Behavioral responses to taxes

Fiscal implications of tax rate changes, their perception and tax privacy

Von der Wirtschaftswissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften
- Doctor rerum politicarum -

genehmigte Dissertation

von

Dipl.-Kfm. Jonathan Bob
geboren am 28. April 1981 in Berlin

2013

Referenten

Professor Dr. Kay Blaufus
Professor Dr. Andreas Wagener

Tag der Promotion

29. August 2013

Keywords

Perceived tax burden · Tax evasion · Value added taxes

Schlagwörter

Gefühlte Steuerbelastung · Steuerhinterziehung · Mehrwertsteuer

Summary

This dissertation is a collection of research articles according to § 2 Promotionsordnung (doctoral regulations) of the School of Economics and Management, Leibniz Universität Hannover. Table 1 provides an overview of these articles, the respective co-authors and the status of publication. In the following, the four papers will be outlined.

Title	Co-authors	Journal / Status
Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy	Kay Blaufus Jochen Hundsdoerfer Dirk Kiesewetter Joachim Weimann	Journal of Economic Psychology 35, 1–16
Perception of income tax rates: evidence from Germany	Kay Blaufus Jochen Hundsdoerfer Christian Sielaff Dirk Kiesewetter Joachim Weimann	European Journal of Law and Economics, online first, doi: 10.1007/s10657-013-9389-9
The effect of tax privacy on tax compliance – An experimental investigation	Kay Blaufus Philipp E. Otto	Under review: Journal of Economic Behavior & Organization
Who bears value added taxes? Evidence from Germany	—	Working Paper

Table 1: Overview of research articles

Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy

Many countries have pursued a tax-cut-cum-base-broadening policy over the past 20–30 years; i.e., these countries have reduced the nominal tax rate and simultaneously increased the tax base. Usually, traditional economic theory assumes individuals to behave rationally. Then, it should make no difference whether their tax rate or (by equal measure) their tax deductions changes. By contrast, if one takes into account the idea that individuals tend to avoid cognitive strain and instead use simplified decision heuristics there may be difference in perception.

The main purpose of this study is to examine whether decision heuristics are used in choosing between tax options that differ, exclusively, in the nominal tax rate and the amount of tax deductible expenses and whether the use of decision heuristics leads to a misperception of changes in nominal tax rates compared to changes in tax deductions. To this aim, both (i)

a survey study (conjoint analysis) with 467 German working individuals that represents the population in terms of age, education, gender, and income and (ii) a lab experiment with 56 working individuals are conducted. By combining these two methods we are able to benefit from the respective strengths of both procedures. First of all, conjoint analysis allows us to draw a large representative sample. Moreover, we are able to measure preferences on the individual level. This incorporates the advantage of conjoint analysis to identify whether individuals choose a decision heuristic over the rational ranking. On the other hand, conjoint analysis is a mere preference measurement whereas a lab experiment sets monetary incentives, i.e., individuals are compensated according to their experimental performance. Besides, the lab experiment enables us to establish a learning environment.

In line with behavioral studies incorporating the burden of cognitive strain, our results reveal that the majority of individuals do not make rational decisions based on the actual tax burden but rather use simple decision heuristics. This leads to an irrationally high impact of changes in nominal tax rates on the perceived tax burden. Taxpayers favor tax options that apply a lower tax rate on a broad base over a higher tax rate applied on a narrow base despite the lower actual tax burden of the latter option. Furthermore, overestimation of tax rate changes increases considerably when information on tax rate is considered first (framing effect). Even after the introduction of performance based incentives we observe that subjects still overestimate nominal tax rate changes. However, we find slight evidence for learning effects. After the implementation of a learning environment, the use of heuristics decreases weak significantly, but even after three decisions the majority of subjects do not make a rational choice.

These results are of particular importance for fiscal policy as they suggest that politicians could combine increasing fiscal revenues and decreasing subjects' tax perception. Furthermore, consistent with the identified framing effect, emphasizing the nominal tax rate when presenting tax options can reduce the perceived tax burden solely through the display format.

Perception of income tax rates: evidence from Germany

As early as 1960 Schmoelders pointed out how important the perceived tax burden is for tax policies. Based on his survey he showed that only one-third of the subjects knew their actual tax burden. However, economic tax research has continued to use almost exclusively actual tax rates.

The main purpose of our study is to examine whether 50 years later the results of Schmoelders are still valid. Like Schmoelders, we are interested in both the perceived tax burden and the relation between the perceived tax burden and the tax burden deemed as fair. In addition, we expand his work by analyzing determinants of the perceived and the fair tax burden in more detail. For this purpose, more than 1,000 employed and self-employed people were interviewed.

Based on this survey data, we measure perceived tax rates by means of four pre-defined income categories as well as subjects' own income. We compare these perceived income tax rates to the actual tax rate, as well as a tax rate that is taken as fair. We can thus determine whether taxpayers were able to accurately estimate their actual tax burdens. Also, we analyze the individual taxpayer's statements regarding whether and to what extent they feel that the tax burden that they perceive can be considered as fair.

We find Schmoelders' results to be quite robust over time. For the majority of subjects the perceived tax burdens significantly deviate from the actual tax burdens. Subjects, on

average, perceive both the average tax rate based on the pre-defined income categories and the marginal tax rate based on the own income as too high at lower income levels and too low at higher income levels. The degree of misperception can be partly explained by the individual's level of education, income, and whether the individual included social security contribution in their income tax rate estimation.

The comparison of the perceived tax burden and the tax burden regarded as fair showed that subjects considered the tax burden as too high for a pre-defined income of EUR 300,000 or lower, while they considered the tax burden as too low for a pre-defined income of EUR 2 million. The most remarkable aspect is the discrepancy of the tax burden that is regarded as fair compared to the perceived tax burden, on the one hand, and to the actual burden, on the other hand. Taxpayers believe it is fair to decrease taxes on the pre-defined income of EUR 10,000 and to increases taxes on the pre-defined income of EUR 2 million. Contrary, the tax burden perceived as fair on an income of EUR 10,000 significantly exceeds the actual tax burden while it is significantly lower for an income of EUR 2 million. Hence, it is of major importance whether recommendations regarding tax policies are based on the individuals' perceived or the actual tax burden. Analyzing the determinants, we can see that the valuation whether the tax burden is considered fair strongly depends on the subjects education and age. Accordingly, the higher the education, the higher the probability that the subject perceives the tax rate as fair. Age affects the perception to the extent that younger subjects in particular perceive the tax burden as too high.

Finally, our results suggest that, generally, individuals prefer a progressive tax tariff due to fairness reasons.

The effect of tax privacy on tax compliance – An experimental investigation

Tax evasion is a problem for all countries. E.g., the tax gap in the United States amounts to $385 billion for the tax year 2006, alone. For more than a century, there has been an ongoing debate whether tax evasion can be fight against by publicly disclose tax return information. While most countries treat tax information confidentially, some countries as Greece and New Zealand publicly list tax evaders. Even more, the Nordic Countries disclose all tax return information. The main reason to disclose tax compliance information publicly is to deter people from evading taxes by threatening them with the shame of being announced as tax evaders. However, whether a strategy of tax publicity is a successful instrument for fighting tax evasion is far from obvious.

Social norms may have a considerable impact on tax evasion, too, with individuals complying as long as they believe that tax compliance equals the social norms. However, public disclosure allows for observing the unethical behavior of others. This is potentially contagious because it may change social norms regarding compliance. Hence, publishing information could be a deterrent to tax evasion but simultaneously it could also enhance it. Due to these potentially opposing effects—increasing shame on the one hand, risk of contagion on the other—the overall effect of public disclosure on tax compliance remains unclear.

We design a tax game with tax privacy as the treatment variable to investigate the impact of public disclosure on tax evasion experimentally. The three different levels of tax privacy ranging from full confidentiality to full publicity. Tax information is either published in an anonymous manner or photos of subjects are paired with each subject's tax behavior. This setting allows for separating both the shame and the contagion effects because only contagion but not shame should not arise under the anonymous disclosure.

The overall result is an increase of non-compliance in the absence of tax privacy if information is disclosed anonymously. This indicates that the pure observation of non-compliant behavior could destroy the social norm of compliance and lead to non-compliance. The results also indicate that the shame effect is not strong enough to override the contagion effect when both effects are present simultaneously. The results are of particular importance for fiscal policy because public disclosure may lead to more evasion instead of less, due to motivational crowding-out of tax morale.

Who bears value added taxes? Evidence from Germany

Value added taxes are levied in all OECD member states with the exception of the United States, usually raising more than one-fifth of total fiscal revenues. Changing value added tax rates is a popular and often used tool in fiscal policy. In Germany, value added tax rates have been changed twice since 2007. First, in 2007 the regular rate was increased from 16 % to 19 % to enlarge fiscal revenues. Secondly, in 2010 the rate on accommodation services was cut from the regular rate of 19 % to the reduced rate of 7 % to advance economic growth after the financial crisis of 2007–2008.

Value added taxes are indirect taxes whose incidence should completely fall to final consumers. However, whether and to what extent tax incidence actually falls on consumers depends on the degree of tax shifting. From a theoretical perspective, it will completely fall to consumers on markets with perfect conditions. In the case of imperfect markets a general prediction cannot be made.

Even though most developed countries levy value added taxes and despite the fact that these taxes raise an important fraction of total fiscal revenues, empirical evidence on the incidence of value added taxes is scarce. In particular, there are no studies on tax incidence for Germany. Therefore, the main purpose of this paper is to estimate the incidence of value added taxes based on both the 2007 and the 2010 tax reforms. To this end, difference-in-differences estimation are applied to German consumer price data.

The results show that consumers have to bear about 44 % of the additional tax burden after the 2007 reform. In addition, there is incidence for tax shifting being dependent on the market competitiveness. As theoretically expected consumer shares are higher in more competitive markets (89 %) compared to less competitive markets (21 %). Finally, in the case of accommodation services prices respond asymmetrically to tax rate changes. After the tax rate increase in 2007 about 38 % of the additional tax burden were shifted to final consumers. In contrast, after the 2010 tax cut the relief is not shifted to consumers at all. Consumer shares are similar to those from previous empirical research for other European countries. Contrary, there is clearly less tax shifting compared to the United States sales tax.

These results are important for fiscal policy because politicians may use them to anticipate consumer price reactions to future tax rate changes beforehand. Moreover, it is possible to conclude useful counteractions if requested.

Zusammenfassung

Bei der vorliegenden Dissertationsschrift handelt es sich um eine kumulative Dissertation gemäß § 2 Promotionsordnung der wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover. Die Dissertation besteht aus vier einzelnen Beiträgen, die zusammen mit den jeweiligen Ko-Autoren und ihrem Publikationsstatus in Tabelle 2 aufgelistet sind. Im Folgenden werden die Beiträge zusammengefasst.

Titel	Ko-Autoren	Zeitschrift / Status
Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy	Kay Blaufus Jochen Hundsdoerfer Dirk Kiesewetter Joachim Weimann	Journal of Economic Psychology 35, 1–16
Perception of income tax rates: evidence from Germany	Kay Blaufus Jochen Hundsdoerfer Christian Sielaff Dirk Kiesewetter Joachim Weimann	European Journal of Law and Economics, online first, doi: 10.1007/s10657-013-9389-9
The effect of tax privacy on tax compliance – An experimental investigation	Kay Blaufus Philipp E. Otto	Im Begutachtungsprozess: Journal of Economic Behavior & Organization
Who bears value added taxes? Evidence from Germany	—	Arbeitspapier

Table 2: Überblick über die Forschungsbeiträge

Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy

In den vergangenen zwanzig bis dreißig Jahren verfolgte eine Vielzahl von Staaten eine Steuerpolitik, bei der nominale Steuersätze gesenkt und gleichzeitig die steuerliche Bemessungsgrundlage verbreitert werden ("tax-cut-cum-base- broadening"). Unter der klassischen Annahmen rational handelnder Individuen ist es für den Steuerpflichtigen unerheblich, ob der nominale Steuersatz oder (in gleichem Umfang) die Bemessungsgrundlage angepasst werden. Eine unterschiedliche Wahrnehmung ist hingegen möglich, wenn man zulässt, dass Individuen vereinfachende Entscheidungsheuristiken zur Vermeidung geistiger Anstrengungen nutzen.

Dieser Beitrag untersucht, ob Individuen bei der Wahl zwischen unterschiedlichen Besteuerungsoptionen, die sich ausschließlich im nominalen Steuersatz und der Höhe der steuer-

lichen Bemessungsgrundlage unterscheiden, vereinfachende Entscheidungsheuristiken verwenden, und ob die Verwendung solcher Entscheidungsheuristiken zu einer verzerrten Wahrnehmung von Änderungen des nominalen Steuersatzes im Vergleich zu Änderungen der Bemessungsgrundlage führt. Zu diesem Zweck werden eine Befragung ("Conjoint Analyse") mit 467 Erwerbstätigen, die in Bezug auf die Merkmale Alter, Bildung, Geschlecht und Einkommen repräsentativ für die deutsche Bevölkerung sind, und ein ökonomisches Laborexperiment mit 56 erwerbstätigen Personen durchgeführt. Dies ermöglicht die Kombination der jeweiligen Stärken beider Verfahren. Auf der einen Seite erlaubt die Conjoint Analyse das Ziehen einer großen repräsentativen Stichprobe. Darüber hinaus können Präferenzen im Rahmen der Conjoint Analyse auf Individualebene erfasst werden. Das ermöglicht eine Identifikation von Individuen, die ihre Entscheidung anhand einer Heuristik und nicht aufgrund einer rationalen Reihung treffen. Auf der anderen Seite handelt es sich bei einer Conjoint Analyse um eine reine Präferenzmessung, wohingegen ein Laborexperiment durch leistungsabhängige Vergütungen monetäre Anreize setzt. Darüber hinaus ermöglicht die Laborumgebung, Lerneffekte zu quantifizieren.

Unsere Ergebnisse befinden sich im Einklang mit anderen Studien, die Belastungen durch kognitive Anstrengungen berücksichtigen. Die Mehrzahl der Individuen entscheidet nicht rational anhand der tatsächlichen Steuerbelastung, sondern nutzt vielmehr vereinfachende Entscheidungsheuristiken. Als Folge daraus haben Veränderungen des nominalen Steuersatzes einen irrational hohen Einfluss auf die gefühlte Steuerbelastung. Steuerpflichtige bevorzugen Besteuerungsoptionen mit niedrigen Steuersätzen und breiter Bemessungsgrundlage gegenüber hohen Steuersätzen bei schmaler Bemessungsgrundlage, selbst wenn die Belastung in letzterem Fall niedriger ist. Darüber hinaus nimmt die Überschätzung von Veränderungen des Steuersatzes erheblich zu, wenn Informationen zum Steuersatz zuerst betrachtet werden ("Darstellungseffekte"). Auch bei leistungsabhängiger Vergütung überschätzen die Individuen Veränderungen des nominalen Steuersatzes. Allerdings sind geringe Lerneffekte erkennbar. Individuen greifen seltener auf Entscheidungsheuristiken zurück, aber selbst nach drei Entscheidungen erfolgt die Mehrheit der Entscheidungen noch immer nicht auf Basis einer rationalen Reihung.

Die Ergebnisse sind für die Fiskalpolitik von großer Bedeutung. Sie deuten an, dass Politiker die Staatseinnahmen erhöhen und gleichzeitig die gefühlte Belastung der Steuerpflichtigen senken könnten. Darüber hinaus erlauben Darstellungseffekte eine Senkung der gefühlten Belastung allein durch eine Fokussierung auf den nominalen Steuersatz.

Perception of income tax rates: evidence from Germany

Bereits 1960 hat Schmölders auf die Bedeutung der gefühlten Steuerbelastung für die Steuerpolitik hingewiesen. Anhand einer Befragung konnte er zeigen, dass nur ein Drittel der Befragten die eigene Steuerbelastung überhaupt kennt. Die ökonomische Steuerforschung arbeitet allerdings bis heute fast ausschließlich mit der tatsächlichen Steuerbelastung.

Dieser Beitrag untersucht, ob die Ergebnisse von Schmölders auch 50 Jahre nach seiner Untersuchung noch Gültigkeit besitzen. Analog zu Schmölders werden sowohl die gefühlte Steuerbelastung als auch die Beziehung zwischen der gefühlten und der für fair gehaltenen Steuerbelastung untersucht. Darüber hinaus erweitert dieser Beitrag die Arbeit von Schmölders durch eine detaillierte Analyse der Einflussgrößen von gefühlter und fairer Steuerbelastung. Zu diesem Zweck wurden mehr als 1000 Erwerbstätige befragt.

Auf Basis dieser Befragungsdaten wird die gefühlte Steuerbelastung anhand von vier vordefinierten Einkommenskategorien sowie des tatsächlichen Einkommens der befragten

Personen gemessen. Die gefühlten Steuersätze werden sowohl mit dem tatsächlichen Steuersatz als auch mit den für fair gehaltenen Steuersätzen verglichen. Dadurch kann festgestellt werden, ob die Steuerpflichtigen ihre tatsächliche Steuerbelastung akkurat einschätzen können. Darüber hinaus kann anhand der Angaben der Steuerpflichtigen untersucht werden, ob und inwieweit die gefühlte Steuerbelastung für fair gehalten wird.

Die Ergebnisse von Schmölders erweisen sich als zeitinvariant. Bei der Mehrheit der Befragten weicht die gefühlte Steuerbelastung signifikant von der tatsächlichen Steuerbelastung ab. Im Mittel nehmen die Befragten den Durchschnittssteuersatz der vier vordefinierten Einkommenskategorien ebenso wie den Grenzsteuersatzes des eigenen Einkommens bei niedrigeren Einkommen zu hoch und bei höheren Einkommen zu niedrig wahr. Das Ausmaß der Fehlwahrnehmung kann zumindest teilweise durch Bildung und Einkommen der Befragten erklärt werden. Darüber hinaus wird die Fehlwahrnehmung durch eine Berücksichtigung von Sozialversicherungsbeiträgen bei der Schätzung der Einkommensteuersätze verstärkt.

Ein Vergleich der gefühlten mit der für fair gehaltenen Steuerbelastung zeigt, dass die Steuerpflichtigen die Steuerlast für ein vordefiniertes Einkommen von EUR 300.000 und weniger als zu hoch erachten, wohingegen sie die Steuerlast bei einem vordefinierten Einkommen von EUR 2.000.000 als zu niedrig erachten. Besonders auffallend ist die Diskrepanz zwischen der für fair gehaltenen und der wahrgenommen Steuerbelastung auf der einen Seite und der tatsächlichen Steuerbelastung auf der anderen Seite. Die Steuerpflichtigen halten es für fair, Steuern bei einem vordefinierten Einkommen von EUR 10.000 zu senken und bei einem vordefinierten Einkommen von EUR 2.000.000 zu erhöhen. Im Gegensatz dazu übersteigt die für fair gehaltene Steuerbelastung bei einem Einkommen von EUR 10.000 die tatsächliche Steuerbelastung signifikant, wohingegen sie bei einem Einkommen von EUR 2.000.000 signifikant niedriger ist. Es ist somit von großer Bedeutung, ob steuerpolitische Empfehlungen auf der gefühlten oder der tatsächlichen Belastung von Steuerpflichtigen beruhen.

Die Analyse der Einflussgrößen zeigt, dass die Beurteilung einer Steuerbelastung als fair besonders von Bildung und Einkommen der Befragten abhängt. Die Wahrscheinlichkeit, dass ein Steuersatz für fair gehalten wird, steigt mit der Bildung der Befragten an. Jüngere Steuerpflichtige empfinden die Steuerbelastung besonders als zu hoch.

Schließlich deuten die Ergebnisse daraufhin, dass Steuerpflichtige üblicherweise aus Fairnessgründen einen progressiven Steuertarif bevorzugen.

The effect of tax privacy on tax compliance – An experimental investigation

Steuerhinterziehung ist ein Problem, dass alle Staaten gleichermaßen betrifft. Allein für das Steuerjahr 2006 beläuft sich beispielsweise das Steuerloch in den Vereinigten Staaten auf $385 Milliarden. Seit mehr als einem Jahrhundert wird darüber diskutiert, ob sich Steuerhinterziehung durch die Offenlegung von Steuererklärungsdaten wirksam bekämpfen lässt. Während die meisten Staaten Steuerdaten vertraulich behandeln, veröffentlichen Staaten wie Griechenland und Neuseeland Listen mit Steuerhinterziehern. Die nordischen Staaten legen darüber hinaus sogar sämtliche steuerlichen Informationen offen. Die Offenlegung soll Steuerpflichtige, aus Scham als Steuersünder an den Pranger gestellt zu werden, von Steuerhinterziehung abhalten. Es ist allerdings vollkommen unklar, ob eine solche Politik im Kampf gegen Steuerhinterziehung wirksam ist.

Auch soziale Normen können einen erheblichen Einfluss auf Steuerhinterziehung haben. Steuerpflichtige verhalten sich steuerehrlich, solange sie davon ausgehen, dass steuerehrliches

Verhalten der sozialen Norm entspricht. Allerdings ermöglicht die Veröffentlichung steuerlicher Informationen, unethisches Verhalten anderer zu beobachten. Das kann unter Umständen ansteckend wirken, da es die sozialen Normen in Bezug auf Steuerehrlichkeit verändern kann. Daher kann die Offenlegung Steuerhinterziehung sowohl eindämmen als auch verstärken. Der Gesamteffekt einer solchen Offenlegungspolitik auf die Einhaltung steuerlicher Regelungen ist aufgrund der gegensätzlichen Effekte – Scham auf der einen Seite, Ansteckung auf der anderen – unklar.

Dieser Beitrag untersucht den Einfluss der Offenlegung auf die Steuerhinterziehung experimentell anhand eines ökonomischen Spiels. Das Steuergeheimnis fungiert als Treatment-Variable, die drei verschiedene Formen annehmen kann und von vollständiger Geheimhaltung zu vollständiger Offenlegung reicht. Es gibt zwei Stufen der Offenlegung. Auf der ersten Stufe findet die Offenlegung anonym statt, wohingegen auf der zweiten Stufe Fotos der Teilnehmer zusammen mit ihrem steuerlichen Verhalten veröffentlicht werden. Aufgrund dieses Designs ist es möglich, den Scham- und den Ansteckungseffekt zu separieren, da bei einer anonymen Veröffentlichung zwar Ansteckungs-, jedoch keine Schameffekte auftreten können.

