MAPE)</i>			0.630	0.044	0.802	0.329	0.351	0.103
	<i>VOT (RMSE)</i>			0.193	0.012	0.272	0.102	0.129	0.030
	<i>VOWT (MAPE)</i>			0.598	0.195	0.846	0.459	0.359	0.112
	<i>VOWT (RMSE)</i>			0.348	0.092	0.546	0.259	0.243	0.063
<i>Predictability:</i>	<i>Training accuracy</i>			0.552	0.006	0.775	0.010	0.741	0.018
	<i>Test accuracy</i>			0.546	0.013	0.716	0.014	0.697	0.016

Figures 4 and 5 support the findings derived from Table 1 and reveal some insightful patterns. First, all models underestimate the VOT and VOWT for most individual groups, except for some peak points caused by misidentification of the interaction effects, but the extent of the underestimation of DCM-LN is smaller than the other models. Second, Figure 5 shows that DCM-LN approximates the non-linear effect much better than DCM-DNN at both population and individual levels. One major issue of DCM-DNN is that it provides almost zero or positive marginal utility of TC for some levels of TC, which may lead to unreasonably high VOT or VOWT estimates, as in the peak of red lines in Figure 4. In contrast, DCM-LN could prevent such misspecification using its monotonicity constraints, also representing relatively stable WTP patterns in Figure 4. Third, the estimated alternative-specific utility functions in DCM-Linear are far from the true DGP. This result implies that there is a need to go beyond hand-crafted utility specifications if predictability is of interest.

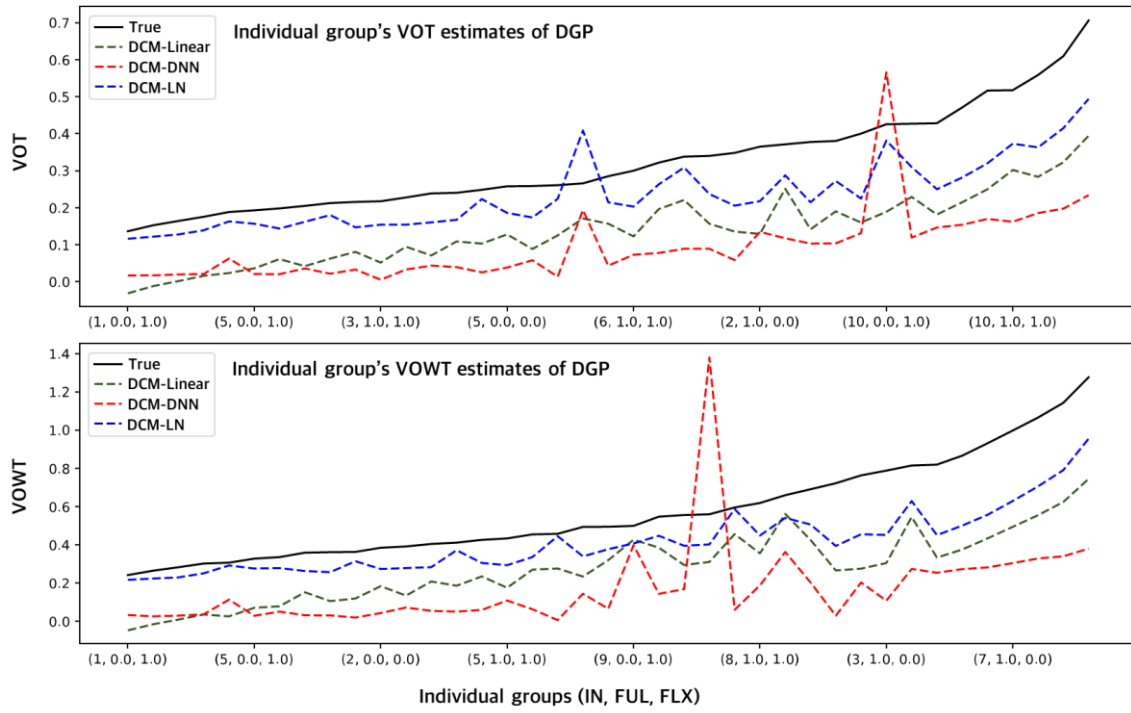


Figure 4. VOT and VOWT estimates for 40 individual groups in DGP.

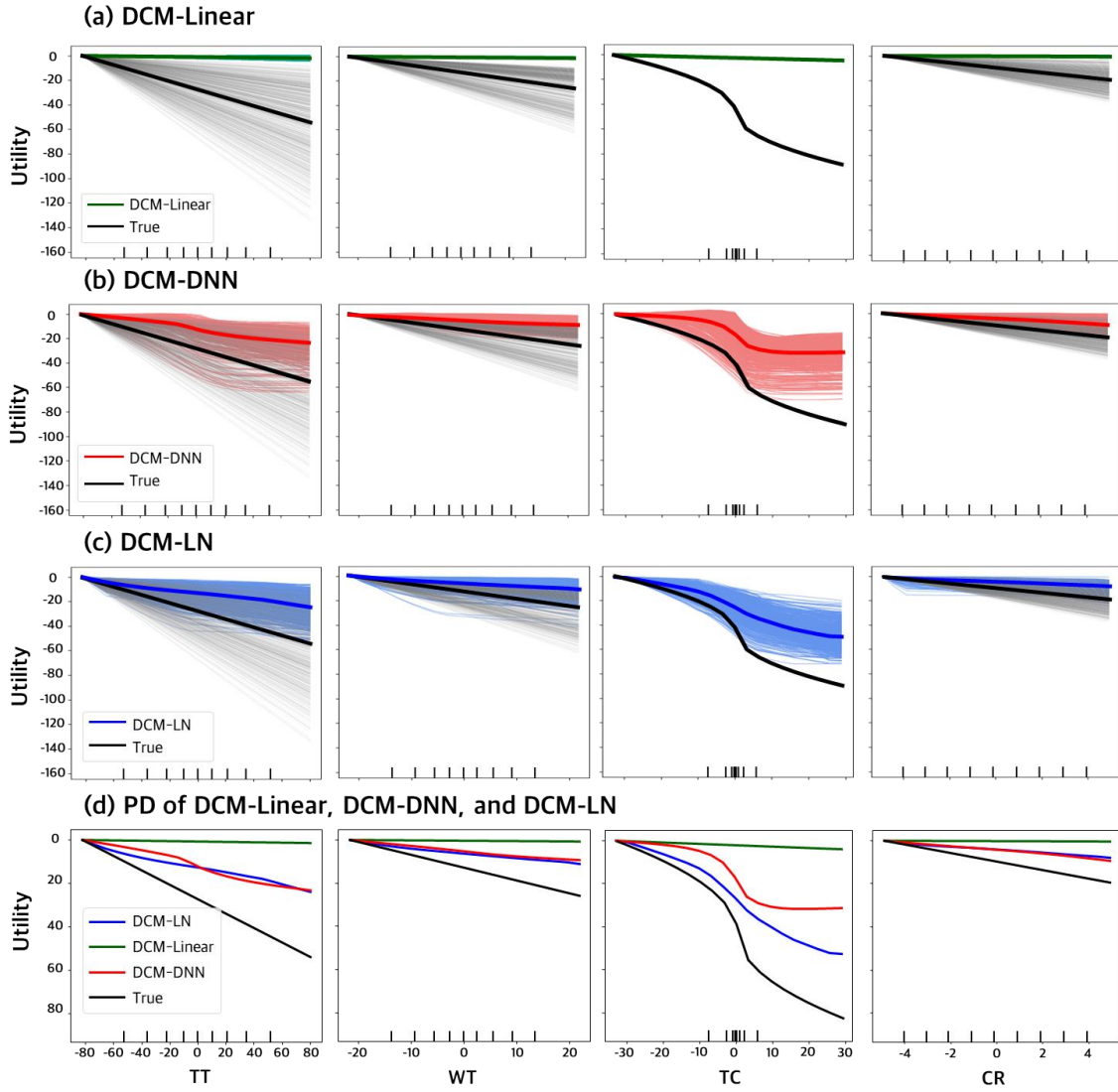


Figure 5. Attribute-specific utility functions estimated by (a) DCM-Linear, (b) DCM-DNN, and (c) DCM-LN at the individual and (d) population-level distribution (PD) of all three models.

4. CONCLUSIONS

In summary, we customize the lattice networks and introduce their first application to the DCMs to achieve a flexible utility specification while maintaining interpretability by imposing theory-driven constraints at an individual level. We benchmark the performance of DCM-LN against DCM-DNN and a parametric DCM (i.e., DCM-Linear) in a simulation study in terms of predictive accuracy and recovery of underlying marginal utility and individual-level WTP values across input space.

The evaluation results show that the DCM-LN highly outperforms the DCM-Linear and DCM-DNN in terms of interpretability, which is measured by the capability to recover the true utility and WTP of the simulation dataset. In contrast, the predictability of DCM-Linear is only slightly outperformed by the DCM-DNN, indicating that its capability to capture the complex

interactions in DGP remains intact after imposing monotonicity constraints. This balanced performance of DCM-LN is quite promising because it suggests that the DCM-LN approximates the true utility function during the training, rather than arbitrary functions that maximize the predictability as DCM-DNN does.

The work for evaluating the DCM-LN on real data is ongoing for further verification. Also, we are incorporating other behavioral mechanisms (i.e., soft constraints) into the DCM-LN, such as the non-compensatory decision rules (e.g., attribute cut-off and attribute non-attendance) and asymmetric marginal utility (e.g., prospect theory).