Insgesamt kann ein Anstieg der Steuerhinterziehung beobachtet werden, wenn steuerliche Informationen nicht vertraulich behandelt werden. Dieses Ergebnis deutet daraufhin, dass das reine Beobachten steuerunehrlichen Verhaltens soziale Normen zerstören und zu vermehrter Steuerhinterziehung führen kann. Darüber hinaus kann geschlossen werden, dass der Schameffekt nicht ausreicht, den Ansteckungseffekt aufzuheben, wenn beide Effekt gleichzeitig auftreten. Diese Ergebnisse sind von erheblicher Bedeutung für die Steuerpolitik, da eine Offenlegungspolitik aufgrund einer Verschiebung der Steuermoral Steuerhinterziehung verstärken kann.

Who bears value added taxes? Evidence from Germany

Mehrwertsteuern werden in sämtlichen OECD-Mitgliedsländern mit Ausnahme der Vereinigten Staaten erhoben. Das Steueraufkommen aus der Mehrwertsteuer beträgt in diesen Staaten regelmäßig ca. 20 % der gesamten Steuereinnahmen. Die Änderung von Mehrwertsteuersätzen ist ein beliebtes und oft genutztes steuerpolitisches Instrument. In Deutschland wurden die Mehrwertsteuersätze seit 2007 zweimal geändert. Zunächst wurde der reguläre Mehrwertsteuersatz zum 1. Januar 2007 zur Erhöhung der Staatseinnahmen von 16 auf 19 % angehoben. Schließlich wurde der Mehrwertsteuersatz auf Beherbergungsdienstleistungen zum 1. Januar 2010 vom regulären Satz von 19 % auf den ermäßigten Satz von 7 % gesenkt, um das Wirtschaftswachstum nach der Finanzkrise 2007–2008 zu beschleunigen.

Die Mehrwertsteuer ist eine indirekte Steuer, deren Inzidenz vollständig auf den Endverbraucher fallen sollte. Ob und inwieweit die Steuerinzidenz allerdings tatsächlich auf den Endverbraucher fällt, hängt vom Ausmaß der Steuerüberwälzung ab. Aus theoretischer Sicht findet auf Märkten mit vollständigem Wettbewerb eine vollständige Überwälzung auf den Endverbraucher statt. Auf unvollständigen Märkten ist eine generelle Prognose hingegen nicht möglich.

Obwohl die meisten Industrieländer Mehrwertsteuern erheben und diese einen erheblichen Anteil an den gesamten Steuereinnahmen dieser Länder haben, existieren nur sehr wenige empirische Studien zur Inzidenz der Mehrwertsteuer. Beiträge zur Inzidenz in Deutschland fehlen sogar gänzlich. Dieser Beitrag schätzt daher die Inzidenz der Mehrwertsteuer auf Basis der beiden Reformen 2007 und 2010. Zu diesem Zweck wird eine Difference-in-Differences-Schätzung mit deutschen Verbraucherpreisdaten durchgeführt.

Die Ergebnisse zeigen, dass die Endverbraucher ca. 44 % der zusätzlichen Steuerbelastung nach der Mehrwertsteuerreform 2007 tragen müssen. Darüber hinaus kann aus den Ergebnissen geschlossen werden, dass das Ausmaß der Steuerüberwälzung vom Marktwettbewerb abhängt. Der Konsumentenanteil ist wie erwartet in wettbewerbsintensiven Märkten höher (89 %) als in wettbewerbsschwachen Märkten (21 %). Schließlich zeigen die Ergebnisse, dass die Preise für Beherbergungsdienstleistungen asymmetrisch auf Steuersatzänderungen reagieren. Nach der Steuersatzerhöhung 2007 werden ca. 38 % der zusätzlichen Steuerlast auf den Endverbraucher überwälzt, wohingegen die Steuersenkung 2010 nicht an die Konsumenten weitergereicht wird. Die geschätzten Konsumentenanteile ähneln denen empirischer Studien anderer europäischer Länder. Im Gegensatz dazu ist die Überwälzung deutlich geringer als im Fall der US sales tax.

Die Ergebnisse sind von besonderer Bedeutung für die Steuerpolitik. Politiker können mithilfe der Ergebnisse die Reaktionen von Verbraucherpreisen bei zukünftigen Steuersatzänderungen antizipieren. Darüber hinaus können bei Bedarf sinnvolle Gegenmaßnahmen beschlossen werden.

Contents

Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy **1**

Perception of income tax rates: evidence from Germany **18**

The effect of tax privacy on tax compliance – An experimental investigation **20**

Who bears value added taxes? Evidence from Germany **39**

Journal of Economic Psychology 35 (2013) 1–16

Contents lists available at SciVerse ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy

Kay Blaufus [a,*], Jonathan Bob [a], Jochen Hundsdoerfer [b], Dirk Kiesewetter [c], Joachim Weimann [d]

[a] Leibniz Universität Hannover, School of Economics and Management, Königsworther Platz 1, D-30167 Hanover, Germany
[b] Freie Universität Berlin, School of Business & Economics, Garystr. 21, D-14195 Berlin, Germany
[c] Julius Maximilian University of Würzburg, Faculty of Economics and Management, Sanderring 2, D-97070 Würzburg, Germany
[d] Otto-von-Guericke University Magdeburg, Faculty of Economics and Management, Universitätsplatz 2, D-39016 Magdeburg, Germany

ARTICLE INFO

Article history:
Received 18 June 2012
Received in revised form 13 December 2012
Accepted 16 December 2012
Available online 7 January 2013

JEL classification:
D03
G11
H20
H30
K34
M41

PsycINFO classification:
2340
2960

Keywords:
Behavioral public finance
Decision heuristics
Framing effects
Perceived tax burden
Tax-cut-cum-base-broadening

ABSTRACT

In this paper, both a conjoint analysis and a lab experiment are conducted to analyze the influence of changes in the tax rate and the tax base on the perceived tax burden. Our results show that the majority of individuals do not make rational tax decisions based on the actual tax burden but rather use simple decision heuristics. This leads to an irrationally high impact of changes in nominal tax rates on the perceived tax burden. Taxpayers favor tax options that apply a lower tax rate on their gross income over a higher tax rate applied on their net income despite the lower actual tax burden of the latter option. This result suggests that politicians could combine increasing fiscal revenues and decreasing subjects' tax perception. Furthermore, overestimation of tax rate changes increases considerably when information on tax rate is considered first (framing effect).

1. Introduction

Following the United Kingdom and the United States, in recent decades, many countries pursue a policy of tax-cut-cum-base-broadening (OECD, 2010: 11), i.e., they reduce the nominal tax rate and simultaneously increase the tax base. One recent example is the introduction of a final withholding tax on interest income in many OECD countries. These countries cut the nominal tax rate on capital income but abolish the possibility to deduct capital income-related expenses (Genser & Reutter, 2007). Moreover, several countries provide explicit tax options that reflect the above mentioned tax-cut-cum-base-broadening policy. Exemplarily, small corporations in Russia may choose between a lower tax rate on their gross

* Corresponding author. Tel.: +49 511 7625660; fax: +49 511 76219858.
E-mail address: blaufus@steuern.uni-hannover.de (K. Blaufus).

0167-4870/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.joep.2012.12.004

income and a higher tax rate on their net income. In France and Germany taxpayers may—under certain conditions—choose whether their dividends are taxed under a progressive tariff based on their net income or under a flat tax based on their gross income. The aforementioned tax reforms are usually reasoned by expected efficiency gains and tax simplification (e.g., Kopczuk, 2005). We propose a behavioral approach as an additional explanation as to why politicians pursue such a policy of tax-cut-cum-base-broadening: the use of decision heuristics that lead to an irrationally high impact of nominal tax rate reductions on taxpayers' perceived burden.

Previous economic and psychological research already shows that attitudes toward tax policies are subject to several biases (McCaffery & Baron, 2004). As early as 1960 Schmölders finds evidence that a high percentage of entrepreneurs overestimate their tax burden (Schmölders, 1960: 84f). Especially, if tax complexity is high or if taxes are non-salient, subjects make substantial decision errors (Boylan & Frischmann, 2006; Chetty, Looney, & Kroft, 2009; Rupert, Single, & Wright, 2003; Rupert & Wright, 1998). In addition, research shows that price complexity influences buyers' price perceptions and demands. Breaking down the price into several components (e.g., base price and shipping costs) leads to a decrease in the perceived price and an increased demand for the corresponding commodity (see Morwitz, Greenleaf, & Johnson, 1998). Krishna and Slemrod (2003) discuss the potential meaning of these price research results for tax policy and Chetty et al. (2009) find that using prices plus sales tax instead of the net amounts leads to a significant reduction in demand. The following article is also based on this idea and is, to our knowledge, the first to examine whether the perceived tax burden is dependent upon which price component (tax rate or tax base) is changed.

The calculation of the tax burden is regarded as a complex task (Kirchler, 2007: 32). As we know from previous research decision heuristics are used in such complex tasks, particularly. Therefore, it seems reasonable to assume that subjects will use heuristics in our tax framing. Even though there are papers on the use of decision heuristics in tax contexts (e.g., McCaffery & Baron, 2003), a paper on the relationship between decision heuristics and a tax-cut-cum-base-broadening policy is missing.

The main purpose of this study is to examine whether the use of decision heuristics leads to a misperception of changes in nominal tax rates compared to changes in tax deductions. We analyze (i) whether decision heuristics are used in choosing between tax systems that differ, exclusively, in the nominal tax rate and the amount of tax deductible expenses, (ii) whether the use of decision heuristics will induce subjects to overemphasize the importance of nominal tax rate changes relative to changes in the amount of tax deductible expenses, and (iii) whether framing effects drive the misperception of nominal tax rate changes compared to tax base changes.

To this aim, we conduct (i) a survey study (conjoint analysis) with a sample of German working individuals that represents the population in terms of age, education, gender, and income and (ii) a lab experiment with working individuals who are compensated according to their experimental performance. By combining these two methods we are able to benefit from the respective strengths of both procedures. First of all, conjoint analysis allows us to draw a large (representative) sample. Moreover, we are able to measure preferences on the individual level. This incorporates the advantage of conjoint analysis to identify whether individuals choose a decision heuristic over the rational ranking. On the other hand, conjoint analysis is a mere preference measurement whereas a lab experiment sets monetary incentives. Besides, while in conjoint analysis the task is performed just once, the lab experiment enables us to execute the task multiple times.

The remainder of this article is organized as follows. First, in Section 2 we discuss the effects that decision heuristics may have on the perceived importance of nominal tax rate changes and derive our hypotheses. The sample, method, and results are presented in Sections 3 and 4, "Conjoint Analysis" and "Experimental Validation", respectively. A discussion of the results and the implications for tax policy and research are carried out in Section 5.

2. Hypotheses

We consider hypothetical tax options and put subjects in the position of a choice between these alternatives that, for a given income, differ only in terms of the nominal tax rate and the allowed deduction of income-related expenses.[1] This setting mirrors the often proposed policy option to eliminate itemized deductions (e.g., Pitt & Slemrod, 1989; Slemrod, 1989). Opposition to itemized deductions is in line with public opinion as 54% of Americans would be willing to give up some deductions to make the tax system simpler (Tax Foundation, 2005). According to Elffers and Hessing (1997) subjects favor higher standard deductions over itemized deductions because taxpayers "prefer being lazy to becoming tired". However, our main purpose is to study whether the use of decision heuristics leads subjects to misperceive nominal tax rate changes compared to tax base changes. We are not interested in the preferences for tax simplification. To control for such preferences subjects are explicitly informed that each tax system from which they can choose leads to the same compliance effort. Therefore, rational decision makers choose among these alternatives based on the actual tax burden B_i, which can be written for the ith alternative as

$$B_i = \tau_i Y - \tau_i D_i \tag{1}$$

where τ_i is the tax rate, D_i is the deduction of income-related expenses, and $Y > D_i$ is taxable revenue (identical for all alternatives).

[1] It is important to note that we are not looking at tax reform models that differ such that certain economic activities are taxable in one alternative but not the other. As is well known from literature (see e.g., Willner & Granqvist, 2002) a base-broadening, rate-reducing policy that taxes previously untaxed opportunities while reducing the tax rate on all taxable opportunities could lead to efficiency gains even if tax payments remain the same. By contrast, we are studying the effect of taxing a single (already taxable) economic activity differently in each tax option.

Traditional theory assumes that economic subjects do not make arithmetic errors and that the calculations included in the decision-making process do not require a great deal of cognitive effort.

In contrast, psychological research shows that a human's cognitive ability is limited and that the calculation of decision variables causes cognitive strain, making it reasonable for individuals to use simple decision heuristics rather than exact calculations (see Tversky & Kahneman, 1974). Using a decision heuristic may lead to a suboptimal solution that is, however, individually satisfactory (Simon, 1990). Particularly in tax settings, it seems reasonable to assume that subjects will use heuristics. The calculation of the tax burden is regarded as a complex task and usually causes enormous cognitive strain. Kirchler summarizes: "People blame the complexity of the tax law for their feelings of tax incompetence." (Kirchler, 2007: 32). In the light of tax complexity the advantage of selecting a decision heuristic consists in lower cognitive strain and less time spent on making the decision.

This leads to our first hypothesis to be tested:

H1. The decision among alternatives is not based exclusively on the actual tax burden but rather on the use of decision heuristics.

One reasonable heuristic is what is known as the anchor heuristic. This heuristic has been observed empirically in several other contexts (Epley & Gilovich, 2006; McCaffery & Baron, 2003; Tversky & Kahneman, 1974). According to the anchor heuristic, individuals who estimate a certain value, such as the actual tax burden, begin with a starting value that serves as an anchor for finding the estimated value. The disadvantage of this heuristic is that the anchor is consistently overweighted, and additional information is not adequately included.

The anchor is often chosen by selecting either the information with which the individual is first confronted (see Hogarth & Einhorn, 1992) or the information which is considered most important (Yadav, 1994). The anchor value is then adjusted (inadequately) based on later information or information that is considered to be less important.

In our opinion, there are several reasons why the tax rate and not the amount of deductible expenses serves as the anchor.

First, information about the nominal tax rate τ_i is, in reality, more readily available than information about the deductibility of single expenditures. According to the availability bias, people overweight evidence that is easily available (Tversky & Kahneman, 1973: 211).

Second, in general, tax liability reacts more elastically to changes in the tax rate than to changes in deductions. A 1% increase in tax rate always leads to a 1% increase in tax liability. However, a 1% decrease in deductions, usually, leads to a tax increase of less than 1%.[2] Thus, the amount of income-related expenses to be deducted could be considered less important by some individuals.

Third, the effects of different tax rates on a given income can be easily recognized such that individuals determine the positive relation between tax rate and income without cognitive strain. By contrast, with the influence of the deduction of income-related expenses, there is a negative relation between the tax base and income-related expenses as well as a positive relation between the tax base and tax liability. The marginal necessary cognitive effort compared to the tax rate effects, thus, supports the assumption that the tax rate, and not income-related expenses, serves as an anchor.

In addition to the anchor heuristic, the use of a lexicographic heuristic is reasonable (e.g., Brandstaetter, Gigerenzer, & Hertwig, 2006). Individuals reduce their cognitive effort by first evaluating the alternatives based on only one criterion, and if no decision is possible with this criterion, they apply other criteria. Based on the high relevance of the tax rate presumed above, we assume that individuals who use a lexicographic heuristic first evaluate alternatives based on the tax rate; only if information regarding tax rates is equivalent they consider deductible expenses.

The use of the mentioned heuristics in combination with the assumption of the perceived high relevance of the tax rate leads to the following hypothesis:

H2. The effect of changes in the tax rate (changes in the tax base) is overestimated (underestimated) compared to the rational benchmark.

Extensive behavioral and psychological research shows that human behavior is subject to framing effects (McCaffery & Baron, 2004; Tversky & Kahneman, 1981). Presenting the same option in different ways can alter subjects' decisions. One example is that individuals tend to choose the piece of information with which they are first confronted as an anchor or as the primary criterion for the lexicographic heuristic (Blaufus & Ortlieb, 2009; Chrzan, 1994; Moran & Meyer, 2006). Thus, we hypothesize:

H3. Confronting subjects first with the tax rate increases the likelihood that subjects overestimate the importance of changes in the tax rate.

3. Conjoint analysis

3.1. Method

To test the hypotheses H1–H3 formulated in the previous section, we conduct a conjoint analysis. Conjoint analysis is based on Luce and Tukey (1964). The aim of this method is to derive the influence of attributes (and their levels) on the total

[2] The elasticity equals $\frac{D_i}{Y-D_i}$ and is always less than one percent if $Y > 2D_i$, which is typically the case.

utility of a combination of attributes (stimuli). For this purpose, subjects are given various stimuli to evaluate. Conjoint analysis is a decomposition method in which the estimation of influence (part-worth utilities of attribute levels, relative importance of attributes) is based on empirically collected total utility of the respective stimuli (see Hair, Black, Babin, & Anderson, 2008).

The subjects' task in conjoint analysis is to rank several of these stimuli according to their personal preferences.[3] The subjects assign the lowest (highest) rank to the stimulus with the highest (lowest) preference. The total utility of the stimuli is derived from the individual ranking by each subject. For this purpose, the stimulus with the lowest (highest) rank is assigned the highest (lowest) utility. Metric part-worth utility for the attribute levels is determined using the calculated total utility and the ordinary least squares method.[4] As is standard in conjoint analysis, we assume an additive model for the relation between total utility and part-worth utility, i.e., the sum of the part-worth utilities of a stimulus corresponds to its total utility. This follows from the basic assumption that the explanatory variables do not interact (see Hair et al., 2008).[5] This leads to the following relationship:

$$U_{i,j} = \mu_i + \sum_{k=1}^{K} \beta_{k,i} x_{k,j} + \varepsilon_j \tag{2}$$

where $U_{i,j}$ represents the total utility of the jth stimulus for the ith subject. μ is the constant, and $\beta_{k,i}$ are the part-worth utilities of the attribute levels. The dummy variables $x_{k,j}$ take on a value of one if the observed stimulus contains the respective attribute level. ε_j is the error term.

As can be seen in Eq. (2), the part-worth utilities are estimated on the subjects' level. To compare and aggregate the part-worth utilities among subjects, they have to be standardized (Green & Krieger, 1985, 3f). For this purpose, the highest total utility is set to the value of one while the lowest total utility is set to zero. Hence, the sum of the most preferred level of each attribute is one, whereas standardized part-worth utilities of the fewest preferred levels are all set to zero. The relative importance of each attribute equals the part-worth utility of the most preferred level for this attribute. The more the total utility of a stimulus changes when the level varies for a given attribute, the higher is the relative importance.

Conjoint analysis is predominantly used in marketing research. However, conjoint analysis has also been used to measure tax effects (e.g. Blaufus & Ortlieb, 2009; Hundsdoerfer & Sichtmann, 2009; Milliron & Toy, 1988; O'Neil, 1982). The idea is to define tax characteristics (e.g. tax rate and deductible expenses) as attributes of products. Traditional conjoint analysis[6] allows to estimate the relative importances of these tax characteristics at the subject level (see Green & Srinivasan, 1978: 104) and compare the measured importances with the importances for a "rational" taxpayer. For this purpose, the part-worth utility of a tax characteristic for a "rational" tax payer will be compared with the actual measured part-worth utility. Thus, we can identify whether subjects misperceive the impact of changes in specific tax characteristics like the nominal tax rates.

A further advantage of conjoint analysis is the simultaneous evaluation of the attributes. Alternatively, one could directly ask subjects for the value they attach to an attribute. This sequential evaluation has the disadvantage that subjects tend to neglect the trade-off effects. All attributes are considered to be very important and hence, the importance of the individual attribute is overrated. Contrary, because of the simultaneous evaluation of the attributes in conjoint analysis subjects must keep in mind the trade-off effects between attributes which also exist in reality.

3.2. Sample

A total of 467 working individuals are interviewed who match the population in terms of the following attributes: gender, age, education, and monthly net income. The selection of working individuals ensures that subjects have experience with income taxation.

Trained interviewers conduct standardized face-to-face interviews that last an average of 20 min. In addition to conjoint analysis, subjects are asked questions regarding demographic attributes, general attitudes toward tax policy, current German income tax law, and tax complexity.

The sample is drawn based on a quota schedule,[7] as a pure random sample was not feasible financially. The quota parameters are based on the following four attributes: age, gender, education, and monthly net income. The corresponding frequency

[3] This ranking is the most common valuation procedure after the rating scale. An overview of various procedures is given in Green and Srinivasan (1978, 104).

[4] The subjects' preference judgments, expressed through their rank ordering, have an ordinal measurement level. Hence, one can apply a monotone variance analysis. However, the least squares method has proven to be very robust in the estimation of part-worth utility values also in the case of ordinally scaled dependent variables (see e.g. Green & Krieger 1993, 478).

[5] Eq. (1) leads to an (at least theoretical) interaction between the two attributes tax rate and income-related expenses. However, by assuming the rational valuation of a homo economicus, this is a no-crossover interaction (see Green & Devita, 1974, 56). Thus, in line with previous research, the interaction can be neglected in the following analysis (see Carmone & Green, 1981, 93).

[6] Other conjoint analysis procedures (e.g. Choice Based Conjoint Analysis) do not allow for an estimation of preferences on the individual level and are therefore unsuitable for the purpose of our study. Related models for preference measurement, such as the Rank-Ordered Logit, allow for estimation on the individual level but require subject specific attributes in addition to the attributes of the stimuli, which distinguish the evaluation of the stimuli attributes by the subjects. See Allison & Christakis, 1994, 202.

[7] Quota samples do not strictly fulfill the requirements of a pure random selection. Nevertheless, it is the most widely used procedure in marketing research and continuously yields good results in comparative studies with pure random selection (see Green, Tull, & Albaum, 1988, 325–327).

A	Tax rate: 30 % Deduction for income-related expenses: € 0
B	Tax rate: 30 % Deduction for income-related expenses: € 2000
C	Tax Rate: 35 % Deduction for Income-Related Expenses: € 0
D	Tax Rate: 25 % Deduction for Income-Related Expenses: € 2000
E	Tax Rate: 35 % Deduction for Income-Related Expenses: € 2000
F	Tax Rate: 25 % Deduction for Income-Related Expenses: € 0

Fig. 1. Stimuli used in conjoint analysis.

in the population is taken from the Statistical Yearbook of the Federal Statistical Office in Germany, which covers the 37 million people who make up Germany's working population.

Compliance with the quota is statistically tested by a chi-squared test. With a margin of error of 5%, no significant difference between population and sample can be detected. In this respect, the sample can be seen as representative of the working population in Germany. The distribution of attributes in the sample is given in Table 11 Appendix A.