REFERENCES

- Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation. *Journal of Computational and Graphical Statistics*, 24(1), 44–65. <https://doi.org/10.1080/10618600.2014.907095>
- Gupta, M., Cotter, A., Pfeifer, J., Voevodski, K., Canini, K., Mangylov, A., Moczydlowski, W., & van Esbroeck, A. (2016). Monotonic Calibrated Interpolated Look-Up Tables. *Journal of Machine Learning Research*, 17, 1–47. <https://doi.org/10.5555/2946645.3007062>
- Han, Y., Pereira, F. C., Ben-Akiva, M., & Zegras, C. (2022). A neural-embedded discrete choice model: Learning taste representation with strengthened interpretability. *Transportation Research Part B: Methodological*, 163, 166–186. <https://doi.org/10.1016/j.trb.2022.07.001>
- Molnar, C. (2018). *A guide for making black box models explainable*. <https://christophm.github.io/interpretable-ml-book>.
- van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., & Walker, J. (2022). Choice modelling in the age of machine learning - Discussion paper. *Journal of Choice Modelling*, 42(December 2021), 100340. <https://doi.org/10.1016/j.jocm.2021.100340>
- Wang, S., Mo, B., & Zhao, J. (2020). Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions. *Transportation Research Part C: Emerging Technologies*, 112(February), 234–251. <https://doi.org/10.1016/j.trc.2020.01.012>
- Wang, S., Wang, Q., & Zhao, J. (2020). Deep neural networks for choice analysis: Extracting complete economic information for interpretation. *Transportation Research Part C: Emerging Technologies*, 118(February), 102701. <https://doi.org/10.1016/j.trc.2020.102701>

Competing on Emission Charges

Nicole Adler¹, Gianmarco Andreana^{*1,2}, and Gerben de Jong^{1,3}

¹Business School, Hebrew University of Jerusalem, Israel

²Department of Economics, University of Bergamo, Italy

³Department of Economics, Vrije Universiteit Amsterdam, Netherlands

SHORT SUMMARY

This research presents a game-theoretic model to analyse market equilibria in the presence of environmental policies at national and supranational levels. In a two-stage game, regulators maximise welfare over their jurisdiction by setting emission charges, whilst airlines compete through frequencies, fares, and fleet choice. Consequently, airlines decide whether to absorb the costs of the environmental charges, pass them on to consumers, replace part of their fleet with more efficient aircraft or redistribute the inefficient fleet to less regulated itineraries. The equilibria outcomes suggest the presence of several distorting forces that can undermine the effectiveness of environmental policies. To assess the robustness of our results, we apply the model to North American and Western European markets, under different regulatory setting, finding that a reduction in the emissions produced comes at the expense of the welfare and that the effectiveness of the policy is limited when regulators interact in their own interests.

Keywords: Decarbonization of transport, Transport economics and policy, Operations research applications, Discrete choice modelling.

1 INTRODUCTION

Within the transportation domain, the aviation industry currently produces 5% of global anthropogenic carbon dioxide (CO_2) emissions, and this is expected to continue to increase by 2050 (Lee et al., 2021; Kwan & Rutherford, 2015). Decision makers at the local, national, and supranational levels have mandated various environmental policies in an attempt to control aviation emissions (Larsson et al., 2019). However, the strength and environmental efficiency of these measures vary widely. Furthermore, since airlines operate globally, policymakers need information on how airlines respond to different, often overlapping, policies, to ensure that their interventions balance the carbon footprint of aviation with its wider economic and connectivity benefits.

Despite increasing understanding in recent years of the negative impacts derived from emissions, no effective and globally accepted emission control mechanism has yet been implemented. Many governments have developed unilateral emissions reduction schemes to regulate emissions production and to limit climate change. However, the lack of coordination among countries' policymakers likely generates suboptimal outcomes. A clear example is given by the presence of multiple, overlapping policies to address aviation emissions, such as the EU-ETS applied alongside Member States' ticket taxes and CORSIA. Another source of inefficiency that arises from the lack of coordination between countries manifests itself in the form of emissions leakage from heavily regulated countries to those jurisdictions in which the schemes are less strict (Baylis et al., 2013; Nordhaus, 2015; Perino et al., 2019). In addition, other market failures, such as firms' market power, will result in a departure from the standard first-best formulation in which government intervention addresses negative externalities by imposing a Pigouvian tax (Pigou, 1924) equal to the marginal external costs (Pels & Verhoef, 2004).

This calls for a game-theoretic framework to analyse how non-cooperative regulators at different administrative levels will set environmental policies strategically and how firms will subsequently react to such mechanisms. Given the complexity and numerous market distortions in the aviation industry, it is necessary to represent a realistic framework capable of including these industry-specific components. Our focus on the case of airline environmental regulation coincides with rapidly growing concerns about the impact of aviation emissions, and fragmentation in the aviation environmental regulatory setting offers an appropriate context for an applied game-theoretic

approach.