Of the 467 polled individuals, 33 favor high tax rates and non-deductibility of income-related expenses. Because the sincerity of this stated preference is doubtful,[8] these individuals are not included in the analysis. The adjusted sample, therefore, includes 434 individuals who also match the population in terms of gender, age, education, and monthly income.

3.3. Research design and operationalization of the hypotheses

Subjects are asked to rank various tax systems according to their individual preferences. For this purpose, subjects are asked to assume they receive taxable earnings (an interest payment) of € 10000 and bear income-related expenses of € 2000. The individual tax systems differ solely in terms of tax rate τ_i and allowable deductions of income-related expenses D_i and, therefore, in terms of tax burden B_i, which can be calculated using Eq. (1).

Interviewers explain this income should be considered as the respondents' own income, and that every tax system leads to the same compliance effort ("time required for the tax return").

The stimuli are presented using the Full Profile Method, i.e., each stimulus exhibits a combination of the two attributes (tax rate and deduction of income-related expenses). Three levels are chosen to express the tax rate (low, medium, high), and two levels are selected for income-related expenses (no deduction, full deduction). This combination of levels yields a maximum of six (3×2) different stimuli.[9] The corresponding complete design is presented in Fig. 1.[10]

The individual stimuli are given to the subjects in the form of randomly ordered laminated cards. The random issuance of the stimuli ensures that the order of presentation has no influence on the valuation (regarding the "sequencing effect" see e.g., Tourangeau & Rasinski, 1988: 304).

Subjects are asked to arrange the cards on a magnetic board in previously numbered positions according to their preferences. Before the subjects begin, the interviewer explains the terms "tax rate" and "income-related expenses" as well as their

[8] For instance, some subjects simply ranked the stimuli alphabetically. An examination showed that the exclusion of these so called reversals had no influence on the presented results.

[9] In addition to the two attributes of tax rate and tax base, the stimuli contained a third attribute ("time required for the tax return"), which is not relevant in the current study, and which only serve as a control in the current study. No significant differences can be detected between the stimuli used and the control stimuli.

[10] Two different stimuli sets are used to test whether the amount of tax rates and income-related expenses influences subject's decision. These two sets differ in the differences between the particular tax rates and expenses. The values on the second stimuli set amount to 25%, 27%, and 29% (tax rate) as well as € 0 and € 1000 (income-related expenses). Of course, rational rankings stay unchanged.

Table 1
Tax burden.

Stimulus	Tax burden (€)
A	3000
B	2400
C	3500
D	2000
E	2800
F	2500

effects on the tax burden using the numbers given in the actual decision task. Furthermore, subjects are made aware of the "objectively best tax alternative", which is shown on card D (see Fig. 1).[11] They then have to rank only the remaining alternatives according to their given income. After the subjects conduct the rankings, the interviewer offers the subjects a chance to review their choices. After the subjects made their final choices, the interviewer records the final preference ranking.

The subjects' rankings serve to test whether the individuals conduct the rankings rationally—according to the actual tax burden—or whether they use decision heuristics (hypothesis H1).

To test whether changes in tax rates are overestimated (hypothesis H2), the ranking is used to estimate part-worth utilities and relative importances. To this aim, we compare the results of our sample with the rational benchmark of the "homo economicus" and consider the relative importance to be overrated if and only if it exceeds the rational benchmark significantly.

For testing hypothesis H3, we use a between-subject design with the independent variable "framing effect". In the first group the tax rate is the first information mentioned, while in the second group it is mentioned last. The stimuli used in the first of the two groups are given in Fig. 1 (tax rate mentioned first). The stimuli used in the second group are the same except for the permutation of the item order.

Subjects are randomly divided into two groups. The 190 subjects in the first group are first presented with the attribute "nominal tax rate", the 244 subjects in the second group are first presented with the attribute "tax deductible expenses".

3.4. Results

3.4.1. Analysis of ranking behavior (Hypothesis H1)

To test hypothesis H1, we analyze whether the empirically observed rank order of the stimuli matches the prediction of the rational model. "Calculating" rational investors minimize their tax burden. Tax burdens of the six stimuli can be calculated using Eq. (1) and are shown in Table 1.

The rational ranking is given as follows:

$$\mathrm{D} \succ B \succ F \succ E \succ A \succ C.$$

To reduce cognitive strain, the individuals may use a simple heuristic instead of computation. In Section 2 the anchor heuristic and lexicographic heuristic are highlighted. If individuals rank lexicographically first according to tax rate and subsequently according to income-related expenses, then the following ranking, which differs from the rational ranking, results:

$$\mathrm{D} \succ F \succ B \succ A \succ E \succ C.$$

Although we predict that the lexicographic ranking uses the tax rate as the first criterion, it should be tested whether there are individuals whose lexicographic ranking uses income-related expenses as the primary decision criterion. If the stimuli are lexicographically ranked first according to deduction of income-related expenses and then to tax rate, we would expect the following ranking results:

$$\mathrm{D} \succ B \succ E \succ F \succ A \succ C.$$

We choose appropriate values permitting a clear distinction between rational and lexicographic ranking. An overview of the proportion of rankings in the sample is shown in Table 2.

From Table 2, it is apparent that a surprising number, more than 90% of the subjects, decide against the rational "homo economicus" ranking. Only 9.4% (41 individuals) follow the predictions based on the model of rational net income maximization.[12] Hypothesis H1 is, therefore, confirmed.

Using a logistic regression, we test for demographic factors as age, gender, education, tax knowledge, and income to distinguish between the different groups of subjects, i.e. those subjects who rank their stimuli rational, those who use a lexicographic heuristic and those who prefer other rankings. Tax knowledge was measured in three different ways. First,

[11] The indication of the objectively best alternative, and in the case of the six control cards (see Fn. 10) also of the objectively worst alternative, served to reduce the work of the subjects.

[12] The probability of randomly achieving a rational sequence is 0.83%. A binomial test shows that the percentage of the rational sequence cannot be the result of pure random selection ($p < 0.01$).

Table 2
Percentage of subjects per ranking.

Sequence	Subjects (%)
Rational	9.4
Lexikographic	
1st tax rate, 2nd income-related expenses	35.0
1st income-related expenses, 2nd tax rate	21.7
Other sequences	33.9
Total	100.0

Table 3
Logistic regression: determinants of a rational ranking.

Dependent variable: rational sequence (dummy)	β coef.	Std. error	β coef.	Std. error	β coef.	Std. error
Constants	−2.217***	0.709	−2.144***	0.688	−2.136***	0.692
With university-entrance qualification	0.193	0.356	0.187	0.358	0.28	0.352
Above average knowledge of tax law (self-assessment)					0.190	0.523
Very good knowledge of tax law (income tax rates)			1.236	0.565		
Tax return self-prepared	0.806**	0.374				
Gender	−0.153	0.344	−0.073	0.341	−0.075	0.346
Age	−0.011	0.015	−0.008	0.014	−0.007	0.014
Self-employed	−0.046	0.531	−0.033	0.537	−0.133	0.532
Net income < 1000 €/month	0.422	0.377	0.233	0.367	0.260	0.365
N	430		430		430	
Nagelkerke's R2	0.031		0.015		0.012	

β coef., β coefficients; Std. error, Standard error.
** Indicates significance on the 5% level.
*** Indicates significance on the 1% level.

subjects' were asked to self-assess their tax knowledge. Second, we asked subjects questions on the current actual income tax tariff. Third, we asked whether subjects self-prepare their tax return.[13] As it can be seen from Table 3, only tax knowledge could explain the subject's decision. Participants who self-prepared their last tax return (who have very good knowledge of tax law) choose around twice (three times) as likely the rational ranking. All other demographic variables are not significant.

Table 2 also shows that more than half of the subjects rank the stimuli lexicographically. One-third of all subjects order the tax systems lexicographically with tax rate the primary criterion. Contrary to our assumption, 21.7% of the individuals use a lexicographic heuristic in which income-related expenses are the dominant criterion. The probability of randomly achieving one of the two lexicographic sequences is 1.7%. Of the group of 434 subjects, seven could have arrived at a lexicographical sequence by randomly ordering the stimuli. Thus, one can assume that the two lexicographical heuristics are consciously chosen.[14]

These explanations point out that we can easily explain the behavior of two-thirds of our respondents. However, one-third chose a ranking that was neither rational nor lexicographical. In the next subsection, we show that the use of the anchor heuristic seems to be the most reasonable explanation for the remaining rankings.

3.4.2. Overestimation of the importance of changes in tax rate (Hypothesis H2)

According to hypothesis H2, the use of heuristics leads to an overestimation of the relative importance of the tax rate. Our approach is to estimate part-worth utilities by Eq. (2) using OLS and to compare the results of our sample with the rational benchmark of the "homo economicus"

$$\widehat{U}_{i,j} = \hat{\mu}_i + \hat{\beta}_{1,i}x_{1,j} + \hat{\beta}_{2,i}x_{2,j} + \hat{\beta}_{3,i}x_{3,j}.$$

$x_{1,j}$ to $x_{3,j}$ are dummy variables for the low tax rate, the medium tax rate and the full deduction of income-related expenses, respectively.

[13] The associated questions are:

1. How do you assess your own knowledge regarding tax law? Possible answers are "No knowledge", "Some basic knowledge", and "Good or very good knowledge".
2. How high do you rate the income tax burden in percent of the respective annual income? Respective incomes are: € 10000; € 40000; € 300000; and € 2000000. We compare subjects' answers with the actual tax burden according to the income tax tariff and compute the absolute value of subjects' error for every income category. The 5% of our subjects with the lowest total error about all four income categories are marked with "very good knowledge of tax law".
3. Who prepared your last tax return?Possible answers are: "On my own", "Someone else in the household", "A tax advisor", "Did not file a tax return", "Other".

[14] In this case, a binomial test also confirms that the percentage of lexicographical sequences cannot be the result of pure random selection ($p < 0.01$).

Table 4
Part-worth and relative importance (sample and rational sequence).

	Estimated part-worth		Relative importance	
	Rational	Sample	Rational	Sample
Low tax rate	0.5625	0.6331	0.5625	0.6331
Medium tax rate	0.2812	0.3166		
High tax rate	0.0000	0.0000		
High deduction of income-related expenses	0.4375	0.3669	0.4375	0.3669
Low deduction of income-related expenses	0.0000	0.0000		

The difference between the samples' relative importance of the tax rate (0.6331) and the rational value (0.5625) is significant as a Mann–Whitney-U test shows ($p < 0.01$).

Table 5
Relative importances of tax rate.

Tax rate named first	Tax rate named last
0.6996	0.6021

Both values are compared to the rational value of 0.5625. The differences are significant according to a Mann–Whitney-U test ($p < 0.01$).

The standardized part-worth utilities, as well as the relative importance that result from the sample of working individuals, are shown in Table 4, along with the part-worth utilities and relative importance of a "homo economicus" which serve as the rational benchmark (Hundsdoerfer, Sielaff, Blaufus, Kiesewetter, & Weimann, in press).

As it is shown in Table 4, the relative importance of tax rate changes is 12.5 percentage points higher than the importance of tax base changes. Therefore, tax rates have a stronger impact on the tax burden than tax deductions. However, by evaluating individuals' preferences we only consider the relative importance of the tax rate to be overrated if it exceeds the rational benchmark significantly. Table 4 shows that the part-worth utility of the low tax rate and the relative importance of the attribute nominal tax rate are higher in the sample. Hence, a change in the nominal tax rate results in a greater change of total utility. Whereas for a rational agent we observe a relative importance of 56% for the tax rate, the corresponding relative importance of the subjects is on average 63%. This difference is highly significant (Mann–Whitney U, $p < 0.01$).[15] Hypothesis H2 is, therefore, confirmed: the importance of changes in tax rates is overestimated, and the importance of changes in the tax base is underestimated.

However, it must be noted that the degree of overestimation can vary widely. The relative importance of the tax rate for individuals who conduct rankings lexicographically based primarily on the tax rate can amount to 80%, whereas the relative importance for the tax rate for other types of ranking only amounts to 63%. In addition, 32.5% of the individuals underestimate the relative importance of tax rates. This includes primarily those who rank the stimuli first according to income-related expenses and, therefore, attach a relative importance of only 40% to tax rates.

One should note the relative importance of the tax rate of 63% for the "other sequences" rankings. These individuals significantly overestimate the relative importance of the tax rate. As we pointed out in the last section, one approach to explain the "other sequences" rankings is the anchor heuristic. By thinking of individuals who use this heuristic and choose the tax rate as an anchor, one should expect a relative importance that is below the importance of lexicographic rankings (those who rank first according to the tax rate) but above the importance of the rational ranking. This is exactly what we observe in this case. Moreover, it becomes obvious that these subjects do not randomly rank the stimuli because the relative importance of the attributes differs significantly from 50% (sign test, $p < 0.01$).

3.4.3. Framing effects (Hypothesis H3)

To identify factors that explain the importance of the tax rate, we test whether framing effects increase the misperception of nominal tax rate changes. Therefore, we compute the relative importance for the two respective groups (tax rate mentioned first/tax rate mentioned last). The framing effect can be quantified as the difference in relative importance for the two settings "tax rate mentioned first" and "tax rate mentioned last". The resulting values are shown in Table 5 below.

As it can be seen from Table 5, the importance of the tax rate is overestimated and the importance of the expense deduction is underestimated in both groups. The deviations from the rational conclusion are highly significant (Mann–Whitney U, $p < 0.01$).

One can also see from Table 5 that the overestimation of the importance of the tax rate is significantly larger (Mann–Whitney U, $p < 0.01$) when the subjects are presented the tax rate first. The relative importance of the tax rate is 70% when

[15] In the following, solely the relative importance of the attribute "nominal tax rate" will be examined, as this can be used to derive all further values. The relative importance values add up to one. Hence, the relative importance of the attribute "allowable deduction of income-related expenses" is given by: 1 minus the relative importance of the "nominal tax rate." In addition, the largest standardized part-worth utility of an attribute level always corresponds to the relative importance of this attribute. The standardized part-worth utility of the middle tax rate can be calculated as half of the relative importance of the tax rate.

Table 6
Choice between tax options.

Position	Stimulus	Choices (%)	Tax burden (€)
1	F	55.10	2500
2	B	44.40	2400
3	E	0.50	2800
4	A	0.00	3000
5	C	0.00	3500

Total percentages according to the first choice model.

mentioned first and only 60% when mentioned last. Thus, hypothesis H3 is confirmed. The framing effect amounts to 10 percentage points. As we mentioned in section "Hypotheses" individuals tend to choose the piece of information with which they are first confronted as an anchor or as the primary criterion for the lexicographic heuristic. This explains the large framing effect. The percentage of subjects with a lexicographical ranking with tax rate as first criterion increases from 25% to 48% when the tax rate is mentioned first. In the same way the percentage of subjects with a lexicographical ranking with tax deductible expenses as first criterion increases from 13% to 28% when the tax rate is mentioned last.

3.4.4. Choice between tax options

The subjects are presented with the various alternatives as possible options for tax policy. Since we focus on tax changes only, we implicitly consider other political factors such as environmental policy, economic policy, or health care as constant among the alternatives. In our view, this assumption is particularly appropriate in situations where taxpayers can choose between different tax options as long as the choice does only affect themselves. Basically, in our setting this choice is whether to tax net income at a higher tax rate or gross income at a lower tax rate. This reflects for example the alternatives of smaller corporations in Russia. These corporations can choose between a tax rate of 6% on their sales (gross income) and a tax rate of 15% on their earnings (net income). Besides, there are several other countries as France and Germany that provide similar tax options with respect to dividend income. In these countries taxpayers can choose between the usual progressive tariff applied on their net dividend and a flat tax on their gross dividend.

To examine how taxpayers will choose between such tax options, a further assumption must be made about the relation between the ranking order (preference) of the individual and the individual's actual choice. We consider a deterministic model ("first choice"). Thus, the probability of choosing the most strongly preferred stimulus is one. All other stimuli have a choice probability of zero. The total percentage of a tax option corresponds to the number of subjects with first preference for this respective tax option divided by the total number of subjects. The percentages using the first choice model are shown in Table 6.[16] It shows that the absolute majority of choices are allotted to stimulus F even though this does not have the lowest actual tax burden. If the first choice model accurately describes choice behavior, then politicians could combine increasing tax revenues with a decrease in the perceived tax burden. This can be achieved by introducing tax options into tax law that allow taxpayers to decide between a higher tax rate on their net income and a lower tax rate on their gross income. Obviously, taxpayers could avoid wrong choices by consulting tax advisors. However, according to our data only 22% of the subjects actually consult a tax advisor for their last tax return.

4. Experimental validation

As we pointed out before, conjoint analysis comes with its own strengths and weaknesses. On one hand, we have the advantages of drawing a large (representative) sample, measuring preferences on the individual level and identify decision heuristics. On the other hand, conjoint analysis in practice does not allow for performance based incentives or repeated task execution. Hence, the results we reveal in the previous section are subject to these objections. Therefore, in the present section we validate our survey results by a lab experiment.

4.1. Method and sample

We apply a computer-based lab experiment that is programmed and conducted with the software z-Tree (Fischbacher, 2007). While the basic setting matches the conjoint analysis, there are two main differences that address the limitations of conjoint analysis. First, we introduce performance based compensation and second, we establish a learning environment.

The experiment took place at the European University Viadrina and the Freie Universität Berlin, Germany. Subjects are 56 non-academic university employees. In contrast to a convenient student sample this subject pool has the advantage that employees come with actual tax experience. Most of the non-academic employees at the European University Viadrina and the Freie Universität Berlin, Germany are women. This is also reflected in our sample as 50 out of 56 are females. Subjects are on average 44.1 years of age. This corresponds to the average age of the subjects that take part in conjoint analysis. The average education level is slightly higher in the lab experiment compared to the subjects that participated in the conjoint analysis.

[16] Stimulus D was eliminated from the analysis because it was given as the objectively best alternative. Thus, stimulus D could not be selected by the subjects.

Table 7
Aggregated ranking of tax options.

Stimulus	Tax rate (in %)	Tax deductible expenses (in %)	Tax Burden (in €)	Rational ranking	Sample ranking
1	25	33.3	2250	2	3
2	25	0	2500	4	2
3	30	100	2100	1	1
4	30	33.3	2700	5	5
5	30	0	3000	6	6.5
6	35	100	2450	3	4
7	35	33.3	3150	7	6.5
8	35	0	3500	8	8

Stimuli 5 and 7 receive the same number of choices. Therefore, they are allotted the same rank which is the average of rank 6 and rank 7.

4.2. Research design

According to conjoint analysis subjects have to choose between tax systems that differ only in terms of the nominal tax rate and the allowed deduction of income-related expenses for a given income.

While income is hypothetical in conjoint analysis, subjects that take part in the experiment have to perform a working task to receive income. This eliminates a potential house money effect.[17]

We design a simple decoding task where subjects have to decode letters into numbers according to a given table. Subjects receive 400 points for every correct entered number and are allowed to accomplish up to 25 successful decodings to receive at most 10000 points, which equals the earnings of € 10000 used in the conjoint analysis.

We limit the time subjects have to complete the task to a maximum of 10 min and charge subjects corresponding to the elapsed time. Thus, we are able to generate income-related expenses. There was a fix charge of 3000 points if subjects successfully earn 10000 points in the first five minutes. For every 30 s surpassing the first five minutes subjects are charged with additional 100 points.

To avoid any income effects on the results, it is required that all subjects have the same income available. Pretests showed that almost everybody should be able to complete the decoding task within five minutes. With only one exception all subjects earn 10000 points and bore income-related expenses of 3000 points.[18]

After completing the task subjects are made aware of their working task income. Afterwards, they are presented eight different tax systems which differ only in tax rates (25%, 30%, and 35%) and deductions allowed (100%, 33%, and 0% of income related expenses).[19] Subjects are told to choose the tax option on the basis of which they want their own income to be taxed and that their decision does not affect the taxation of any other participant in the experiment.[20]

While subjects have to rank all tax options in conjoint analysis, they only have to choose one in the experiment. This is due to the fact that a performance based reward is required. To ensure that these rewards control for individuals' preferences saliency is needed. This means subjects must be aware of the relationship between decisions and incentives. Moreover, the relationship between decisions and incentives has to be understandable rather than too complex. Then and only then, subjects can recognize how their payoff is affected by their actions. At first glance, it seems to be pretty easy to find a performance rewarding payoff structure that leads to higher payoff for the rational ranking than for other rankings. However, one has to consider the 120 different ways to rank the stimuli from our conjoint analysis. There are too many issues to adapt a payoff structure that fulfills the demand for saliency. Therefore, subjects only choose their personal best tax option.

As in conjoint analysis the objectively best tax option is excluded, i.e., subjects choose only among the remaining eight options. Subsequent to their choice, subjects receive an extensive feedback. They are not only informed on their pre-tax, taxable, and net income, but even calculations are displayed. So the participants receive all the information needed to find the tax option with the lowest tax burden. The working task as well as the choice of the tax option is repeated three times in exactly the same manner. Screenshots of the experiment are displayed in Figs. 2 to 4 in Appendix B. The instructions are presented in Appendix C.

Subjects are compensated according to their net income. Net income is calculated as earned points less income-related expenses less taxes. Since the first two components are the same for all subjects, the choice of the tax option determines the payoff. To enforce the incentive, there are large differences in compensation between the respective tax options. Choosing the tax option with the lowest tax burden yields a return of € 8.00, while choosing the second best option is awarded with a return of € 5.00. For choosing any other than these two tax options subjects are paid € 2.00. Hence, choosing the tax option

[17] As it is shown in the literature subjects are more involved if their income is not given but earned. This is called the house money effect (Clark, 2002).

[18] Since results were not affected by this subject, we did not exclude her from our study.

[19] In conjoint analysis we used three levels for the attribute "tax rate" and two levels for the attribute "income related expenses". Some studies report that attributes with more levels achieve a higher relative importance (e.g. Wittink, Lakshman, & Reibstein, 1990). To ensure that our results are unaffected by such a "level effect" we decided to validate our conjoint results with a design that contains the same number of levels for both attributes.

[20] In a post experimental manipulation check we asked whether the subjects understand that their choice for a certain tax option does only affect their own experimental income and is independent from the other participants. On a scale from 1 (does not apply) to 7 (does apply) 40 out 56 participants name a 6 or higher. The average (median) answer is 5.63 (6) with a standard deviation of 1.75.

Table 8
Aggregated relative importance (sample and rational benchmark).

	Rational	Sample
Tax rate	0.4615	0.5600
Deduction of income-related expenses	0.5385	0.4400

Table 9
Impact of repeated execution.

	Decision in favor of tax system with lowest tax burden (in%)	
	Yes	No
Before first feedback	32.1	67.9
After first feedback	44.6	55.4

Above the percentages of tax burden minimizing decisions are displayed for the first choice and the following choices in the experiment. The complete overview of the experimental results is given in Table 12 (Appendix B).

with the lowest tax burden thrice leads to the maximum total payment of € 24.00. Since subjects are paid at least € 2.00 for every decision, the minimum total payment amounts up to € 6.00. On average subjects earned a total of € 15.67 (Std. Dev. € 5.86) for an experiment lasting on average 50 min.

The experiment is designed as a validation whether the conjoint analysis results regarding hypothesis H2 are robust with respect to performance based incentives and learning effects. To test whether performance based incentives affect the results from the conjoint analysis, we again use a conjoint analysis to look at the impact of changes in nominal tax rate and tax base.

To test whether the implementation of a learning environment changes the results we use a within-subject-design and apply generalized estimation equation to compare subjects' first choice with their subsequent choices.

4.3. Results

4.3.1. Overestimation of the importance of changes in tax rate (Hypotheses H1 and H2)

In the analysis of subjects' ranking behavior in Section 3, we point out that only 9.4% of our subjects rank all of the stimuli rationally. In the experiment subjects do not have to rank all of them but only have to choose one. Due to this change in our design an individual ranking cannot be obtained. Thus, to conduct a conjoint analysis we aggregate subjects' choices. We obtain the ranking by assigning the stimulus with the highest total percentage of choices the first rank, the stimulus with the second highest total percentage the second rank, and so on. To avoid any impact of learning effects we only consider choices from the first of the three periods. The number of choices for each stimuli is given in Table 12 in Appendix B.

As it can be seen from Table 7, the stimulus with the lowest tax burden is chosen most often. However, the aggregated sample ranking deviates from the rational ranking. In fact, only roughly a third of the subjects (18 out of 56) choose the tax option with the lowest tax burden. The result regarding hypothesis H1 is unaffected by the implementation of performance-based incentives. Even after the implementation of performance-based incentives the majority of the subjects does not choose their tax option based on the actual tax burden.

To test whether performance-based incentives change our results regarding hypothesis H2, we estimate part-worth utilities by Eq. (2) using OLS and compare the results of our aggregated ranking with the rational benchmark. The results are given in Table 8.

As it can be seen from Table 8, the relative importance of the nominal tax rate for the rational benchmark amounts to 46%, while the corresponding relative importance of the sample is 56%. The relative importance of the rate is 10 percentage points higher in the sample than for the homo economicus. As it is for hypothesis H1, the results regarding hypothesis H2 remain unaffected by the implementation of performance-based incentives.

4.3.2. Learning effects

Subjects have to choose their individual tax option thrice. As pointed out before, subjects receive an extensive feedback after each decision. The feedback does not only contain information on subjects' income, but even on calculations. For our analysis, we divide subjects' decisions into two groups. The first group contains the subjects' first decisions, since these choices are carried out before the first feedback. The second group consists of the second and third decisions, respectively. These choices are made after receiving at least one feedback.[21]

In the first period, 18 out of 56 subjects choose the tax option with the lowest tax burden (see Table 9 below). This equals 32%. After receiving feedback on their decisions 50 out of 112 (45%) decisions in the last two runs are in favor of the tax option with the lowest tax burden. The difference in favor of the performance in the last two periods is weakly significant (chi squared test, $p = 0.082$) (see Table 9).

[21] We compared only the first with the last decision, too, but the results remain the same.

Table 10
GEE results: determinants of learning effects.

	Coefficient	Std. error
Period	0.533*	0.309
Education	−0.103	0.476
Tax return self-prepared	0.116	0.461
Age	−0.760	0.481
Constant	0.197	0.420

Dependent variable: rational ranking.
* indicates significance on the 10% level, standard errors are robust.

To test for determinants of learning effects, we use a generalized estimation equation. The parameter estimations and robust standard errors are shown in Table 10. We choose this technique because we are interested in estimation of population average impact. A logistic link function is formulated for the dependent variable "lowest tax burden (yes/no)". As explanatory variables period, age, education and tax knowledge (self-assessment) are used.[22] While we find the demographic variables not to have any impact on the ranking, period has a weak significant positive influence ($p = 0.084$). The likelihood to choose the tax option with the lowest tax burden is 70% higher in the third period compared to the first one.

Hence, the implementation of a learning environment leads to a slight enhancement of the conjoint analysis results. However, even after learning the majority of subjects do not make a rational choice. Most subjects still overrate the nominal tax rate.

5. Discussion

If individuals behave rationally according to traditional economic theory, it makes no difference whether their tax rate or (by equal measure) their tax deductions change. By contrast, if one takes into account the idea that individuals tend to avoid cognitive strain and instead use simplified decision heuristics changes in the tax rate may influence the perceived tax burden more strongly than do changes in the tax base.

To empirically test the hypothesis regarding the use of heuristics, the resulting overestimation of the importance of changes in tax rates, and the framing effect, we conduct both a conjoint analysis and a lab experiment. In line with behavioral studies incorporating the burden of cognitive strain, our results reveal that the majority of the subjects use decision heuristics. The use of heuristics leads to a noticeable overestimation of the relative importance of changes in tax rates and to an underestimation of the importance of changes in the tax base. We show that the overestimation increases considerably when information about the tax rate is mentioned first (framing effect).

Even after the introduction of performance based incentives we observe that subjects still overestimate nominal tax rate changes. However, we find slight evidence for learning effects. After the implementation of a learning environment, the use of heuristics decreases weak significantly, but even after learning the majority of subjects do not make a rational choice. Most subjects still overrate the importance of the nominal tax rate.

Our results indicate that politicians could simultaneously increase tax revenues and reduce the perceived tax burden for taxpayers. Taxpayers favor tax options that apply a lower tax rate on their gross income over a higher tax rate applied on their net income despite the lower actual tax burden of the latter option. This result remains stable in a lab experiment where performance-based incentives and a learning environment are established. Consistent with the identified framing effect, emphasizing the nominal tax rate when presenting tax options can reduce the perceived tax burden solely through the display format.

Our study offers starting points for future research. First, the present study only considers the effect of non-specified income-related expenses. It would be interesting to analyze whether subjects also underestimate the importance of changes in the deductibility of emotionally loaded costs like commuting expenses. Second, we study the behavior of the German working population not the behavior of tax experts. Thus, it would be worth to explore if our results also hold for specialized decision makers in companies. Third, we do not vary complexity. According to current research (e.g., Blaufus & Ortlieb, 2009), subject's performance decreases with increasing complexity. Fourth, we abstract from distribution and equity effects that are usually affected by introducing new tax options. Finally, even though considering changes in tax rate and tax base simultaneously is specific to taxes, it would be interesting to see, whether our results regarding overestimation are of relevance for non-tax frames, too. For example, our results could be relevant for consumer price research. According to our results the value of a relative price discount should be perceived as higher when it is applied at a higher discount rate on a narrow base (e.g. product price excluding extras) compared to a lower discount rate on a broad base (e.g. product price including extras).

[22] One might ask why sex is not used as explanatory variable. Out of the 56 subjects that take part in the experiment only 6 are male. Because of this enormous mismatch sex was excluded of the analysis.

Appendix A

Table 11
Distribution of quota attributes in sample.

Attribute	Value	Frequency	Percent	Percent (Population)
Gender	Male	209	44.8	45.2
	Female	258	55.2	54.8
Age	Under 20	15	3.2	3.5
	20–29	96	20.6	17.5
	30–39	112	24.0	24.0
	40–49	125	26.8	29.6
	50–59	94	20.1	20.5
	60 and older	25	5.4	5.1
Education	University degree	81	17.3	16.0
	University-entrance qualification	80	17.1	14.5
	Secondary school leaving certificate	126	27.0	26.8
	Lower secondary school leaving certificate	136	29.1	31.4
	No school leaving certificate	8	1.7	2.2
	Other	36	7.7	9.1
Monthly net Income	Under € 1000	161	34.5	32.7
	€ 1000–2000	206	44.1	44.8
	€ 2000–3000	65	13.9	14.6
	Above € 3000	30	6.4	8.0
	Not stated	5	1.1	–

Table 12
Distribution of choices in the experiment (in%).

	Tax rate		
	25%	30%	35%
Deductible expenses			
100%	–	32.1 (44.6,44.6)	7.1 (3.6,1.8)
33%	25.0 (25.0,26.8)	5.4 (1.8,7.1)	1.8 (3.6,3.6)
0%	26.8 (19.6,8.9)	1.8 (1.8,5.4)	0.0 (0.0,1.8)

Percentage of choices that the respective tax system received in the first (second, third) period. The tax system with the lowest tax burden (tax rate 25%, deductible expenses 100%, top left) was given. From top to bottom and from left to right tax burdens decrease. The worst tax system is at bottom right (tax rate 35%, deductible expenses 0%).

Appendix B

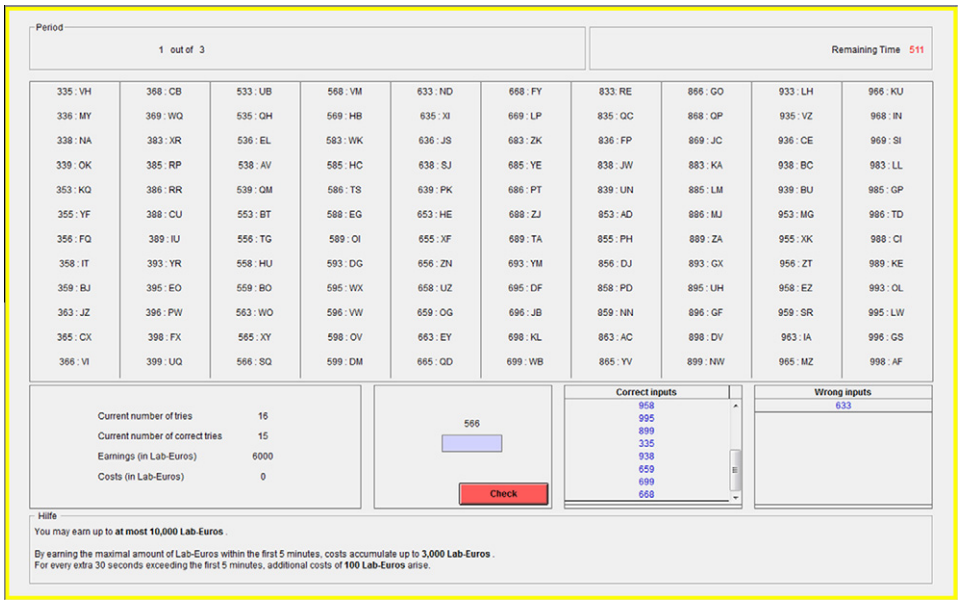

Fig. 2. Screenshot from the experiment – working task.

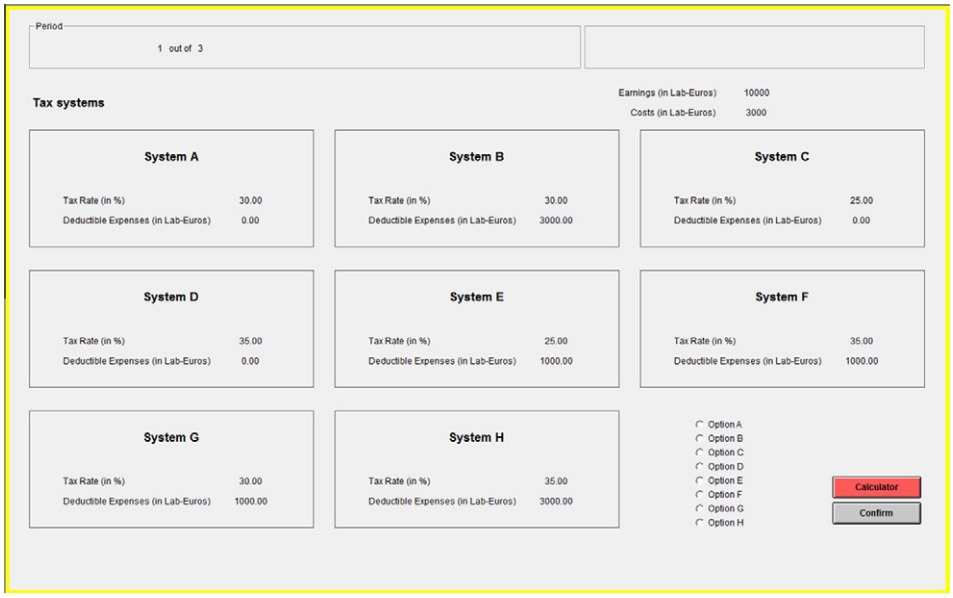

Fig. 3. Screenshot from the experiment – choosing individual tax system.

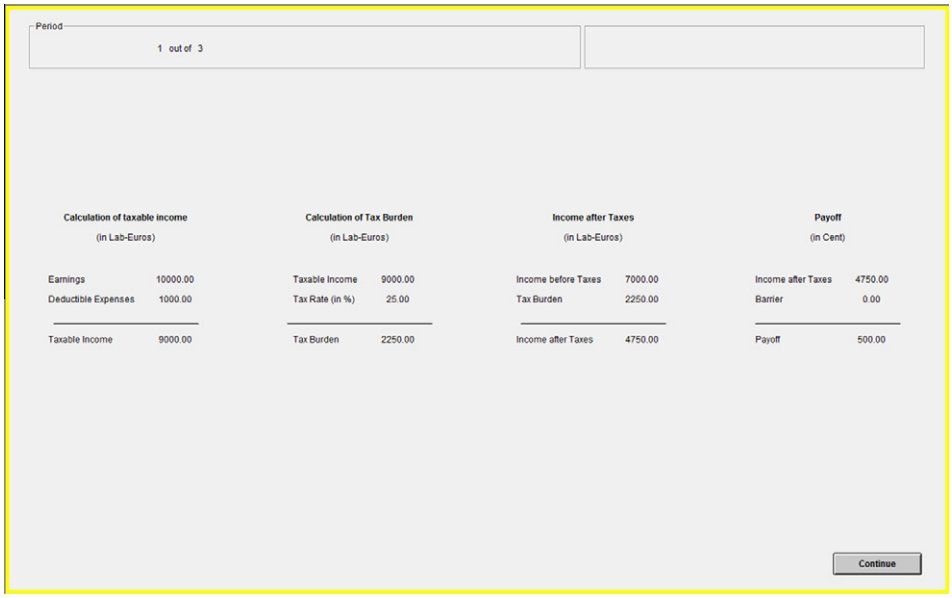

Fig. 4. Screenshot from the experiment – feedback.

Appendix C

C.1. Experimental instructions, part 1 (presented before the working task)

Thank you very much for your willingness to participate in this experiment (see Figs. 3 and 4).

By these instructions the procedure of the experiment will be explained. Please read the instructions carefully. Take as much time as you need to read and understand them.

The experiment consists of three periods. To make you familiar with the software and your task, there will be a trial period that has no impact on the outcome of the experiment. Please take sufficient time to grasp the information presented on the screen. The trial period lasts 5 min.

Every new period won't be started until all participants are ready. This can lead to short waiting times. In the meantime you are presented with current headlines.

During the entire experiment laboratory-Euros are used as currency unit. The exchange rate of laboratory-Euros to actual Euros is explained in detail later.

Your task is to convert three-digit numbers into two letters. Every number-letter combination is given in a list. For example "356: CF".

After entering two letters and a subsequent click on the "check" button, the program compares your input with the data given in the list. Correct inputs are then shown in the table "correct inputs", while incorrectly entered letters are shown in the table "incorrect inputs". Please use only upper case letters.

For every correct input, you are credited with 400 laboratory-Euros. In every period you can reach up to 25 correct inputs, i.e. you can earn up to 10000 laboratory-Euros. To reach the maximum of 25 correct entries, you get at most 10 min of time. If you reach the maximum number of correct entries before time elapses, the task ends automatically. Otherwise it ends after 10 min.

Costs arise while you use the software. The amount of these costs depends on how long you need to complete the working task. Within the first five minutes, you are charged with a fixed amount of 3000 laboratory-Euros. If you need longer than five minutes, you are charged with an additional 100 laboratory-Euros for every additionally 30 s.

The remaining time (in seconds) is shown at the top right corner on the screen.

Your income is subject to an income tax (earnings–costs). You will have to choose one out of eight tax systems that differ, exclusively, in the nominal tax rate (25%, 30%, 35%) and the amount of tax deductible expenses (full deduction allowed, partial deduction allowed, no deduction allowed). You choose your tax system by your own. Your decision does not affect the taxation of the other participants.

Your payment in the experiment depends on your net income:

Earnings–Costs–Taxes (in laboratory-Euros)

At the end of each period, your actual payment (in Cent) is displayed. For every 100 laboratory-Euros that exceed the threshold of 4500 laboratory-Euros, you receive a payoff of 200 Cents. If your net income does not exceed the threshold by at least 100 laboratory-Euros, you will receive the minimum payment of 200 Cents.

If you have any questions regarding the course of the experiment, please do not hesitate to contact the experimental supervisor.

By clicking the "Continue" button, you proceed to the first stage of the experiment. It will start as soon as all participants are ready.

C.2. Experimental instructions, part 2 (presented after the working tax)

Your income is subject to an income tax. Therefore, you have to choose one out of eight taxation options. You choose your tax system by your own. Your decision does not affect the taxation of the other participants. Your decision is in effect only for the current period. You may choose other tax systems at the end of the following periods. The eight taxation options differ in two ways.

First, the various options differ in the nominal tax rate at which your income is taxed.

Second, a distinction is made on the amount of deduction allowance. If the deduction allowed equals zero, you have to pay taxes on your earnings without deducting any expenses, i.e. on 10,000 laboratory-Euros. If deduction allowance is not limited, you just have to pay taxes on your earnings less your actual costs, i.e. on 10000 laboratory-Euros − 3000 laboratory-Euros = 7000 laboratory-Euros. If deduction allowance is limited to one third of your actual costs, you have to tax your income less the one third of your actual costs, i.e. on 10000 laboratory-Euros − 1000 laboratory-Euros = 9000 laboratory-Euros.

A calculator is available if needed.

An example of a possible taxation option is given below. The tax rate equals 25%, deductions are allowed up to 3000 laboratory-Euros. Hence, you have to pay taxes in the amount of 1750 laboratory-Euros.

References

Allison, P. D., & Christakis, N. A. (1994). Logit models for sets of ranked items. *Sociological Methodology, 24*, 199–228.

Blaufus, K., & Ortlieb, R. (2009). Is simple better? A conjoint analysis of the effects of tax complexity on employee preferences concerning company pension plans. *Schmalenbach Business Review, 61*, 60–83.

Boylan, S. J., & Frischmann, P. J. (2006). Experimental evidence on the role of tax complexity in investment decisions. *Journal of the American Taxation Association, 28*, 69–88.

Brandstaetter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113*, 409–432.

Carmone, F. J., & Green, P. E. (1981). Model misspecification in multiattribute parameter estimation. *Journal of Marketing Research, 18*, 87–93.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *The American Economic Review, 99*, 1145–1177.

Chrzan, K. (1994). Three kinds of order effects in choice-based conjoint analysis. *Marketing Letters, 5*, 165–172.

Clark, J. (2002). House money effects in public good games. *Experimental Economics, 5*, 223–231.

Elffers, H., & Hessing, D. J. (1997). Influencing the prospects of tax evasion. *Journal of Economic Psychology, 18*, 289–304.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*, 311–318.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*, 171–178.

Genser, B., & Reutter, A. (2007). Fiscal policy in action: Moving towards dual income taxation in Europe. *FinanzArchiv/Public Finance Analysis, 63*, 436–456.

Green, P. E., & Devita, M. T. (1974). A complementarity model of consumer utility for item collections. *Journal of Consumer Research, 1*, 56–67.
Green, P. E., & Krieger, A. M. (1985). Models and heuristics for product line selection. *Marketing Science, 4*, 1–19.
Green, P. E., & Krieger, A. M. (1993). Conjoint analysis with product positioning applications. In J. Eliashberg & G. L. Lilien (Eds.), *Handbook of Operations Research Series. Marketing* (Vol. 5, pp. 467–515). New York, NY: North-Holland.
Green, P. E., & Srinivasan, V. S. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5*, 103–122.
Green, P. E., Tull, D. S., & Albaum, G. (1988). *Research for marketing decisions* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2008). *Multivariate data analysis* (7th ed.). New York, NY: Pearson Education.
Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief adjustment model. *Cognitive Psychology, 24*, 1–55.
Hundsdoerfer, J., Sielaff, C., Blaufus, K., Kiesewetter, D., & Weimann, J. (in press). The influence of tax labeling and tax earmarking on the willingness to contribute – A conjoint analysis. Schmalenbach Business Review.
Hundsdoerfer, J., & Sichtmann, C. (2009). The importance of taxes in entrepreneurial decisions: An analysis of practicing physicians' behavior. *Review of Managerial Science, 3*, 19–40.
Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge: Cambridge University Press.
Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics, 89*, 2093–2119.
Krishna, A., & Slemrod, J. (2003). Behavioral public finance: Tax design as price presentation. *International Tax and Public Finance, 10*, 189–203.
Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1*, 1–27.
McCaffery, E. J., & Baron, J. (2003). The Humpty Dumpty blues: Disaggregation bias in the evaluation of tax systems. *Organizational Behavior and Human Decision Processes, 91*, 230–242.
McCaffery, E. J., & Baron, J. (2004). Framing and taxation: Evaluation of tax policies involving household composition. *Journal of Economic Psychology, 25*, 679–705.
Milliron, V. C., & Toy, D. R. (1988). Tax compliance: An investigation of key features. *Journal of the American Taxation Association, 9*, 84–104.
Moran, S., & Meyer, J. (2006). Using context effects to increase a leader's advantage: What set of alternatives should be included in the comparison set? *International Journal of Research in Marketing, 23*, 141–154.
Morwitz, V. G., Greenleaf, E. A., & Johnson, E. J. (1998). Divide and prosper: Consumers' reactions to partitioned prices. *Journal of Marketing Research, 35*, 453–463.
OECD (2010). Choosing a broad base – low rate approach to taxation. *OECD Tax Policy Studies*, No. 19, OECD Publishing.
O'Neil, C. J. (1982). The targeted jobs credit: An evaluation of its impact on the employment decision process. *Journal of the American Taxation Association, 3*, 15–22.
Pitt, M. M., & Slemrod, J. (1989). The compliance cost of itemizing deductions: Evidence from individual tax returns. *The American Economic Review, 79*, 1224–1232.
Rupert, T. J., Single, L. E., & Wright, A. M. (2003). The impact of floors and phaseouts on taxpayers' decisions and understanding of marginal tax rates. *Journal of the American Taxation Association, 25*, 72–86.
Rupert, T. J., & Wright, A. M. (1998). The use of marginal tax rates in decision making: The impact of tax rate visibility. *Journal of the American Taxation Association, 20*, 83–99.
Schmölders, G. (1960). *Das Irrationale in der öffentlichen Finanzwirtschaft*. Hamburg: Rowohlt.
Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–20.
Slemrod, J. (1989). The return to tax simplification: An econometric analysis. *Public Finance Quarterly, 17*, 3–28.
Tax Foundation (2005). Tax Foundation Annual Survey of US Attitudes on Tax and Wealth.
Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*, 299–314.
Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458.
Willner, J., & Granqvist, L. (2002). The impact on efficiency and distribution of a base-broadening and rate-reducing tax reform. *International Tax and Public Finance, 9*, 273–294.
Wittink, D. R., Lakshman, K., & Reibstein, D. J. (1990). The effect of differences in the number of attribute levels on conjoint results. *Marketing Letters, 1*, 113–123.
Yadav, M. S. (1994). How buyers evaluate product bundles: A model of anchoring and adjustment. *Journal of Consumer Research, 21*, 342–353.