The purpose of this paper is to develop a game-theoretic model that assesses the impact of environmental policies in the aviation industry, taking into account both airline and regulatory competition. Specifically, our objective is to investigate how airlines respond to policies instigated by multiple non-cooperative policymakers at different administrative levels that set rules according to their own objectives. Our model allows to analyse and understand the policy implications deriving from the competition of multiple regulators and compare them to the implementation of an optimal global policy. We identify the cases in which a carbon charge may result in effective environmental policy and those in which the implementation of such a policy would fail due to divergence in regulator objectives. This style of game represents a novelty in the (air) transportation literature and may be used to analyse environmental and regulatory issues in other network industries.

2 METHODOLOGY

We define our game-theoretic model as a two-stage Nash game with perfect information. The set of players in the first-stage is characterised by the different regulators of the countries in which airlines are based and/or supranational decision-makers. In the first-stage, each regulatory body aims to maximise the social welfare of the area under its control by setting the level of environmental taxation to be applied. The regulator may reduce (global) environmental damages by setting a relatively high environmental tax, but this may come at a cost to both (local) consumer and producer surplus. Consequently, regulators compete on the entire level of emissions produced considering how much they are susceptible to the environmental damage resulting from these emissions. In the second-stage, airlines compete with each other by setting airfares and service frequencies through their best response functions, pursuing profit maximisation. To respond to changes in the climate policies, airlines may replace inefficient aircraft with more environmentally friendly technologies, fly their higher-emission aircraft less, reallocate their higher-emission aircraft to routes with less environmental taxation or reduce frequencies on regulated routes.

We define a hub-and-spoke network, $G(\mathcal{N}, \mathcal{K})$, where the nodes are connected to the spokes through ordered legs within the set \mathcal{K} , allowing indirect connections between the spokes passing through the hub airport.

Given the network configuration, airlines are subject to different levels of climate policies imposed by regulators. The sets belonging to the area of influence of a specific regulator are defined as:

$$\mathcal{N}^r = \{i^r, j^r | i^r, j^r \in \mathcal{N}, i^r \text{ and } j^r \text{ are nodes in the area regulated by } r\}$$

$$\mathcal{A}^r = \{a^r | a^r \in \mathcal{A}, a^r \text{ is an airline serving the area regulated by } r\}$$

$$\mathcal{K}^r = \{k^r | k^r \in \mathcal{K}, k^r \text{ is a network leg served by an airline based in the area regulated by } r\}$$

In the first stage, regulators maximise the social welfare of the area under their influence. Welfare is composed of four main components: passenger surplus, producer profits, government income from environmental taxation and environmental damages.

$$\begin{aligned} \underset{\theta_r}{Max} SW_r = & \sum_{i^r j^t} d_{i^r j^t} \frac{1}{-\beta_{2t}} \ln \left(e^{V_0 + \sum_a V_{i^r j^t a} (f_{k^r v a}^* p_{i^r j^t a}^*)} \right) + \sum_{a^r \in \mathcal{A}^r} \pi_{a^r} (f_{k^r v a^r}^*, p_{i^r j^t a^r}^*, x_{h v a^r}^*, \theta_r) \quad (1) \\ & + \sum_{k^r v a} \varepsilon_{k^r v} f_{k^r v a}^* \theta_r - \eta_r \sum_{k v a} \varepsilon_{k v} \xi f_{k v a}^* \end{aligned}$$

where

$$\varepsilon_{k v} = \gamma_k \phi_{h v} CO_2$$

is the ton of CO_2 produced on a flight leg k by a specific version of aircraft v . In 1, the first element represents the consumer surplus, expressed as the log-sum of the utility of passengers departing

from the jurisdiction of the regulator (Small & Rosen, 1981). The second element represents the profit generated by airlines certified within the jurisdiction of the regulator. The third element is the income from the carbon charge imposed on CO_2 generated over the regulated arcs and the last element expresses the share η_r of the overall social cost of emissions that the regulator takes into account. The decision variable for the regulatory entity is the charge per ton of carbon (θ_r) originating from a flight departing from its jurisdiction, taking into account the behavior of the other regulatory agencies and the airlines' responses to carbon charges in the second-stage.

Regulators are encouraging the internalization of environmental externalities, which are public goods. The aviation industry belonging to each regulator contributes to the total amount of emissions produced while providing connectivity between regions. However, not all regions are affected in the same way by emissions. Specifically, we allow for different degrees of risk exposure through the parameter η_r . In this way, regulators have the incentive to free-ride on the emissions reduction achieved by the actions of competing regulators. All CO_2 emissions generated by civil aviation bear a social cost common to all regulators, namely the (global) social cost of carbon represented by the parameter ξ . This social cost is homogeneous over the regions given the global impact of carbon emissions on the environment, however the exposure or distribution of this impact varies over regions. In our game, regulators can trade off environmental externalities with the surplus of passengers and carriers (profits) in their region by deciding the level of taxation on CO_2 in their jurisdiction. Consequently, the regulator's decisions are strictly connected to other regulators' actions, creating competition across jurisdictions.