Blaufus, Kay, Bob, Jonathan, Hundsdoerfer, Jochen, Sielaff, Christian, Kiesewetter, Dirk, Weimann, Joachim, 2013. Perception of income tax rates: evidence from Germany. European Journal of Law and Economics. http://dx.doi.org/10.1007/s10657-013-9389-9.

The effect of tax privacy on tax compliance – an experimental investigation

Kay Blaufus[a], Jonathan Bob[a] and Philipp E. Otto[b]

[a]*Leibniz Universität Hannover*
[b]*European University Viadrina Frankfurt (Oder)*

Abstract In this paper, a tax game with audit costs as a public bad is designed to investigate the impact of public disclosure on tax evasion experimentally. Three different types of tax privacy are tested, ranging from complete privacy to full disclosure. We expect to observe two different effects: first, a contagion effect, arising when an individual observes non-compliance of other individuals and therefore reduces his own tax compliance; second, a shame effect of increased tax compliance due to the anticipated shame of being declared a tax evader. We find evidence of increasing tax evasion with reduced tax privacy if information is disclosed anonymously. Our results also indicate that the shame effect is not strong enough to override the contagion effect when both effects are present. Our results are of particular importance for fiscal policy because public disclosure may lead to more evasion instead of less, due to motivational crowding-out of tax morale.

Keywords Tax privacy · Tax evasion · Public bad · Social Norm · Conditional cooperation · Economic experiment

JEL classification H23 · H24 · H26 · H30

1 Introduction

Death is certain, but paying taxes is definitely not—at least, not for everyone. According to the Internal Revenue Service, the tax gap in the United States amounts to $385 billion for tax year 2006 alone, mainly due to underreported income. Non-compliance reduces both public revenue and the availability of public services and also discriminates against honest taxpayers (Alm, 2012). Therefore, fighting tax evasion is an important issue for policy agendas. A number of countries (e.g., Greece and New Zealand) publicly list tax evaders to combat tax evasion. Others (e.g., Finland, Iceland, Norway, and Sweden) disclose all tax return information. However, the majority of countries treat tax information confidentially.

The main reason to disclose tax compliance information publicly is to deter people from evading taxes by threatening them with the shame of being announced as tax evaders. In addition to the imposition of monetary penalties, shame should be effective as a non-monetary sanction. However, it is far from obvious that a strategy of tax publicity is a successful instrument for fighting tax evasion. Previous research has demonstrated that

social norms have a considerable impact on tax evasion (Cullis et al., 2012). Individuals comply as long as they believe that compliance is the social norm (Alm, 2012). Gino et al. (2009) show that the observation of unethical behavior of another person is potentially contagious because it may change the social norms regarding dishonesty. Therefore, at the same time that publishing information could be a deterrent to tax evasion (via the shame effect), it could also destroy the social norm of compliance. This effect would be in keeping with a strand of literature that shows that taxpayers are conditionally cooperative, i.e., people are willing to comply as long as others do (e.g., Frey and Torgler, 2007; Traxler, 2010). Due to these potentially opposing effects—increasing shame on the one hand, risk of contagion on the other—the overall effect of public disclosure on tax compliance remains unclear.

Our study investigates different levels of publicized tax evasion to determine whether the shame effect or the contagion effect dominates as an overall behavioral response. Using a lab experiment enables a controlled variation of tax privacy as a treatment variable. We design a specific tax game with a baseline treatment of no public disclosure. In one treatment, individual tax information is disclosed publicly in an anonymous manner; therefore, only the contagion effect may arise. In another treatment, public disclosure occurs by displaying the pictures of the players next to their tax information, introducing both shame and contagion effects.

Overall, we find evidence of increasing tax evasion with reduced tax privacy if information is disclosed anonymously. In the case of full public disclosure, i.e., when both the contagion and shame effects are present, shame is not strong enough to override contagion. The policy implications are clear: universal tax return publicity should not be implemented because more evasion could result due to motivational crowding-out of tax morale.

The remainder of this paper is organized as follows. In section 2, we give a brief literature review with its implications for our study. In section 3, the tax game and the experimental design of the study are described. The results are provided and discussed in section 4, and section 5 concludes.

2 Tax evasion theory

The theory of tax evasion is primarily based on the work of Allingham and Sandmo (1972), which assumes that each individual maximizes expected utility after taxes, applying a certain audit probability and penalty level. However, it has been shown that actual tax compliance often differs from the predictions of this model (e.g., Dhami and Al-Nowaihi, 2007), and the model has therefore been modified in several aspects. With respect to tax publicity, the following two modifications seem particularly important. First, Erard and Feinstein (1994) account for moral sentiments (such as shame) and empirically show that sentiments can be important determinants of compliance. In their model, individuals experience utility-reducing shame when they evade taxes and are subsequently audited. Shame reduces the benefits of evasion and decreases its occurrence. If tax publicity increases shame, disclosure could help to reduce evasion. Second, Traxler (2010) incorporates tax morale, as internalized social norm of tax compliance, into the Allingham and Sandmo (1972) standard model. Taxpayers are assumed to conditionally cooperate because their level of evasion depends on others' compliance. The results imply that strategies that increase belief in high compliance levels reduce tax evasion. Consequently, publishing information about actual tax evasion could alter this belief in a high compliance level and thus destroy the corresponding social

norm. This effect would conform to the observation of the contagion effect of unethical behavior in Gino et al. (2009) or the widely discussed broken window hypothesis (Wilson and Kelling, 1982).

Hence, tax publicity might trigger two opposing effects simultaneously—the shame effect and the contagion effect. The sum of the overall effect of tax publicity is theoretically unclear; therefore, its positive or negative impact is a matter for empirical investigation. Despite its great policy importance, direct evidence on the effects of tax privacy is scarce. To our knowledge, there exists neither a theoretical study that incorporates both these effects nor an empirical study of public tax disclosure that simultaneously evaluates the contagion and shame effects on tax compliance. However, there are two experimental and two empirical studies on the effects of tax publicity.

Laury and Wallace (2005) investigate the impact of tax confidentiality experimentally in a between-subjects design. Individuals decide how much of a given income to report to the tax authority under two different treatments. Subjects are informed about the tax rate, the (exogenous) audit probability, and the relevant fine. In the first treatment, full confidentiality is employed. In the second treatment, only partial confidentiality is used, and 25% of the subjects' decisions are disclosed to all other participants. However, the decisions are anonymous and cannot be tracked to the actual person making the decision. The results show that reported income is typically higher under the partial confidentiality treatment; however, this difference is significant only in five out of twenty periods. Moreover, when controlling for demographic variables (particularly gender, marital status, student of economics, raised in North America) the treatment effect becomes insignificant. By and large, the results of this study are ambiguous.

Coricelli et al. (2010) use a within-subjects design to study the impact of tax publicity on compliance. Subjects, in groups of eight players, decide individually how much income to declare. The declared income is subject to a uniform tax rate. Again, taxpayers are informed about audit probability and fines. The treatment variable is the publication of individual pictures. In half of the trials, if an audit reveals that a player underreported his income, a picture of the detected evader is displayed on the group members' screens. The results show that tax publicity reduces both the number of evaders and the amount of tax evaded. The risk of being "named and shamed" as an evader diminishes the probability that an individual will evade taxes by 8.2%.

In addition to these experimental studies, two archival studies were recently published. Hasegawa et al. (2012) use Japanese data where disclosure of both individual and corporate income tax information was mandatory from 1950 until 2004. These data show no evidence that companies reduced declared taxable income after the disclosure requirement was abolished in 2004. However, companies prefer to avoid public disclosure and therefore decreased reported taxable income to be below the threshold beyond which disclosure is required. In contrast to Hasegawa et al. (2012), Slemrod et al. (2013) report evidence that public disclosure of tax returns on the Internet increases reported income. These authors' estimate of the effect of public disclosure on tax compliance based on a quasi-experimental study using data from Norwegian income tax statistics is an increase in reported income of 3%.

Overall, evidence on the effect of tax publicity is rare and the data available are ambiguous. While Laury and Wallace (2005) find only a weak effect from tax publicity and Hasegawa et al. (2012) find no effect from tax publicity, the results of Slemrod et al. (2013) and Coricelli et al. (2010) indicate that abolishing tax privacy laws could increase tax compliance. Moreover, the consequences of different sorts of tax publicity are unclear. For example, the weak results of Laury and Wallace (2005) could be due to the anonymous form in which tax

return information is announced in these studies. There, it can be expected that anticipated shame is no deterrent because participants remain anonymous. If, however, a person can be identified by the other participants—e.g., by displaying photos of the subject as in the experiment by Coricelli et al. (2010)—shame should be exacerbated. This might explain why, in contrast to Laury and Wallace (2005), Coricelli et al. (2010) find a strong positive effect of disclosure on compliance.

Interestingly, neither of the experimental studies on tax compliance behavior implements a public good or any refund of taxes. This neglects the fact that public good games provide a standardized opportunity for studying social interactions within groups (Frey and Torgler, 2007). Outside of such a public good context, externalities may not arise; hence, there is no need for social norms (Huck et al., 2012). This possibility raises the question of whether these experiments underestimate the contagion effect. Conditional cooperators do contribute as long as others contribute, but without a public good context, there is neither necessity nor an opportunity to cooperate. Without a direct public good context, it is unclear whether participants perceive their payment of taxes as contributions to a public good or simply as a cost. If one's own income is not affected by the decisions of other group members, rules of reciprocity or conformity may seem less important. Therefore, the contagion effect loses its bite. To answer this question, we conduct a tax experiment that is designed to investigate the responses to different levels of tax publicity in a public good context, which also allow us to draw conclusions about whether a shame or contagion effect dominates overall contributions.

3 Experimental design

We adapt a standard public good game with repetition, inverting the public good by introducing audit costs as a public bad. The rationale behind this is that evasion requires the state to implement a costly tax audit administration that must be equally shared by all subjects. The costs of this administration reduce the endowment of each individual upfront, where each of the N subjects bears $1/N$ of the costs. In a world without evasion, there would be no need for tax audits. Thus, by complying with the tax law, subjects reduce the costs of tax audits, i.e., they reduce the public bad. However, if some subjects do not contribute their share of taxes, the state will need to increase the audit probability. That, in turn, increases the public bad. This idea is implemented in the following way:

A group of six players receives an aggregate endowment of 450 cents income in each round (each individual receives 75 cents).[1] Given a nominal tax rate of 30%, the group in total must pay income taxes of 135 cents (450×0.3) in each round. All taxes paid are removed from the game (there is no refund or compensation) and only tax evasion determines the public bad in the form of changing audit costs. In the case of equal contribution, every member of the group would have to pay 22.5 cents, but subjects can decide individually how much tax they wish to pay. Hence, subjects are given the chance to evade taxes, where the individual evasion is given by

$$E_{i,t} = \max\left[22.5 - \tau_{i,t}; 0\right] \tag{1}$$

with $\tau_{i,t}$ being the tax paid by subject i in round t, and $E_{i,t}$ being the resulting evasion.

[1] Cent is the experimental currency and for every cent subjects earn they receive a euro-cent as payoff.

Total evasion determines the cost incurred in the next round. Hence, the cost of audit administration is equal to the sum of taxes evaded by all subjects of one group in the previous round given by

$$C_t = \sum_{i=1}^{6} E_{i,t-1} \tag{2}$$

with C_t being the cost of round t, and $E_{i,t-1}$ being the evasion of individual i in round $t-1$. This cost may be perceived as damage done that must be borne equally by all subjects. As long as there is no individual tax evasion, the cost of audit administration remains zero. Therefore, by paying taxes, subjects reduce the implemented public bad. If some subjects choose to ride completely free or not fully contribute their sixth of the demanded tax, the cost of audit administration increases. Because cost must not be negative, i.e., $C_t \geq 0$, the overpayment of taxes is not refunded. Audits to combat tax evasion occur with a given probability that is dependent on the tax evasion present in the previous round. Thus, audit probability is endogenous and defined by

$$p_t = 0.005 \times C_t \tag{3}$$

with p_t being the audit probability of round t, and C_t being the cost from (2). This restricts the audit probability to the interval 0 to 67.5%. An overview of these relations between taxes paid, cost, and audit probability is provided in Table 3 of Appendix A.

Audits occur at the individual level based on the probability given in (3). After every round, it is determined whether declared taxes are sufficient for each subject. If an audit occurs and tax evasion is detected (i.e., if $E_{i,t} > 0$), the subject has to pay immediately the amount $E_{i,t}$ plus a penalty of 50% of $E_{i,t}$. Hence,

$$\theta_{i,t} = \begin{cases} 1.5 \times E_{i,t} & \text{if an audit occurs,} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

with $\theta_{i,t}$ being the sanction for subject i in round t. Note also that we follow the recommendation of Alm (2010) and describe the game in neutral language to avoid subjects using individual scripts when interpreting loaded terms (i.e., instead of the term "tax", the term "fee" is used).

In each period, subjects' experimental income is equal to the original endowment of 75 cents less taxes, costs, and sanctions. The overall income equals the sum of the 20-period earnings and the payoff from the experiment is given by

$$\pi_i = \sum_{t=1}^{20} 75 - 22.5 + E_{i,t} - \frac{C_t}{6} - \theta_{i,t}. \tag{5}$$

In a between-subjects design, we implement three treatments that differ in their degree of tax privacy. In our baseline treatment (no information treatment), subjects are only aware of the group's overall tax gap, but have no information on individual tax evasion. This setting reflects tax privacy as it is in most countries. The degree of tax evasion is (at least roughly) known, but individual misconduct is not. In the second and third treatments, there is no tax privacy: all subjects are directly informed about the individual behavior of their group members after each round—everyone knows who evades taxes and to what extent.

The difference between treatments two and three is the type of publication. In the second treatment (partial information treatment), only subject numbers are stated, but numbers and individuals cannot be matched. This treatment corresponds to the partial confidentiality treatment in Laury and Wallace (2005). Moreover, because desks are individually separated, other subjects and their screens are not visible. Feelings of shame should not arise under the anonymous disclosure used in this treatment. However, because subjects are provided with information regarding the individual behavior of the other subjects, a contagion effect could result. In the third treatment (full information treatment), photos of subjects are paired with each subject's tax behavior.[2] This treatment is equivalent to the picture treatment in Coricelli et al. (2010). Due to individual disclosure, both shame and contagion effects can be expected. By comparing both the partial and the full information treatments, we can identify whether the shame or contagion effect predominates when ceding tax privacy.

The experimental procedure is as follows: After entering the laboratory, subjects are randomly assigned to their group.[3] The subjects remain in the same group and are provided the same treatment throughout the whole experiment. Subjects are given instructions (see Appendix A) after being seated and as much time as they require to fully understand the procedure. Only after all subjects confirm that they fully understand the experimental instructions and do not have any remaining questions does the tax game begin.

Each experimental session consists of 20 rounds. Screenshots of the different stages of the experiment are provided in Appendix C. At the beginning of each round, subjects are informed about the taxes evaded in the previous round and about the resulting consequences, i.e., the group's total cost according to (2) and the current audit probability according to (3).[4]

Then, subjects are informed about their endowment (75 cents), their share of the group's total costs, their endowment after costs are deducted, the required amount of taxes (22.5 cents), and the current audit probability. At this point, subjects must decide (without time restrictions) how much in taxes they wish to contribute.

After deciding their contribution, subjects are informed of the group's total result. Subjects are also informed whether their contributions are subject to an audit[5] and, if applicable, on the amount of penalties they must pay.

At the end of each round, subjects are informed about their results from the current round as well as their total income so far.

When all 20 rounds are completed, subjects must answer a short questionnaire that seeks information regarding demographic variables including age, gender, and area of expertise. Then, individual risk aversion is measured using lottery decisions based on the procedure of Holt and Laury (2002). An extract of the full questionnaire is given in Appendix B.

A total of 198 undergraduate and graduate students participated in the computer-based experiments, programmed and conducted under z-Tree (Fischbacher, 2007). The experiment

[2] These pictures were taken before the experiment inside the laboratory. After the experiment was finished, all photos were deleted in the presence of the participants.

[3] 20 sessions were conducted with 13 sessions made up of 2 groups and another 7 sessions made up of one group. In 10 of the 13 sessions with 2 groups, subjects were assigned to either the no information treatment or the partial information treatment, while in the other 3 sessions all subjects were assigned to the same treatment. In 5 of the 7 sessions with only one group, individuals were assigned to the full information treatment, while in the other sessions they were all assigned to the no information treatment.

[4] Since there was no round 0, costs and audit probability for round 0 were given exogenously with $C_1 = 80$ and $p_1 = 0.40$.

[5] A uniformly distributed random number between 0 and 1 was drawn. If this number did not exceed the current audit probability, an audit occurred.

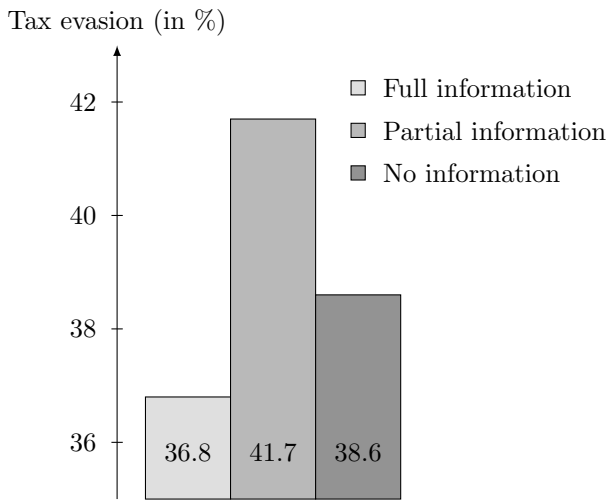

Figure 1: Average tax evasion in the three treatments

took place in the ViaLab of the European University Viadrina Frankfurt (Oder). Subjects earned an average of €9.99 (std. dev. €1.05). They were on average 22 years old and 85% were German. The majority of the subjects (2/3) were economics students.

4 Results

The overall result is an increase of non-compliance in the absence of tax privacy if information is disclosed anonymously. Our results from the partial information treatment indicate that this result is due to the contagion effect. The results from the full information treatment show that shame does not weaken this effect significantly. Tax evasion is not lower in the full information treatment than in the baseline treatment, suggesting that the shame effect is not strong enough to override the contagion effect when both effects are present.

4.1 Descriptive statistics

Tax evasion is higher in the absence of tax privacy, as shown in Figure 1. In the no information treatment, the average tax evasion is 8.28 cents (37% of the required taxes), whereas tax evasion amounts to 9.38 cents (42%) in the partial information treatment. This difference in average tax evasion is highly significant (two-sample t-test, $p < 0.01$).[6] Observing non-compliance causes conditionally cooperative taxpayers to reduce their contributions, and thus the contagion effect is positive. Tax evasion is higher when information on other subjects does not contain personal information. In the full information treatment, tax evasion amounts to 8.69 cents (39%). The difference stemming from the shame effect is not significant in comparison to the evasion observed in the partial information treatment (two-sample t-test, $p = 0.15$), and the shame effect compensates for the contagion effect as tax evasion

[6] We tested for difference in tax evasion between all three treatments and, hence, performed multiple testing. All p-values are adjusted according to Holm (1979).

	No information		Partial information		Full information	
Tax evasion (in %)	36.78	(41.40)	41.66	(44.15)	38.61	(41.73)
Audit Probability	0.23	(0.20)	0.25	(0.21)	0.25	(0.21)
Audit Frequency	0.25		0.25		0.26	
Penaltiy (in cent)	18.77	(12.38)	19.92	(12.95)	19.18	(11.74)
Final profit (in €)	9.94	(1.06)	10.03	(1.24)	9.99	(0.72)
Females (in %)	62.5		52.8		57.4	
Age (in years)	22.26	(2.59)	22.17	(2.33)	22.07	(2.33)
Economic students (in %)	63.9		70.8		70.4	
Risk aversion	5.26	(2.63)	5.38	(1.83)	5.02	(2.31)
Subjects	72		72		54	

Values shown above are means with standard deviations in parentheses of cents per period. Audit frequency is simply the overall percentage of audits that occurred. Subjects' final profit is their total experimental income accumulated over 20 rounds. Risk aversion was measured based upon Holt and Laury (2002). Subjects were given 10 paired lottery-choice decisions. Risk aversion is the average number of decisions subjects made in favor of the risk free alternative. A more detailed overview on the measure of risk aversion, as well as the other variables, is given in Appendix B. The subjects row simply shows the number of subjects per treatment.

Table 1: Descriptive statistics

in the full information treatment is not significantly higher than in the baseline treatment ($p = 0.29$).[7]

Descriptive statistics are shown in Table 1. Audit probability and penalty are calculated from tax evasion as illustrated in Table 3 of Appendix A.

Audit probability is a linear function of tax evasion (3) and, on average, is higher in both the partial and the full information treatment (25% each) than in the no information treatment (23%). The frequency of audits is very similar in all three treatments (25% in the no and partial information treatments and 26% in the full information treatment).

Even though tax evasion is the highest in the partial information treatment, penalties are the lowest, indicating that subjects in the partial information treatment use low audit probabilities to fully evade taxes. Accordingly, final profit is slightly higher in the partial information treatment than in the other two treatments, but this difference is not significant (analysis of variances, $p = 0.17$). Overall, subjects earned on average €9.99. Subjects could easily raise average incomes by full compliance.[8]

[7] Separating the contagion effect from the shame effect is not possible in this experimental design because the shame component always also leads to contagion.

[8] Note that subjects had to bear total costs in the amount of $C_1 = 80$ in the first round (see section 3) and under full compliance their maximum profit in the first round would be $75-80/6-22.5 = 39.2$. In all other rounds the maximum profit under full compliance is $75-22.5 = 52.5$. Hence, $\pi_i = 39.2+1952.5 = 1036.7$ and their incomes would increases by on average 37.8, which is about 4% of their actual mean income.

	Coefficient	Robust Std. Err.	p-value	
Intercept	9.13	0.77	<0.01	***
Partial information treatment	1.36	0.38	<0.01	***
Full information treatment	0.94	1.50	0.53	
Audit probability	-25.45	1.19	<0.01	***
Female	-1.83	0.49	<0.01	***
Student of economics	1.39	0.54	0.01	**
Risk aversion	-0.45	0.17	<0.01	***
Income	0.01	0.00	<0.01	***

Dependent variable: tax evasion, 3960 observations, adj. $R^2 = 0.29$
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Standard errors are clustered at the group level

Table 2: Random effects model (OLS)

4.2 Panel analysis

What influences tax evasion behavior? As noted in section 3, a total of 198 students participated in experimental sessions comprising 20 rounds each, resulting in $198 \times 20 = 3960$ observations. We apply a one-way random effects panel model on these data. This panel model offers a major advantage: it enables us to include unobserved heterogeneity among individuals in our model, implicitly assuming that there are differences in behavior across subjects. We choose random, instead of fixed, effects because fixed effects only allow consideration of varying regressors. In our design, only audit probability and income vary between subjects. Because we are not only interested in audit probability but also in the impact of tax privacy and other determinants on tax evasion, we must use random effects. Our one-way random effects model is given by

$$E_{i,t} = \mu + \alpha_i + \gamma_j T_j + \sum_i \delta X_{i,t} + \varepsilon_{i,t} \tag{6}$$

with $E_{i,t}$ as the tax evasion of subject i in round t, μ being the average tax evasion and α_i being the individual specific effect. T_j are the dummy variables for both the partial and the full information treatment, and the matrix X_{it} contains the covariates with gender, subject of study, risk aversion, and income. The no information treatment (full anonymity) is the baseline treatment. In the partial information treatment, public disclosure is anonymous; in the full information treatment, real public disclosure occurs. Subjects' total profit accumulated at the beginning of each round is their income. The other model variables are briefly described in section 4.1, and the results are given in Table 2.

The descriptive results are further confirmed by the panel regression: tax evasion increases with decreasing tax privacy. The coefficients of both the partial and the full information treatments are both positive, but only the coefficient of the partial information treatment is significant. Because disclosure is anonymous in the partial information treatment, the strong effect of this treatment can only stem from the simple observation of others' behavior, i.e., a contagion effect. As in our descriptive results, the difference in tax evasion between the

baseline and the full information treatments is not significant, indicating that shame is not strong enough to override the contagion effect. This is further supported by the fact that the difference between the two information treatment coefficients is not significant.[9]

The impact of control variables is as expected from the previous literature. Audit probability has a highly significant negative impact on tax evasion. Male subjects evade more taxes than female subjects, which comports with Kastlunger et al. (2010), and economics students are less compliant than others, as has been shown before by Cullis et al. (2012). Tax evasion increases with decreasing risk aversion and risk aversion decreases with higher income—together, these effects lead to more tax evasion at higher incomes. These patterns prevail even if the public resource is crowded out, and all must share the burden of the public bad.

5 Conclusion

To determine the effect of tax privacy on tax compliance, we designed a tax game with tax privacy as the treatment variable. Tax privacy ranged from full confidentiality to full publicity, combining the different forms of tax publicity from Laury and Wallace (2005) and Coricelli et al. (2010). Theoretically, two opposing effects—a contagion effect and a shame effect—can occur in response to public disclosure. We investigated these effects at three different levels of tax privacy.

In the partial information treatment, tax information is publicly disclosed in an anonymous manner. The only difference from the baseline treatment of full tax privacy is that the behavior of the other subjects in one's group is observable. Because subjects remain anonymous in this treatment, the shame effect is not expected, and the only impact on compliance results from the contagion effect. In contrast, in the full information treatment, public disclosure of all individuals is employed: tax evaders' photos are shown, which presents the potential for the shame effect. The shame effect only arises if subjects' contributions are known and linked to each subject. Overall, we find higher tax evasion with public disclosure if information is published anonymously. This indicates that the pure observation of deviant behavior could destroy the social norm of compliance and lead to one's own non-compliance. Additionally, the shame effect appears to be too small to override the contagion effect when both are present simultaneously. This is of particular importance to fiscal policy: general public disclosure could lead to more, instead of less, evasion, especially if information is published anonymously.

These results should be interpreted with caution. While Alm et al. (1992) do not find any impact from terminology used (loaded vs. neutral) on tax compliance, other studies provide some evidence that subjects are more compliant in a tax, as opposed to a neutral, context (Baldry, 1986; Wartick et al., 1999; Durham et al., 2012). This could potentially influence the relative strength of the different effects. The shame effect could also be more effective in natural settings where long-term economic consequences often result. However, the work of Coricelli et al. (2010) indicates that accounting for emotions in the laboratory is possible and that emotion is the main driver of the shame effect. There is no direct reason why emotion should not be present in the laboratory, but stronger financial consequences might increase the importance of fear in the tax evasion decision. The broader question here concerns the impact an initiative might have on social behavior where the motivational

[9] The difference of the coefficients is $1.36 - 0.94 = 0.42$ and the corresponding standard error amounts to 1.12, which leads to $p = 0.71$.

crowding literature (i.e., Benabou and Tirole, 2006 or Bolle and Otto, 2010) stresses the importance of the current state for predicting potential changes.

In contrast to Laury and Wallace (2005) and Coricelli et al. (2010), tax publicity leads to higher tax evasion in our study (at least) in the case of anonymous disclosure. This strong difference might be due to the missing public good context in the previous studies, which leads to an underestimation of the contagion effect. Outside a public good context, there exists neither a necessity nor an opportunity to cooperate. Within this standard of cooperative behavior, the effects of tax publicity remain ambiguous with no sustained effect of copying positive behavior, as can be observed, for instance, in a fundraising context (Andreoni and Petrie, 2004). In addition to the negative impact of shame on tax evasion, there is an opposing contagion effect that might, overall, lead to less compliance.

A Experimental instructions

Thank you very much for participating in this experiment on the evaluation of group behavior. Overall, the experiment will last about one hour. Your payoff depends both on your own decisions and on the decisions of the other players. In the following, we will give you a short overview on how the experiment is carried out. If you have any questions, please do not hesitate to contact the experimental supervisor.

1. Initially, you are randomly assigned to one out of two groups. Both groups are made up of six players each. Throughout the whole experiment, the groups' composition remains unchanged.

2. There will be a total of 20 rounds. At the beginning of each round, every group member receives an endowment of 75 cents. Thus, the group's total endowment is 450 cents for every round.

3. In each of the 20 rounds, you must pay fees on your endowment. You may decide on the amount of your fees yourself. However, in each round your group's total fees should sum up to 135 cents. In the case of a uniform distribution, every member of your group would have to pay 22.5 cents.

4. If your group's total fees amount to less than 135 cents, the consequences for the following round are twofold. First, there will be a cost for your group in the amount of the underpaid fee. Secondly, the amount of the underpaid fee will determine the probability of your individual fee being audited in the next round. If you are audited, and it is found that you paid less than the required 22.5 cents, your underpaid fee plus a penalty of half of your less underpaid fees will be collected automatically.

5. The consequences resulting from underpaid fees by your group are given in Table 3. The table reads as follows: There will not be any cost in the next round, if you and your group members paid at least the required fee of 135 cents in total. Moreover, your individual fee in the following round will not be subject to an audit. For instance, there will be a cost of 6 cents in the following round if the group's total fees only sum up to 129 cents. These costs will be split equally among every group member. Thus, in the next round, there will be a deduction of 1 cent from the endowment of every member of your group. The corresponding audit probability for the next round amounts to 3%.

Total fees	Total costs	Audit probability (in %)	Total fees	Total costs	Audit probability (in %)	Total fees	Total costs	Audit probability (in %)	Total fees	Total costs	Audit probability (in %)
135	0	0.0	101	34	17.0	67	68	34.0	33	102	51.0
134	1	0.5	100	35	17.5	66	69	34.5	32	103	51.5
133	2	1.0	99	36	18.0	65	70	35.0	31	104	52.0
132	3	1.5	98	37	18.5	64	71	35.5	30	105	52.5
131	4	2.0	97	38	19.0	63	72	36.0	29	106	53.0
130	5	2.5	96	39	19.5	62	73	36.5	28	107	53.5
129	6	3.0	95	40	20.0	61	74	37.0	27	108	54.0
128	7	3.5	94	41	20.5	60	75	37.5	26	109	54.5
127	8	4.0	93	42	21.0	59	76	38.0	25	110	55.0
126	9	4.5	92	43	21.5	58	77	38.5	24	111	55.5
125	10	5.0	91	44	22.0	57	78	39.0	23	112	56.0
124	11	5.5	90	45	22.5	56	79	39.5	22	113	56.5
123	12	6.0	89	46	23.0	55	80	40.0	21	114	57.0
122	13	6.5	88	47	23.5	54	81	40.5	20	115	57.5
121	14	7.0	87	48	24.0	53	82	41.0	19	116	58.0
120	15	7.5	86	49	24.5	52	83	41.5	18	117	58.5
119	16	8.0	85	50	25.0	51	84	42.0	17	118	59.0
118	17	8.5	84	51	25.5	50	85	42.5	16	119	59.5
117	18	9.0	83	52	26.0	49	86	43.0	15	120	60.0
116	19	9.5	82	53	26.5	48	87	43.5	14	121	60.5
115	20	10.0	81	54	27.0	47	88	44.0	13	122	61.0
114	21	10.5	80	55	27.5	46	89	44.5	12	123	61.5
113	22	11.0	79	56	28.0	45	90	45.0	11	124	62.0
112	23	11.5	78	57	28.5	44	91	45.5	10	125	62.5
111	24	12.0	77	58	29.0	43	92	46.0	9	126	63.0
110	25	12.5	76	59	29.5	42	93	46.5	8	127	63.5
109	26	13.0	75	60	30.0	41	94	47.0	7	128	64.0
108	27	13.5	74	31	30.5	40	95	47.5	6	129	64.5
107	28	14.0	73	62	31.0	39	96	48.0	5	130	65.0
106	29	14.5	72	63	31.5	38	97	48.5	4	131	65.5
105	30	15.0	71	64	32.0	37	98	49.0	3	132	66.0
104	31	15.5	70	65	32.5	36	99	49.5	2	133	66.5
103	32	16.0	69	66	33.0	35	100	50.0	1	134	67.0
102	33	16.5	68	67	33.5	34	101	50.5	0	135	67.5

Table 3: Consequences of less paid fees

6. At the beginning of the first round, a hypothetical total fee for the fictional round 0 will be determined randomly. This fee will be between 1 and 135 cents. The cost and audit probability for the first round will be determined according to the table above. In all later rounds, cost and audit probability are determined on the basis of the actual total fees paid by your group in the previous round.

7. At the beginning of each round, every player will be informed of the amount of your group's total fees paid in the previous round. This determines the current audit probability according to the table above. (The following applies only in the partial and the full information treatments: Every player is informed of the amount of fees paid by every single group member.)

8. At the end of the 20 rounds, your individual income for every round is summed up. The payoff of your participation in this experiment corresponds to your total income.

9. Subsequent to the 20 rounds, there will be a lottery. Afterward, you have to answer a short questionnaire. The resulting payoff for your participation in this experiment may be picked up during the office's opening hours.

The experiment starts as soon as all participants are ready. Good luck!

B Questionnaire (extract)

1. Age (in years)
2. Gender (male/female)
3. Student of economics (yes/no)
4. Semester
5. Risk aversion (see Table 4)

C Screenshots from the experiment

Below are screenshots from the experiment. The experiment was conducted in German, hence, everything shown below was translated into English.

Please choose always the one alternative A or B you prefer. One of the ten lotteries you choose will be selected randomly. You will be credited the respective result.

Option A	Option B	Expected payoff difference
10/10 of €0.525	1/10 of €0.75, 9/10 of €0.41	€0.08
10/10 of €0.525	2/10 of €0.75, 8/10 of €0.41	€0.05
10/10 of €0.525	3/10 of €0.75, 7/10 of €0.41	€0.01
10/10 of €0.525	4/10 of €0.75, 6/10 of €0.41	€-0.02
10/10 of €0.525	5/10 of €0.75, 5/10 of €0.41	€-0.05
10/10 of €0.525	6/10 of €0.75, 4/10 of €0.41	€-0.09
10/10 of €0.525	7/10 of €0.75, 3/10 of €0.41	€-0.12
10/10 of €0.525	8/10 of €0.75, 2/10 of €0.41	€-0.16
10/10 of €0.525	9/10 of €0.75, 1/10 of €0.41	€-0.19
10/10 of €0.525	10/10 of €0.75, 0/10 of €0.41	€-0.23

Table 4: The ten paired lottery-choice decisions (Holt and Laury, 2002)

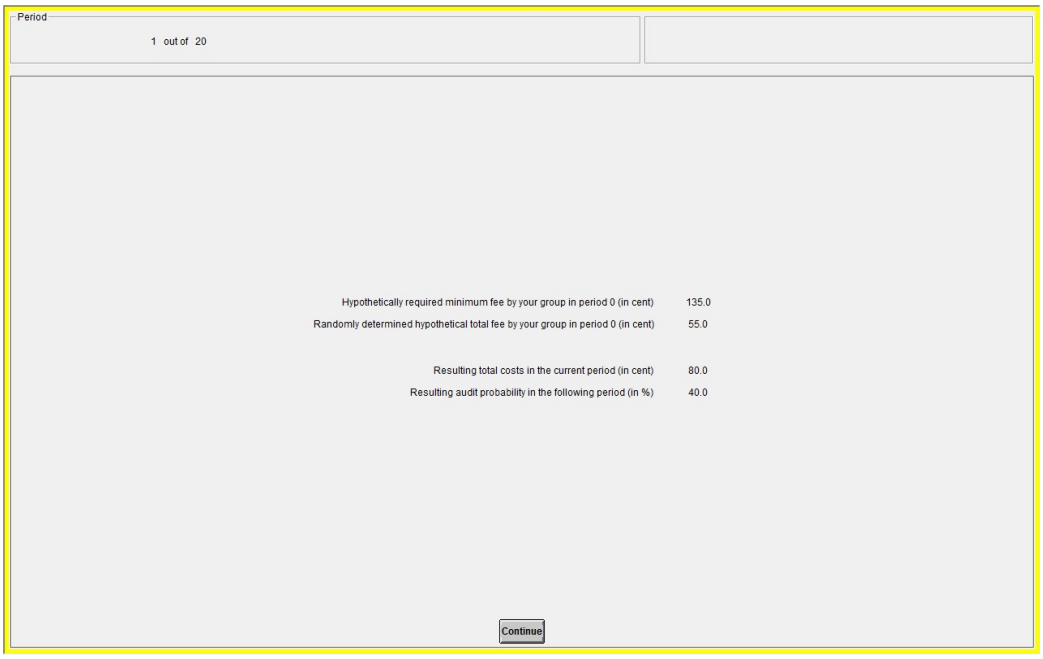

Figure 2: Information stage

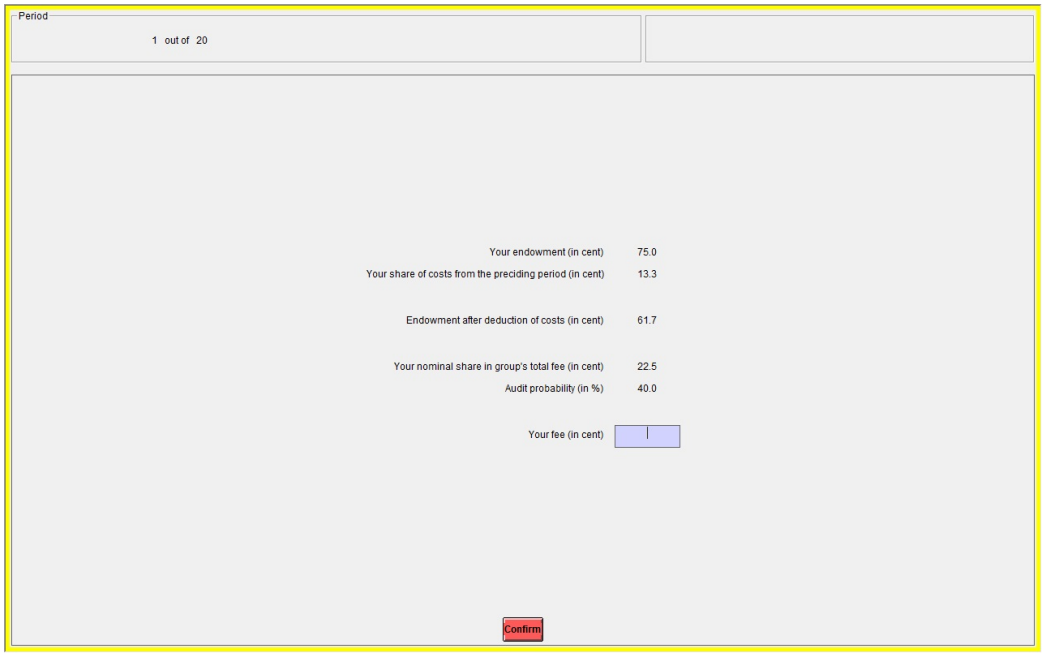

Figure 3: Contribution entry

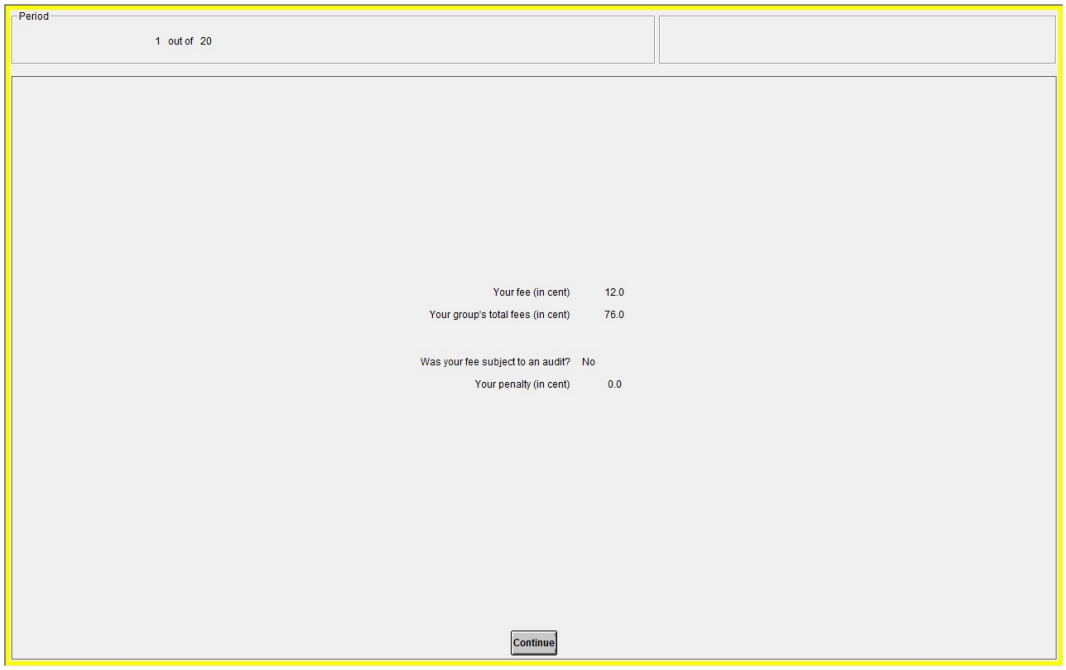

Figure 4: Audit stage

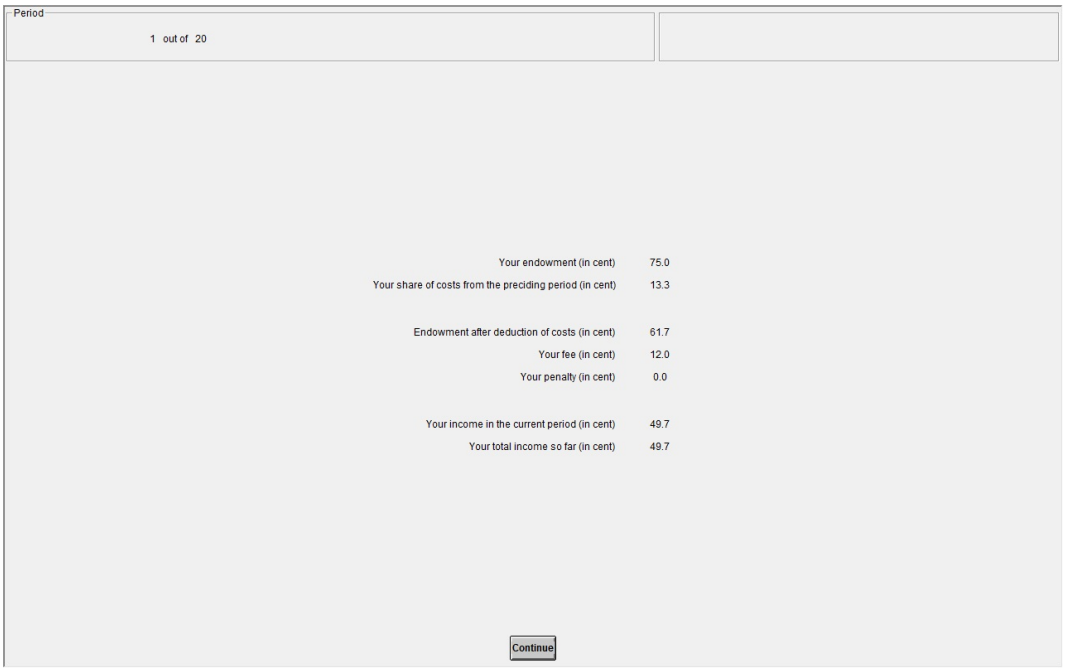

Figure 5: Profit display

The screenshot shown reflects the situation in the no information treatment, where no individual information is available. In both the partial and the full information treatment individual taxes of every subject are visible.

References

Allingham M.G., Sandmo A., 1972. Income tax evasion: a theoretical analysis. Journal of Public Economics 1, 323–338.