We assume that passengers are utility maximizers when selecting the airline and the itinerary for their trip. According to McFadden (1974), utility can be decomposed into a systematic component, V_{ijta} and a random element, ϵ_{ijta} :

$$U_{ijta} = V_{ijta} + \epsilon_{ijta}, \quad \forall i, j \in \mathcal{N}, t \in \mathcal{T}, a \in \mathcal{A} \quad (2)$$

The systematic component is defined in the following way:

$$V_{ijta} = \beta_{0t}\delta_{ija} + \beta_{1t}\ln\left(1 + \min_{k^* \in \mathcal{K}}(f_{ka})\right) + \beta_{2t}p_{ijta} + \beta_{3t}\tau_{ija}, \quad (3)$$

$$\forall i, j \in \mathcal{N}, t \in \mathcal{T}, a \in \mathcal{A}$$

where δ_{ija} is the component of the utility associated with a direct connection and the second term represents the utility of a higher service frequency. When flying indirectly, only the lower frequency of the two legs is taken into account (Hansen (1990)). The third element represents the disutility from paying the ticket fare and the last represents the loss of utility generated by the travel time τ_{ija} . Consequently, demand is shared between airlines through a multinomial-logit model (MNL) that determines market shares:

$$m_{ijta} = \frac{e^{V_{ijta}}}{e^{V_0} + \sum_{a' \in \mathcal{A}} e^{V_{ijta'}}}, \quad \forall i, j \in \mathcal{N}, t \in \mathcal{T}, a \in \mathcal{A} \quad (4)$$

where the term V_0 is the utility associated with the outside-option from not flying.

According to Swan & Adler (2006), the direct operating cost of the airline is defined through a cost function that differentiates between long- and short-haul flights.

$$C_{kv} = \begin{cases} (\gamma_k + 722)(s_{kv} + 104)\$0.019 & \text{if } k \in \mathcal{K}^s \\ (\gamma_k + 2200)(s_{kv} + 211)\$0.0115 & \text{if } k \in \mathcal{K}^l \end{cases} \quad (5)$$

where

$$\mathcal{K}^s = \{k^s | k^s \in \mathcal{K} \text{ are the short-haul legs served}\}$$

$$\mathcal{K}^l = \{k^l | k^l \in \mathcal{K} \text{ are the long-haul legs served}\}$$

The monthly cost of owning an aircraft (o_{hv}) is approximated by the equivalent annual capital costs divided by the number of months per year.

In the second stage, airlines maximize their profits, given the environmental charges imposed by regulators in the first stage. Each airline strategically sets the service frequency of each version of aircraft f_{kva} per leg, the fares p_{ijta} on the itineraries between an origin and destination and the optimal number and version of the aircraft x_{hva} to operate given their network.

$$\text{Max}_{p_{ijta}, f_{kva}, x_{hva}} \pi_a = \sum_{\substack{i,j,t \\ i \neq j}} d_{ijta} m_{ijta} p_{ijta} - \sum_{k,v} C_{kv} f_{kva} - \sum_{k^r,v} \varepsilon_{k^r v} \theta_r f_{k^r va} - \sum_{h,v} o_{hv} x_{hva} \quad (6)$$

where m_{ijsa} is the market share function specified in 4, representing the share of demand served by a specific airline a for each city pair and passenger type, C_{kv} represents the operating costs, defined in 5, incurred by the airline for serving a specific leg, o_{hv} is the monthly ownership cost of the type of aircraft h version v and x_{hva} is the number of aircraft of type h and version v that carrier a operates in its network.

The competition framework for regulators and airlines is structured as an extensive form game with complete and perfect information (Osborne & Rubinstein, 1994). In this model, players make strategic decisions sequentially in two stages. This allows second-stage players to decide their strategy in response to the decisions of first-stage players. The actions of the regulators, in the first-stage, are represented by the environmental charges imposed on airlines, while, in the second-stage, airlines react by choosing service frequency, ticket fares and the number of new and old version aircraft to deploy.