Alm J., 2010. Testing behavioral public economics theories in the laboraty. National Tax Journal 63, 635–658.

Alm J., 2012. Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. International Tax and Public Finance 19, 54–77.

Alm J., McClelland G.H., Schulze W.D., 1992. Why do people pay taxes? Journal of Public Economics 48, 21–38.

Andreoni J., Petrie R., 2004. Public goods experiments without confidentiality: a glimpse into fund-raising. Journal of Public Economics 88, 1605–1623.

Baldry J.C., 1986. Tax evasion is not a gamble: A report on two experiments. Economics Letters 22, 333–335.

Benabou R., Tirole J., 2006. Incentives and prosocial behavior. The American Economic Review 96, 1652–1678.

Bolle F., Otto P.E., 2010. A price is a signal: on intrinsic motivation, crowding-out, and crowding-in. Kyklos 63, 9–22.

Coricelli G., Joffily M., Montmarquette C., Villeval M.C., 2010. Cheating, emotions, and rationality: an experiment on tax evasion. Experimental Economics 13, 226–247.

Cullis J., Jones P., Savoia A., 2012. Social norms and tax compliance: Framing the decision to pay tax. The Journal of Socio-Economics 41, 159–168.

Dhami S., Al-Nowaihi A., 2007. Why do people pay taxes? prospect theory versus expected utility theory. Journal of Economic Behavior & Organization 64, 171–192.

Durham Y., Manly T.S., Ritsema C., 2012. The effects of income source, context, and income level on tax compliance decisions in a dynamic experiment. Journal of Economic Psychology (forthcoming).

Erard B., Feinstein J.S., 1994. The role of moral sentiments and audits perceptions in tax compliance. Public Finance 49, 70–89 (Supplement).

Fischbacher U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171–178.

Frey B.S., Torgler B., 2007. Tax morale and conditional cooperation. Journal of Comparative Economics 35, 136–159.

Gino F., Ayal S., Ariely D., 2009. Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. Psychological Science 20, 393–398.

Hasegawa M., Hoopes J.L., Ishida R., Slemrod J., 2012. The effect of public disclosure on reported taxable income: Evidence from individuals and corporations in Japan. SSRN Working Paper.

Holm S., 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6, 65–70.

Holt C.A., Laury S.K., 2002. Risk aversion and incentive effects. The American Economic Review 92, 1644–1655.

Huck S., Kübler D., Weibull J., 2012. Social norms and economic incentives in firms. Journal of Economic Behavior & Organization 83, 173–185.

Kastlunger B., Dressler S.G., Kirchler E., Mittone L., Voracek M., 2010. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). Journal of Economic Psychology 31, 542–552.

Laury S., Wallace S., 2005. Confidentiality and taxpayer compliance. National Tax Journal 58, 427–438.

Slemrod J., Thoresen T.O., Bø E.E., 2013. Taxes on the internet: Deterrence effects of public disclosure. CESifo Working Paper: Public Finance, No. 4107.

Traxler C., 2010. Social norms and conditional cooperative taxpayers. European Journal of Political Economy 26, 89–103.

Wartick M.L., Madeo S.A., Vines C.C., 1999. Reward dominance in tax-reporting experiments: The role of context. Journal of the American Taxation Association 21, 20–31.

Wilson J.Q., Kelling G.L., 1982. Broken windows. Atlantic Monthly 249, 29–38.

Who bears value added taxes? Evidence from Germany

Jonathan Bob[a]

[a]*Leibniz Universität Hannover*

Abstract Since 2007 value added tax rates were changed twice in Germany. On January 1st 2007, the regular rate was increased from 16 % to 19 % while the reduced rate remained unchanged. On January 1st 2010, the rate on accommodation services was cut from the regular rate of 19 % to the reduced rate of 7 %. Both tax rate changes are used to examine who bears value added taxes. Therefore, difference-in-differences estimation is applied to German consumer price data. The results show that consumers have to bear about 44 % of the additional tax burden. In addition, I find tax incidence to be depending on the market competitiveness. Finally, in the accommodation services industry prices respond asymmetrically to tax rate changes.

Keywords Value added taxes · Tax incidence · Difference-in-differences · Consumer price index

JEL classification H22 · H32

1 Introduction

Value added taxes are consumption taxes levied in more than 130 countries around the world. This includes all OECD member states with the exception of the United States (OECD, 2006). In these countries value added taxes usually raise more than one-fifth of total fiscal revenues (Keen and Lockwood, 2010). Value added taxes offer many advantages as they feature administrative simplicity and are commonly considered as economically efficient (Atkinson and Stiglitz, 1972).

Value added tax reforms are a popular and often used tool in fiscal policy. According to European Commission (2013), 20 out of 27 member states of the European Union varied their respective tax rates at least once within the last five years. Typically, there are two reasons for tax rate changes. On the one hand, governments increase rates to enlarge fiscal revenues. This was the main purpose of the German government when on June 29th 2006 then president Horst Köhler signed a bill into law that provided an increase in the regular tax rate from 16 % to 19 %. The law entered into force as from January 1st 2007 leaving the reduced rate of 7 % unchanged. The tax rate increase was reasoned by the seriousness of the fiscal situation in Germany. Spending exceeded revenues by far for many years (Deutscher Bundestag, 2006). On the other hand, tax rates are cut to kick-start the economy. This

was the intention when the rate on accommodation services was cut from the regular rate of 19 % to the reduced rate of 7 % as from January 1st 2010 to advance economic growth after the financial crisis of 2007–2008 (Deutscher Bundestag, 2009).

Value added taxes are indirect taxes. According to the basic principles in OECD (2006) tax incidence should completely fall to consumers, i.e., the consumer share of the tax burden should equal 100 %. However, it depends on the degree of tax shifting whether and to what extent tax incidence actually falls on consumers. The main purpose of this paper is to use both the 2007 and the 2010 tax reforms to estimate the incidence of value added taxes. The results show that the consumer share amounts to about 44 % of the additional tax burden with incidence being significantly higher in more competitive markets. Another finding is that in the accommodation services industry prices respond stronger to tax rate increases than to tax cuts.

These results are of particular interest for fiscal policy. Tax rate changes may lead to changing consumer prices and hence, affect consumers' real purchasing power. For instance, in the case of tax rate increases the intended effect of increasing fiscal revenues may change to the opposite when higher consumer prices lead to less consumption expenses. If tax incidence had fallen completely on consumers after the 2007 tax reform, prices of commodities that are taxed at the regular rate would have been increased by about 2.59 percentage points.[1] The German government wanted to prevent a slump in consumption expenses due to increasing consumer prices. Therefore, unemployment insurance contributions were cut for financial relief.[2] However, without studies on the incidence of value added taxes in Germany it is possible neither to conclude useful counteractions nor to evaluate whether the cut in unemployment insurance contributions was appropriate in size.

The remainder of this article is structured as follows. First, in section 2 a brief literature review is presented. In section 3 the data is introduced. In section 4 I illustrate how I measure the consumer share before the results are presented in section 5. Section 6 concludes.

2 Literature Review

Most research on tax incidence is theoretical work. These studies can be distinguished depending on the competitiveness of the market. Under perfect conditions tax incidence can be shown easily as it is well known from literature that marginal costs equal price. Since value added taxes are additional producers' costs, assuming a perfect elastic supply leads to consumer prices responding one-for-one to tax rate changes (e.g., Alm et al., 2009).

In contrast, the case of imperfect markets is more complicated with ambiguous results regarding consumer shares. Probably, the two most simple imperfect markets are the Cournot-Nash and the Bertrand oligopoly. Both oligopolies assume identical firms competing in quantities (Cournot-Nash) and prices (Bertrand), respectively. Price competition in the Bertrand oligopoly leads to prices equal marginal costs as it is under perfect conditions. Thus, taxes are fully shifted to final consumers. In contrast, many different incidence scenarios may result in a Cournot-Nash-model ranging from undershifting to even overshifting, meaning that prices increase more than taxes with primarily elasticities and the number of firms determining the actual degree of tax shifting. There are further imperfect markets as monopolies and spatial competition. Prices will usually respond one-for-one in monopolies

[1] The regular tax rate was increased from 16 % to 19 %. If prices respond one-for-one, they would increase by $100 \times 1.19/1.16 - 100 = 2.59$ percentage points.

[2] In Germany, unemployment insurance contributions are mandatory for all employees.

while spatial competition typically leads to undershifting with decreasing consumer share in less competitive markets (Katz and Rosen, 1985; Stern, 1987; Delipalla and Keen, 1992; Hamilton, 1999). A general overview on tax incidence for imperfect markets can be found in Fullerton and Metcalf (2002).

Even though most developed countries levy value added taxes and despite the fact that these taxes raise an important fraction of total fiscal revenues, there exist only few empirical studies on the incidence of value added taxes. Carbonnier (2005) applies difference-in-differences estimation to consumer price data to examine the impact of three valued added tax reforms in France. His results show that price responses are stronger for increasing than for decreasing tax rates. For less competitive markets he finds a consumer share of 91 % after a tax rate increase and 22 % after a tax rate cut. In contrast, on more competitive markets the consumer share amounts to 52 % after the tax rate increase and 130 % after the tax cut.

Carbonnier (2007) again uses difference-in-differences estimation and consumer price data to examine the effect of two value added tax reforms in France with the second one being already included in his 2005 study. In 1987 the tax rate on car sales was cut from 33.33 % to 18.6 % and in 1999 the tax rate on housing repair services was cut from 20.6 % to 5.5 %. For both reforms he finds evidence for undershifting with 57 % of the tax relief on car sales and 77 % of the tax relief on housing repair services passed to consumers. Due to the higher competitiveness in the market for housing repair services his results suggest a higher consumer share in more competitive markets.

Zapal (2007) applies difference-in-differences estimation to consumer price data to study the 2004 Czech value added tax reform that included tax rate increases as well as tax cuts. He shows that producers do not let consumers participate in tax cuts. For tax rate increases, the consumer share of the additional tax burden is positive but small.

There are some more studies for United States sales taxes. Poterba (1996) uses clothing prices for fourteen cities from 1925 to 1939 and for eight cities from 1947 to 1977 to examine the incidence of state and local retail sales taxes. He finds evidence for undershifting for the early time period and full tax shifting for the later period. Besley and Rosen (1999) use twelve commodities to estimate tax incidence in 155 cities around the United States. Their data ranges from 1982 to 1990. Using panel regression with fixed effects, they find full shifting as well as overshifting. Both Poterba (1996) and Besley and Rosen (1999) find evidence that prices fully respond to tax rate changes within the first three to four months.

3 Data

The data used in this study is taken from the German consumer price index. It depicts the average consumer price changes in Germany and is published as time series on a monthly basis by the German Federal Statistical Office.[3] The index is based on a basket of commodities that contains all goods and services German households purchase for consumption purposes.

Currently, the basket of commodities consists of 667 different goods and services that are combined into 106 classes, 42 groups and 12 divisions building up the overall index.[4] Table 1 provides an overview of the structure of the consumer price index that is an adapted version

[3]Data is free for download at the GENESIS-online database of the German Federal Statistical Office, `https://www-genesis.destatis.de/genesis/online`.

[4]See `https://www.destatis.de/EN/Meta/abisz/VPI_e.html` for details.

coicop-cpi	item
07	Transport
071	Purchase of vehicles
0711	Motor cars
	New cars
	Utility vans
	⋮
0712	Motor cycles
⋮ ⋮ ⋮	⋮ ⋮ ⋮ ⋮

Table 1: Structure of the cosumer price index

of coicop by United Nations Statistics.[5] As it can be seen from table 1 different commodities as new cars and utility vans are wrapped up in the class of motor cars. The classes of motor cars, motor cycles and other are a component of the purchase of vehicles group. This group itself is featured in the transport division.

The price of each commodity is collected every month. To calculate the overall index, the average price of each product is weighted according to a given pattern, resulting in a weighted average. This pattern reflects the proportion of expenses that households allocate for the respective product. Such a pattern is available for divisions, groups and classes.[6]

The consumer price index offers two major advantages for my purpose. First, all prices are gross prices and hence, value added taxes are included. Thus, tax rates and price responses on tax rate changes are implicitly incorporated in the data. Secondly, the weighting pattern allows for disassembling and reassembling the index. This is particularly important because the tax rate change should only affect those commodities that are subject to it. Therefore, regarding the 2007 tax reform I must separate commodities that are taxed at the regular rate from commodities that are taxed at the reduced rate or are even tax-exempt.[7]

Some classes have to be excluded from the analysis. The exclusion reasons are outlined in the following.

In most of the 12 divisions the respective goods and services are taxed at different tax rates. As it can be seen in table 1, division 07 consists of transport commodities. E.g., the purchase of vehicles is taxed at the regular rate but many other transport services as short distance train tickets or cab rides are taxed at the reduced rate. However, at the four-digits classes level a unique tax rate can mostly be assigned. The 10 classes that are not subject to a unique tax rate are excluded from the analysis.[8]

[5] Coicop is the abbreviation for Classification of Individual Consumption According to Purpose.

[6] The weighting pattern is available for download on the website of the German federal statistical office (in German only), `https://www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/Preise/Verbraucherpreisindizes/WarenkorbWaegungsschema/Waegungsschema.pdf?__blob=publicationFile`.

[7] In the case of the 2010 tax reform this is much easier as only accommodation services have to be separated. However, the following remarks apply for both the 2007 and the 2010 tax rate changes.

[8] I call a tax rate unique when each commodity of one class is taxed at the same tax rate. 58 classes are completely taxed at the regular rate, 16 at the reduced rate and another 13 are tax-exempt. 9 further classes show almost unique tax rates, i.e., at least 90 % of the commodities are taxed at the same rate. These 9 classes are included in the analysis, too. A robustness check shows that excluding these 9 classes leads to a slightly lower consumer share.

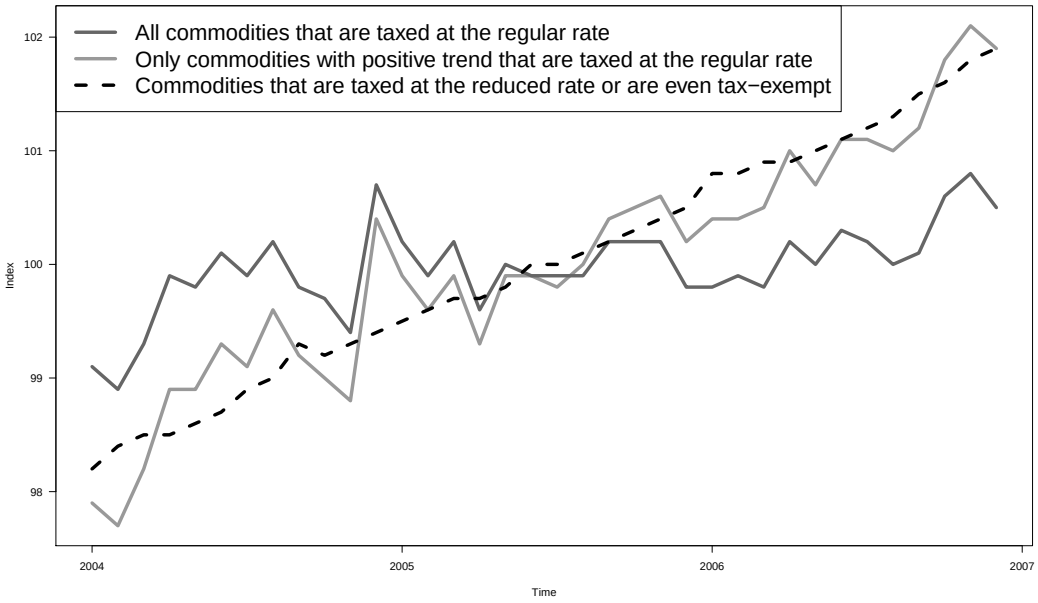

Figure 1: Impact of classes with negative trend on consumer prices

Many studies recommend to measure inflation based on consumer prices excluding food and energy because their higher volatility may result in noised data (e.g., Cogley, 2002; Rich and Steindel, 2005). Indeed, food and energy feature a very high volatility and an exceptional increase over time, especially for energy as four of the eight classes with the highest increase from 2004 to 2009 belong to energy. Therefore, both food and energy are excluded from the analysis.[9]

Moreover, I exclude nine classes that show a negative trend. All of these classes are taxed at the regular rate. As it will be pointed out in section 4 I use differences between commodities that are taxed at the regular rate and commodities that are not to determine the effect of the tax reform. The basic assumption is that both types of commodities share a common trend. As it can be seen in figure 1 this assumption is only fulfilled without these nine classes. Therefore, they are excluded from the analysis.

Overall, 33 of the 106 classes are excluded. According to the weighting pattern, these classes account for about 31 % of total consumption expenses of German households. With regard to the 2007 tax reform the ratio of commodities that are taxed at the regular rate and commodities that are not remains almost unchanged. This ratio equals 53 % for all commodities and 51 % for commodities that are included in the study. An overview of all 106 classes, their respective (average) tax rates and whether the class is included in the analysis is given in table 5 in Appendix A.

[9]Food is almost exclusively taxed at the reduced rate while energy is taxed at the regular rate. Including both food and energy leads to a consumer share of more than 100 %. Excluding only food leads to an even higher consumer share of about 145 %, while there is no significant consumer share at all if only energy is excluded.

As it is typical for time series data on a monthly basis, consumer prices show seasonal patterns. I seasonally adjust the data at the classes level assuming a additive deterministic component model.[10] Thus, I adjust for seasonal effects before they get mixed up reassembling the index.

The following analysis is based on the seasonal adjusted data for the remaining 73 classes. I use data ranging from 3 years before the respective tax rate change to 3 years after.

4 Model

The main purpose of this paper is to estimate the tax incidence of value added taxes based on both the 2007 and the 2010 value added tax reforms. Hence, I am interested in how much of the additional tax burden is shifted to the final consumer. Therefore, I proceed as follows. First, I estimate the changes in consumer prices due to the tax rate change only. Secondly, I calculate the changes in consumer prices as if prices would respond one-for-one. Then, consumer share can be calculated as ratio of both price changes.[11]

The change in consumer prices due the tax rate change only is called the causal or treatment effect of the tax rate change. 48 of the 73 classes included in the analysis receive the treatment, i.e., these classes are subject to the tax rate change. They are said to be in the treatment group. The remaining 25 commodities are assigned to the control group. Assuming that all consumer prices share a common trend,[12] no matter whether or not they belong to the treatment group, difference-in-differences (DID)[13] can be used to estimate the treatment effect. If the common trend assumption holds the treatment effect is just the deviation from this trend. This is visualized in figure 2.

In figure 2 two hypothetical time series for the 2007 tax reform are shown. One series represents the treatment group and the other the control group. As it can be seen, there are no systematic difference between the slope of both series neither before nor after the 2007 tax reform. Hence, the treatment effect is given by the offset, i.e., the deviation from the common trend.

DID can be estimated by OLS. Then, the regression equation is given by

$$p_i = \alpha + \beta\, \mathit{trend} + \gamma_1\, \mathit{treatment} + \gamma_2\, \mathit{time} + \gamma_3\, (\mathit{treatment} \times \mathit{time}) + \varepsilon_i. \tag{1}$$

The response variable is the consumer price p_i. The dummy variable *treatment* equals 1 for commodities in the treatment group. Analogously, the dummy variable *time* equals 1 when the treatment is in effect, i.e., after the respective tax reform. The treatment effect is given by the coefficient γ_3 of the interaction *treatment* $\times$ *time*. The *trend* variable is included to control for the usual trend in consumer prices.[14] The error term ε_i is assumed

[10] Such an additive component model is given by $Y_t = G_t + S_t + I_t$ with Y_t begin the response variable, G_t the trend, S_t the seasonal component and I_t the irregular or random component. First, G_t is calculated using a centered 12-month moving average. Seasonal and irregular components are obtained by $Y_t - G_t = S_t + I_t$. The irregular component can be subtracted by averaging the values of the combined seasonal and irregular components for each month. The monthly averages correspond to the seasonal component S_t. Finally, seasonally adjusted data can be calculated by $Y_t - S_t = G_t + I_t$.

[11] Consumer share equals 1 if consumer prices respond one-for-one. Hence, dividing by this change results in the actual consumer share.

[12] The common trend assumption states that prices in both the treatment and the control groups would be identical without the treatment.

[13] See Angrist and Pischke (2008) for more details on difference-in-differences estimation.

[14] This variable equals 1 for the first observation, 2 for the second, and so forth.

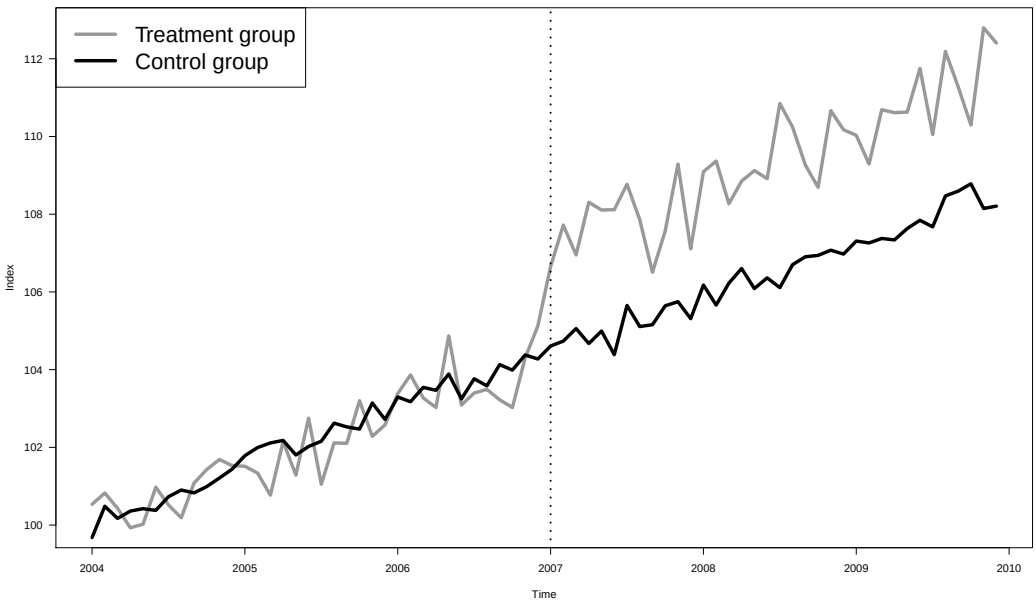

Figure 2: Common trend assumption (hypothetical time series)

to be white noise. Generally, this assumption does not hold for time series data because of heteroscedasticity. Therefore, I correct the estimated standard errors based on White.

As noted above, the consumer share is given by the ratio of the treatment effect γ_3 and the theoretical effect if consumer prices respond one-for-one. The latter one can be calculated easily. In 2007 the regular tax rate was increased from 16 % to 19 %. If consumer prices that are taxed at the regular rate had responded one-for-one they would have increased by

$$100 \times 1.19/1.16 - 100 = 2.59$$

percentage points. Analogously, the change in prices for accommodation services after the 2010 tax reform is given by

$$100 \times 1.07/1.19 - 100 = -10.08$$

percentage points.

Then, the consumer share θ of the additional tax burden is given by

$$\theta = \begin{cases} +\dfrac{\gamma_3}{2.59} & \text{for the 2007 tax reform,} \\ \\ -\dfrac{\gamma_3}{10.08} & \text{for the 2010 tax reform.} \end{cases} \tag{2}$$

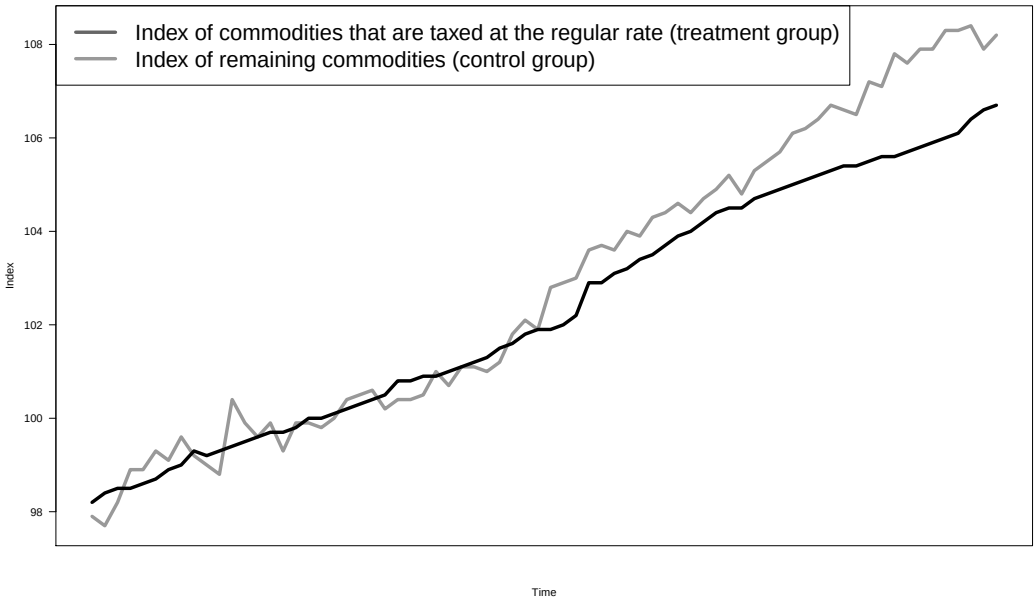

Figure 3: Overall effect of the 2007 tax reform on consumer prices

5 Results

5.1 Tax incidence after the 2007 tax reform

First, I look at the the 2007 tax reform. Therefore, I use the seasonally adjusted data described in section 3 ranging from January 2004 to December 2009. I split the data into two subsets with the first one containing the commodities of the treatment group. These commodities are affected by the tax reform. The second subset includes the commodities of the control group. The resulting time series are shown in figure 3.

As it can be seen from figure 3, before the tax reform there seems to be no systematic difference between both indices. They fluctuate around each other with the index of the treatment group showing higher variation. In January 2007 the prices of the treatment group increases by about 1 percentage point more than the prices of the control group, immediately resulting in an offset between both indices. Even though there is a similar increase in the control group shortly after, the offset persists clearly. This indicates that prices immediately respond to tax rate changes.

In 2009, the offset seems to be getting even bigger, indicating an increasing slope in the treatment group that may result in a slight overestimation of the treatment effect.

To test this effect statistically, I apply equation (1) on the data. The results are shown in table 2. As it can be seen from this table, the treatment effect is highly significant.[15] Consumer prices increase by about 1.13 percentage points due to the tax rate change only.

[15] The treatment effect is given by the coefficient γ_3 of the interaction between time and treatment, see section 4.

	Coef.	Std.err.	Sign.	
Intercept	97.78	0.08	< 0.01	***
Trend	0.12	0.00	< 0.01	***
Time	0.24	0.15	0.11	
Treatment	-0.01	0.09	0.90	
Time × Treatment	1.13	0.13	< 0.01	***
	Consumer share: 44 %			

Response variable: consumer price

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Robust standard errors, adj. $R^2 = 0.98$

Table 2: Effect of the 2007 tax reform on consumer prices

As noted before, consumer prices would have increased by about 2.59 percentage points if they had responded one-for-one. Thus, the consumer share θ equals $1.13/2.59 = 0.44$. The additional tax burden resulting from the 2007 tax rate increase is almost equally split between producers and consumers. Consumer share is somewhat lower than in Carbonnier (2005, 2007) and well below the results of Poterba (1996) and Besley and Rosen (1999).

This is an important results for fiscal policy. Based on θ, it would be possible to calculate the net effects of the tax rate increase and the cut in unemployment insurance contributions on consumers' real purchasing power.[16] Moreover, this result enables politicians to estimate price responses on future tax rate changes beforehand.

As noted before, about 51 % of all consumption expenses of German households are apportioned to commodities that are taxed at the regular rate. This leads to an overall increase in consumer prices of $0.51 \times 1.13 = 0.57$ percentage points due to the tax rate change only. In the case of full tax shifting consumer prices would have increased by $0.51 \times 2.59 = 1.32$ percentage points.

One should note the R^2 of almost 1. This is mostly due to the fact that the upward trend is almost identical for both indices. Unsurprisingly, the coefficient of this deterministic trend is highly significant. The average monthly increase in consumer prices amounts to 0.12 percentage points. Both the coefficients of the time and the treatment dummy variable are insignificant which implies that there are differences neither in the control group over time nor between the treatment and control group before the tax reform.

5.2 Consumer share and market competitiveness

In section 2 it was pointed out that market competitiveness may affect tax incidence. Usually, the consumer share of the tax burden will increases with increasing market competition.

Normally, goods are capital intensive and services are labor intensive. Assuming that competitiveness in labour intensive markets is higher because of higher market entry barriers this would lead to a higher consumer share for services compared to goods (Carbonnier, 2007). To examine this I split the subset of commodities from the treatment group into further subsets with the first subset containing goods and the second subset containing

[16] However, this will not be covered here.

	Goods				Services			
	Coef.	Std.err.	Sign.		Coef.	Std.err.	Sign.	
Intercept	97.92	0.06	< 0.01	***	97.50	0.13	< 0.01	***
Trend	0.11	0.00	< 0.01	***	0.14	0.01	< 0.01	***
Time	0.51	0.14	< 0.01	***	-0.30	0.22	0.18	
Treatment	-0.04	0.07	0.62		0.04	0.14	0.75	
Time × Treatment	0.55	0.11	< 0.01	***	2.30	0.22	< 0.01	***
	Consumer share: 21 %				Consumer share: 89 %			

Robust standard errors
Adj. $R^2 = 0.99$ (goods) and adj. $R^2 = 0.96$ (services)

Table 3: Consumer share and market competitiveness

services. The control group remains unchanged. Again, I will test this effect statistically by estimating equation (1) based on these data. The results for goods and services are shown in the left and right panel of table 3, respectively.

As it can be seen in table 3 the treatment effect is highly significant for both goods and services. Prices increase by about 0.55 and 2.30 percentage points due to the tax reform, respectively. The consumer share of the additional tax burden amounts to 21 % for goods and 89 % for services. Hence, most of the overall treatment effect stems from services.[17]

The higher consumer share for services is theoretically expected and in line with the findings of Carbonnier (2005, 2007). The difference in γ_3 between services and goods is highy significant (2-sample t-test, $p < 0.01$). Moreover, the γ_3 coefficient for services is only weakly significant smaller than 2.59 (t-test, $p = 0.095$), i.e., the additional burden of services is almost fully shifted to final consumers.

Again, R^2 is very high and the deterministic trend is highly significant in both estimations.

5.3 Asymmetric price responses

The tax rate on accommodation services was changed twice. First, in 2007 the tax rate was increased from 16 % to 19 %. In 2010, this tax rate was cut from 19 % to 7 %. This allows to test whether consumer prices respond asymmetric to different kinds of tax rate changes.

Regarding the 2007 tax reform, accommodation services remain the only commodity in the treatment group. As in sections 5.1 and 5.2 all commodities that are taxed at the regular rate or are tax exempt are included in the control group. Analogously for the 2010 tax reform, the treatment group consists only of accommodation services but the control group comprises all other commodities no matter at which rate they are taxed. Again, (1) is applied to the data. The results for 2007 and 2010 are shown in the left and right panel of table 4, respectively.

As it can be seen from table 4 the treatment effect on accommodation services after the 2007 tax rate increase is highly significant. Prices for accommodation services increased by about 0.99 percentage points. The consumer share amounts to 38 %. As before, R^2 is very high and the deterministic trend is highly significant. Regarding the 2010 tax reform the

[17] 66 % of all commodities from the treatment group are goods. Hence, the overall effect is given by the weighted average $0.66 \times 0.21 + 0.34 \times 0.89 = 0.44$. This holds for the other coefficients as well.

	2007 tax reform				2010 tax reform			
	Coef.	Std.err.	Sign.		Coef.	Std.err.	Sign.	
Intercept	97.68	0.14	< 0.01	***	102.86	0.15	< 0.01	***
Trend	0.13	0.01	< 0.01	***	0.16	0.01	< 0.01	***
Time	0.06	0.26	0.81		-1.23	0.30	< 0.01	***
Treatment	-0.00	0.19	0.99		-0.28	0.21	0.19	
Time × Treatment	0.99	0.27	< 0.01	***	0.43	0.27	0.11	
	Consumer share: 38 %				Consumer share: 0 %			

Robust standard errors
Adj. $R^2 = 0.93$ (2007) and adj. $R^2 = 0.93$ (2010)

Table 4: Price responses of accommodation services to opposing tax rate changes

positive sign of the coefficient γ_3 might indicate that prices of accommodation services are even increased after the tax cut. However, the coefficient is not significant. The tax relief is not passed to consumers at all. Consumer prices respond stronger to tax rate increases compared to tax rate cuts. This suggest asymmetric responses in consumer prices.

Note that the 2010 tax cut amounts to 12 percentage points. In contrast the 2007 tax increase amounts to only 3 percentage points. Hence, effects of the 2010 tax rate cut should be easier to identify.

I neglect the fact that not all prices of accommodation services are part of the tax cut as breakfast is still taxed at the regular rate. However, this should only have a very small impact on the results and cannot account for the missing price decease.

6 Conclusions

The main purpose of this paper was to determine the incidence of value added taxes based on both the 2007 and the 2010 tax reforms. Therefore, I conducted difference-in-differences estimation on German consumer price data. Using data ranging from 3 years before the tax reform to 3 years after it, the results show that consumers have to bear about 44 % of the additional tax burden after the 2007 reform.

In addition, I find incidence to be depending on the market competitiveness. As theoretically expected consumer shares are higher in more competitive markets (89 %) than in less competitive markets (21 %). Finally, in the case of accommodation services price respond asymmetrically to tax rate changes. After the 2007 reform 38 % of the additional tax burden are shifted to final consumers. In contrast, the tax relief after the 2010 reform is not shifted to consumers at all.

These results are in line with previous studies regarding the incidence of value added taxes by Carbonnier (2005, 2007) and Zapal (2007). As compared to the United States sales tax consumer shares of value added taxes are clearly smaller. This may be due to tax salience as usually, price tags show net prices (prices excluding sales taxes) in the United States. Net prices do not need to be changed to shift taxes.

There are some limitations to my study. The tax rate increase in 2007 amounts to just 3 percentage points. Effects should be easier to identify for higher tax rate changes. Moreover,

the offset between consumer prices that are subject to the 2007 tax reform and those that are not seem to raise over time. This may lead to an overestimation of actual consumer shares.

The results presented in this paper provide starting points for further research. It would be interesting to examine whether incidence is constant over time by studying the 1993 and 1998 German value added tax reforms. Moreover, since most European Union member states changed their value added tax rates within the last five years, it could be tested whether incidence is constant across these countries. This is suggested by the results of Carbonnier (2005, 2007) and Zapal (2007). Moreover, it could be tested experimentally whether salience actually has an impact on tax incidence. Finally, the results are important for fiscal policy. Consumer price reactions to future tax rate changes can be anticipated and hence, useful counteractions may be applied.

A Classes of the consumer price index

Class	Label	Weight (in ‰)	tax rate (in %)	incl.
0111	Bread and cereals	16.44	7.0	no
0112	Meat	21.54	7.0	no
0113	Fish	3.11	7.0	no
0114	Milk, cheese and eggs	14.44	7.0	no
0115	Oils and fats	2.55	7.0	no
0116	Fruit	9.23	7.0	no
0117	Vegetables	10.60	7.0	no
0118	Sugar, jam, honey, chocolate and confectionery	7.70	7.0	no
0119	Food products n.e.c.	4.38	7.0	no
0121	Coffee, tea and cocoa	3.95	7.0	yes
0122	Mineral waters, soft drinks, fruit and vegetable juices	9.61	19.0	yes
0211	Spirits	1.88	19.0	yes
0212	Wine	5.56	19.0	yes
0213	Beer	9.03	19.0	yes
0214	Ready-mixed drinks with an alcohol content of < 6 %	0.09	19.0	yes
0220	Tobacco	22.43	19.0	yes
0312	Articles of clothing garments	37.10	19.0	yes
0313	Other articles of clothing and clothing accessoires	1.12	19.0	yes
0314	Cleaning, repair and hire of clothing	1.20	19.0	yes
0321	Footwear	9.17	19.0	yes
0322	Repair of footwear	0.29	19.0	yes
0411	Actual rentals for housing	203.30	0.0	yes
0431	Materials for the maintenance and repair of the dwelling	6.80	19.0	yes
0432	Services for the maintenance and repair of the dwelling	5.04	19.0	yes
0441	Water supply	11.09	7.0	yes
0442	Refuse collection	6.84	0.0	yes
0443	Effluent disposal	9.35	0.0	yes
0444	Other services relating to the dwelling n.e.c.	5.76	11.9	no
0451	Electricity	24.61	19.0	no
0452	Gas	12.85	19.0	no
0453	Liquid fuels	9.21	19.0	no

Class	Label	Weight (in ‰)	tax rate (in %)	incl.
0454	Solid fuels	0.79	13.2	no
0455	Heat energy	12.36	19.0	no
0511	Furniture and furnishings	22.16	19.0	yes
0512	Carpets and other floor coverings	3.11	19.0	yes
0513	Repair of furniture, furnishings and floor coverings	1.23	19.0	yes
0520	Household textiles	4.07	19.0	yes
0531	Major household appliances whether electric or not	6.76	19.0	no
0532	Small electric household appliances	1.04	19.0	yes
0533	Repair of household appliances	0.79	19.0	yes
0540	Glassware, tableware and household utensils	4.05	19.0	yes
0551	Tools and equipment for house	1.83	19.0	yes
0552	Tools and equipment for garden	3.10	19.0	yes
0561	Non-durable household goods	4.96	19.0	yes
0562	Domestic services and household services	2.77	19.0	yes
0611	Pharmaceutical products	9.83	17.6	yes
0612	Other medical products	1.21	19.0	yes
0613	Therapeutic appliances and equipment	6.51	19.0	yes
0621	Medical services	8.32	0.0	yes
0622	Dental services	5.59	4.0	yes
0623	Paramedical services	2.19	0.0	yes
0630	Hospital services	6.62	0.0	yes
0711	Motor cars	34.84	19.0	yes
0712	Motor cycles	1.23	19.0	yes
0713	Bicycles	1.43	19.0	yes
0721	Spare parts and acces. for pers. transp. equip.	6.62	19.0	yes
0722	Fuels and lubricants for pers. transp. equip.	35.91	19.0	no
0723	Maintenance and repair of pers. transp. equip.	20.32	19.0	yes
0724	Other services in respect of pers. transp. equip.	12.72	8.5	no
0731	Passenger transport by railway	5.53	16.0	no
0732	Passenger transport by road	1.15	7.0	yes
0733	Passenger transport by air	2.58	1.6	yes
0734	Passenger transport by sea and inland waterway	0.62	7.0	yes
0735	Combined passenger transport	8.44	7.0	yes
0736	Other purchased transport services	0.51	19.0	yes
0810	Postal services	2.28	2.2	yes
0820	Telephone and telefax equipment	1.60	19.0	no
0830	Telephone and telefax services	27.12	19.0	no
0911	Equip. f. reception, rec. and reprod. of sound and pict.	4.80	19.0	no
0912	Photogr. and cinematogr. equip. and opt. instr.	2.43	19.0	no
0913	Information processing equipment	6.98	19.0	no
0914	Recording media	3.40	19.0	no
0915	Repair of aud.-vis., photogr. and inform. proces. equip.	1.40	19.0	yes
0921	Other major durables for recreation	1.84	19.0	yes
0922	Other major durables for culture	0.82	19.0	yes
0931	Games, toys and hobbies	5.72	19.0	no
0932	Equipment for sport, camping and open-air recreation	3.04	19.0	no
0933	Gardens, plants and flowers	7.64	9.0	no
0934	Pets and related products	3.60	11.3	no

Class	Label	Weight (in ‰)	tax rate (in %)	incl.
0935	Veterinary and other services for pets	1.53	19.0	yes
0941	Recreational and sporting services	7.38	14.0	no
0942	Cultural services	15.80	6.4	no
0943	Games of chance	5.81	0.0	yes
0951	Books	6.29	7.0	yes
0952	Newspapers and periodicals	7.76	7.0	yes
0953	Miscellaneous printed matter	0.93	15.4	no
0954	Stationery and drawing materials	2.58	19.0	yes
0960	Package holidays	25.93	19.0	yes
1010	Primary education services	3.16	0.0	yes
1040	Secondary education services	2.00	0.0	yes
1050	Tertiary education services	2.24	19.0	yes
1111	Restaurants, cafes and the like	28.81	17.9	yes
1112	Canteens	3.31	18.4	yes
1120	Accommodation services	11.87	18.5	yes
1211	Hairdressing salons and personal grooming establishm.	9.58	19.0	yes
1212	Electric appliances for personal care	0.49	19.0	yes
1213	Other appliances, articles and prod. for pers. care	11.47	19.0	yes
1231	Jewellery, clocks and watches	3.02	19.0	yes
1232	Other personal effects	2.80	19.0	yes
1240	Social protection	11.81	0.1	yes
1252	Insurance connected with the dwelling	1.98	0.0	yes
1253	Insurance connected with health	9.39	0.0	yes
1254	Insurance connected with transport	9.53	0.0	yes
1255	Other insurance	3.98	0.0	yes
1262	Financial services n.e.c.	5.34	4.5	yes
1270	Other services n.e.c.	5.08	12.5	no
		1000.00	11.2	

Weights are according to the 2005 weighting pattern of the German federal statistical office. Tax rates are according to Elbel and Werner (2008) and after the 2007 tax reform. "Include" indicates whether a class is included in the estimation. Exclusion reasons are outlined in section 3. Food comprises classes 0111 to 0119, while energy comprises classes 0451 to 0455 and 0722. Classes 0531, 0820, 0830, 0911 to 0914 and 0931 to 0932 show a negative trend. All remaining exclusions are due to mixed tax rates.

Table 5: Classes of the consumer price index

References

Alm, J., Sennoga, E., Skidmore, M., 2009. Perfect competition, urbanization, and tax incidence in the retail gasoline market. Economic Inquiry 47, 118–134.

Angrist, J. D., Pischke, J.-S., 2008. Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Atkinson, A. B., Stiglitz, J. E., 1972. The structure of indirect taxation and economic efficiency. Journal of Public Economics 1, 97–119.

Besley, T. J., Rosen, H. S., 1999. Sales taxes and prices: An empirical analysis. National Tax Journal 52, 157–178.

Carbonnier, C., 2005. Is tax shifting asymmetric? Evidence from French VAT reforms, 1995–2000, Paris-Jourdan Sciences Economiques, Working Paper 2005-34.

Carbonnier, C., 2007. Who pays sales taxes? Evidence from french vat reforms, 1987–1999. Journal of Public Economics 91, 1219–1229.

Cogley, T., 2002. A simple adaptive measure of core inflation. Journal of Money, Credit and Banking 34, 94–113.

Delipalla, S., Keen, M., 1992. The comparison between ad valorem and specific taxation under imperfect competition. Journal of Public Economics 49, 351–367.

Deutscher Bundestag, 2006. Gesetzentwurf der Bundesregierung – Entwurf eines Haushaltsbegleitgesetzes 2006, Drucksache 16/752 vom 17.03.2006.

Deutscher Bundestag, 2009. Gesetzentwurf der Fraktionen der CDU/CSU und FDP – Entwurf eines Gesetzes zur Beschleunigung des Wirtschaftswachstums, Drucksache 17/15 vom 09.11.2009.

Elbel, G., Werner, P., 2008. Die Mehrwertsteuererhöhung zum 1. Januar 2007 – Auswirkungen auf den Verbraucherpreisindex für Deutschland. Wirtschaft und Statistik 2008, 693–703.

European Commission, 2013. VAT rates applied in the member states of the European Union.

Fullerton, D., Metcalf, G. E., 2002. Tax incidence. In: Auerbach, A. J., Feldstein, M. (Eds.), Handbook of Public Economics 4. Elsevier, Amsterdam, pp. 1787–1872.

Hamilton, S. F., 1999. Tax incidence under oligopoly: A comparison of policy approaches. Journal of Public Economics 71, 233–245.

Katz, M., Rosen, H., 1985. Tax analysis in an oligopoly model. Public Finance Quarterly 13, 3–19.

Keen, M., Lockwood, B., 2010. The value added tax: Its causes and consequences. Journal of Development Eonomics 92, 138–151.

OECD, 2006. International VAT/GST guidelines, February 2006.

Poterba, J. M., 1996. Retail price reactions to changes in state and local sales taxes. National Tax Journal 49, 165–176.

Rich, R., Steindel, C., 2005. A review of core inflation and an evaluation of its measures, Staff Report, Federal Reserve Bank of New York, No. 236.

Stern, N., 1987. The effects of taxation price control and government contracts in oligopoly and monopolistic competition. Journal of Public Economics 32, 133–158.

Zapal, J., 2007. Consumption tax incidence: Evidence from the natural experiment in the Czech Republic, Working Paper.