It is possible to solve this two-stage simultaneous game using a Kuhn-Zermelo-type backward induction algorithm (Schwalbe & Walker, 2001), as described in 1. The algorithm starts by initialising the values for the first and second-stages. Successively, the algorithm solves the first-stage problem for each regulator in the set \mathcal{R} , moving to estimate the sub-game perfect Nash equilibrium (SPNE) in the second stage for each airline and continuing until no airline changes the values of their decisions variables. The second stage, non-linear mathematical programs are solved using *IPOPT* (Wächter & Biegler (2006)). Following the approach used in Adler et al. (2022), the first-stage algorithm performs a line search around each regulator's incumbent solution starting at + and - 50% and gradually decreasing to the point in which a further reduction would not improve the integer solution. A cycle is completed once all regulators have chosen their current optimal carbon tax. The equilibrium of the game is found when two cycles are completed such that no actor in the game changes the values of their decision variables.

The robustness of the results is tested by selecting different starting points and sequences of players within each specific set of players.

3 RESULTS AND DISCUSSION



Figure 1: Selected nodes in North America and Western Europe.

Algorithm 1 Solve the two-stage game (pseudo-code)

```
1: Start
2: initialise values of competitors' decision variables and their network, for regulators and
   airlines
3: while first-stage solution  $>$  optimal threshold do:
4:   while first-stage solution is not a best response for all regulators do:
5:     for each regulator do:
6:       create point grid around previous first-stage solution
7:       for each point in grid do:
8:         while second-stage solution not a best response for airlines do:
9:           for each airline do:
10:            solve mathematical program using IPOPT
11:            assess whether second-stage solution is a best response for all airlines
12:         return second-stage solution
13:       return second-stage solution for each point
14:       select the point that maximises welfare
15:     return first-stage solution for each regulator
16:   shrink grid radius
17: return first and second-stage solutions
18: Stop
```

In this section, we analyse a game that describes the aviation markets of North America and Europe. We assume a social cost of carbon of €200, according to the latest IPCC report (Pörtner et al., 2022). Our network covers 9% of the monthly traffic within and between Europe and North America.

The baserun, presented in Table (1) aims to replicate the 2019 transport equilibria outcome taking into account the European carbon charge in order to replicate the EU-ETS scheme. After accounting for free permits, we assume that the cost of carbon in 2019 was approximately €22 per ton of CO_2 produced. The results from the baserun case show that, despite the higher demand in North America, the European market generates a higher surplus than that of North American. This discrepancy between the two regions is due to the disutility faced by North American passengers who paid a higher fare than their European counterparts in 2019 and have a regional network characterized by longer distances. Thus, higher fares are the result of operating costs in North America and the presence of fewer alternative modes of transport, resulting in North American passengers being more dependent on aviation. Given the higher demand in North America, both LCCs and legacy carriers operate more flights in this region compared to Europe. As a result of higher fares and higher demand, North American carriers are more profitable than European airlines, despite the higher operating costs incurred by North American airlines. Regarding CO_2 emissions, we note that European carriers spent around €8 million in the monthly game covering 9% of the European market (equivalent to €1 billion for the entire market in 2019) and emission damages amounted to a total of €29 billion across the two regional markets (EU and NA) and Trans-Atlantic (TA) routes.

Table 1: Baserun scenario

	<i>Baseline</i>	
	EU	NA
θ_r (€)	22	
Government surplus (€ M)	8	
Emissions (€M)	-110	-110
Consumer surplus (€M)	942	683
Producer surplus (€M)	102	139
Welfare (€M)	942	712

Table 2: Validation (real world values in brackets)

	CASK (€c)	RASK (€c)	Demand, two-way (pax. M)		
			EU	NA	TA
EU legacy	7.2 (7.1)	9.1 (7.7)			
NA legacy	6.4 (6.4)	7.9 (8.7)			
EU LCC	4.2 (4.3)	5.3 (4.8)	4.7 (4.7)	5.7 (4.9)	0.5 (0.4)
NA LCC	5.3 (5.9)	6.2 (6.5)			

The second scenario explores the impact of a global regulator who sets a single charge per ton of CO_2 generated in all aviation markets. There is no possibility of free-riding in this scenario because the single regulator fully bears the costs of all generated emissions ($\eta_r = 1$). The results of the model, presented in Table (3), suggest that the optimal charge set by the regulator is much lower than the expected Pigouvian tax, which should compensate for the social costs of the carbon equal to €200 per tonne of CO_2 produced. This is due to the airlines' market power and the Mohring effect, both of which induce the regulator to lower the tax. In the case of market power, the regulator is attempting to counter the output reduction of hubbing carriers that choose to serve fewer passengers with higher fares, thereby increasing their own profits but decreasing social welfare. The Mohring effect captures the idea that each additional passenger contributes towards higher frequency, hence the quality of the air travel services for all other passengers (Mohring, 1972). As these benefits are external to the passengers (i.e. they are positive externalities), too few passengers choose to travel from a societal perspective, which the regulator can address through subsidies. Consistent with the economic literature, these two effects lead to a lower carbon charge compared to the (marginal) social cost of carbon in our game.

It is also important to observe that, in our framework, a regulator is not able to discriminate across routes, and the charge is the same for all operations. Such a limitation may result in a sub-optimal tax, because charges cannot be tailored to local conditions (Benoot et al., 2013). As a result of this global scope policy, we observe a slight departure from the *base-run* scenario. Specifically, as a consequence of this marginal global charge, we do not observe significant changes in airline strategies. With regard to the environment, the imposition of a charge on the North American market too leads to a small reduction in the emissions generated.

Table 3: Single regulator scenario

	<i>Baseline</i>			<i>1REG</i>	Δ
	EU	NA	Sum	REG	
θ_r (€)	22	0	22	8	
Government surplus (€M)	8	0	8	10	
Emissions (€M)	-110	-110	-220	-219	1
Consumer surplus (€M)	942	683	1,624	1,623	-2
Producer surplus (€M)	102	139	241	241	0
Welfare (€ M)	942	712	1,654	1,655	-1

We now define scenario, *2REG*, in which two regulators, based in different regions, compete by setting emission charges on all flights departing from their jurisdiction. We assume that one regulator sets charges for all flights departing within and from North America, and similarly, within and from Europe. The environmental risk is distributed equally between the two regions. Given the round-trip assumption of each flight, operations within a region are charged twice by the same regulator. Trans-Atlantic flights are subject to both regulators' charges, one per direction. The results of our model for this scenario are reported in Table (4). We observe that competing regulators decide to free-ride on each other, resulting in charges that are much lower than the social cost of carbon. In this way, regulators protect the surplus of both passengers and carriers under their jurisdiction. Indeed, the results of this competing regulator case closely reflect the charges currently imposed by Europeans (EU-ETS price of €22) and North Americans (€0) in the real world in 2019. In North America, a higher charge would be welfare-damaging given the longer routes flown and the lower surplus of passengers resulting from higher fares. In Europe, where airfares and distances are lower and alternative modes compete with aviation, the regulator has a greater incentive to set a positive charge. As a consequence of the implementation of a small but positive tax on both continents, the airlines respond by slightly increasing airfares and moderately reducing service frequency.

Table 4: Two-regulator scenario

	Baseline		<i>2REG</i>		Δ	
	EU	NA	EU	NA	EU	NA
θ_r (€)	22	0	38	4	16	4
Government surplus (€M)	8	0	14	3	6	3
Emissions (€M)	-110	-110	-109	-109	1	1
Consumer surplus (€M)	942	683	936	678	-6	-4
Producer surplus (€M)	102	139	101	139	-1	-0
Welfare (€M)	942	712	943	711	0	-1

Finally, we note that the only path to reducing emissions substantially would appear to be an increase in the social cost of carbon, as demonstrated in Figure (2). Once the cost of carbon is above € 500, the airfares increase by one sixth, the frequencies drop by one third and the social welfare accordingly but so too the emissions.

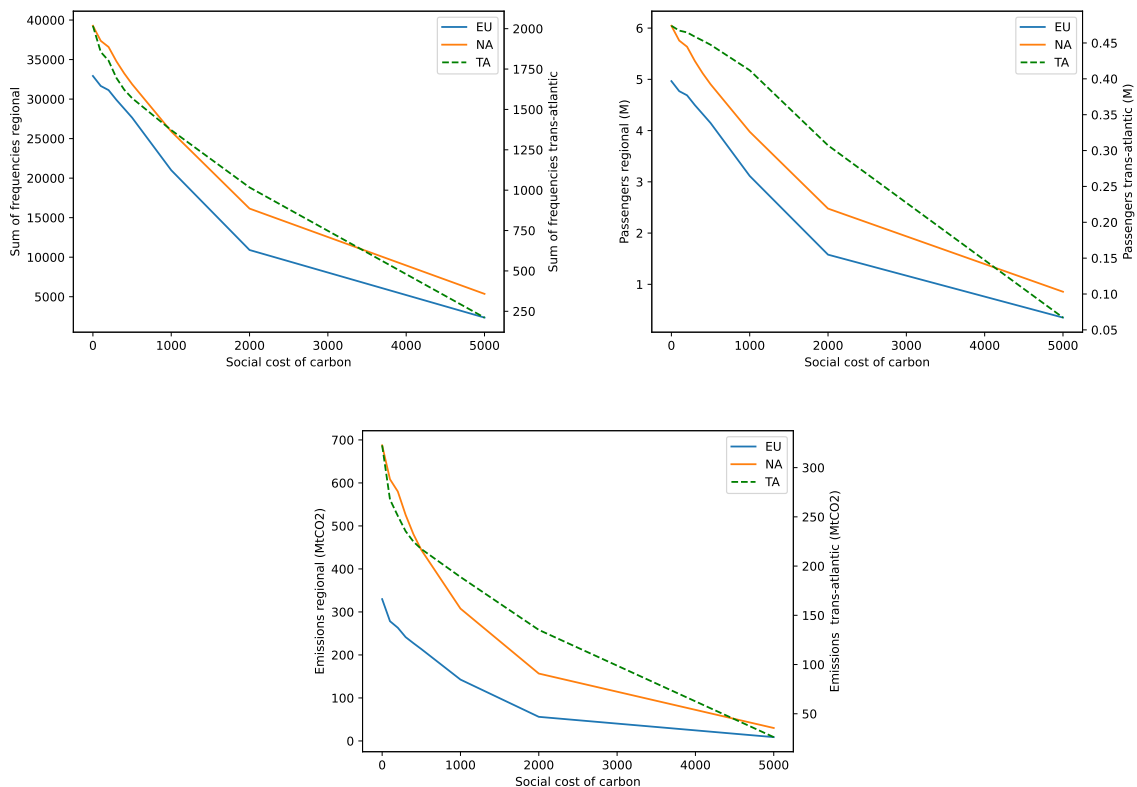


Figure 2: Sensitivity analysis over the social cost of carbon

4 CONCLUSIONS

In this paper, we develop a two-stage model capable of representing competition between regulators and airlines under different emission charges. By comparing scenarios with a 2019 baserun case, we assess the impacts of the different regulators' interactions on welfare and the environment. Our analysis suggests that imposing an environmentally optimal carbon charge on the aviation industry can lead to unexpected and welfare-detrimental outcomes. Specifically, we have assessed that the carbon charge imposed by a single regulator results in a level that is well below the social impact of emissions. We further show that when regulators are free to set their charges, they enter into regional surplus protectionism, which undermine the effectiveness of the mechanism. The outcomes we present in our paper are the result of several distorting forces in the aviation industry and offer

an explanation behind the reasons for the absence of an international cooperative carbon policy.

ACKNOWLEDGEMENTS

We would like to thank the Netherlands Organisation for Scientific Research Award 019.201SG.019, Israel Science Foundation Grant 2441/21, INFORMS Aviation Applications Section, and Goldman Data-Driven Innovation Research Centre for funding this research.

REFERENCES

- Adler, N., Hanany, E., & Proost, S. (2022, 4). Competition in congested service networks with application to air traffic control provision in europe. *Management Science*, *68*, 2751-2784.
- Baylis, K., Fullerton, D., & Karney, D. H. (2013, 5). Leakage, welfare, and cost-effectiveness of carbon policy. *American Economic Review*, *103*, 332-337.
- Benoot, W., Brueckner, J., & Proost, S. (2013). Intercontinental-airport regulation. *Transportation Research Part B: Methodological*, *52*, 56-72.
- Hansen, M. (1990, feb). Airline competition in a hub-dominated environment: An application of noncooperative game theory. *Transportation Research Part B: Methodological*, *24*(1), 27-43.
- Kwan, I., & Rutherford, D. (2015). Assessment of U.S. domestic airline fuel efficiency since 2010. *Transportation Research Record*, *2501*, 1-8.
- Larsson, J., Elofsson, A., Sterner, T., & Akerman, J. (2019, jul). International and national climate policies for aviation: a review. *Climate Policy*, *19*(6), 787-799.
- Lee, D. S., Fahey, D., Skowron, A., Allen, M., Burkhardt, U., Chen, Q., ... Fuglestedt, J. (2021). The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric Environment*, *244*, 117834.
- McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303-328.
- Mohring, H. (1972). Optimization and scale economies in urban bus transportation. *The American Economic Review*, *62*(4), 591-604.
- Nordhaus, W. (2015, apr). Climate Clubs: Overcoming Free-riding in International Climate Policy. *American Economic Review*, *105*(4), 1339-1370.
- Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. MIT press.
- Pels, E., & Verhoef, E. T. (2004). The economics of airport congestion pricing. *Journal of Urban Economics*, *55*(2), 257-277.
- Perino, G., Ritz, R. A., & Van Benthem, A. (2019). Understanding overlapping policies: Internal carbon leakage and the punctured waterbed. *National Bureau of Economic Research*.
- Pigou, A. C. (1924). *The economics of welfare*. Macmillan.
- Pörtner, H.-O., Roberts, D. C., Adams, H., Adler, C., Aldunce, P., Ali, E., ... Biesbroek, R. (2022). Climate change 2022: Impacts, adaptation and vulnerability. *IPCC Sixth Assessment Report*.
- Schwalbe, U., & Walker, P. (2001). Zermelo and the early history of game theory. *Games and economic behavior*, *34*(1), 123-137.
- Small, K. A., & Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, 105-130.
- Swan, W. M., & Adler, N. (2006, mar). Aircraft trip cost parameters: A function of stage length and seat capacity. *Transportation Research Part E: Logistics and Transportation Review*, *42*(2), 105-115.

Wächter, A., & Biegler, L. T. (2006, mar). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1), 25–57.