Angewandte Mathematik und Optimierung Schriftenreihe
Applied Mathematics and Optimization Series
AMOS # 35(2015)

Ingmar Vierhaus, Roel van Veldhuizen, and Armin Fügenschuh

Strategic Uncertainty in Non-renewable Resource Markets: a Level-k Approach

Herausgegeben von der
Professur für Angewandte Mathematik
Professor Dr. rer. nat. Armin Fügenschuh

Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg
Fachbereich Maschinenbau
Holstenhofweg 85
D-22043 Hamburg

Telefon: +49 (0)40 6541 3540
Fax: +49 (0)40 6541 3672

e-mail: appliedmath@hsu-hh.de
URL: http://www.hsu-hh.de/am

Angewandte Mathematik und Optimierung Schriftenreihe (AMOS), ISSN-Print 2199-1928
Angewandte Mathematik und Optimierung Schriftenreihe (AMOS), ISSN-Internet 2199-1936

Strategic Uncertainty in Markets for Non-renewable Resources: a Level-k Approach

I. Vierhaus[1], R. van Veldhuizen[2], A. Fügenschuh[3]

[1] Zuse Institute Berlin, Department Optimization, Germany

[2] WZB Berlin Social Science Center, Germany

[3] Helmut Schmidt University / University of the Federal Armed Forces Hamburg, Germany

Abstract:
Existing models of non-renewable resources assume that sophisticated agents compete with other sophisticated agents. This study instead uses a level-k approach to examine cases where the focal agent is uncertain about the strategy of his opponent, or predicts the opponent will act in a non-sophisticated manner. We show that lower levels of strategic reasoning are closer to the Walras and collusive benchmark, whereas higher level strategies converge to the Nash-Hotelling equilibrium. These results are then fitted to experimental data, suggesting that the level of sophistication of participants increased over the course of the experiment.

Keywords:
Level-k, behavioral economics, non-renewable resources, laboratory experiment, optimization.
JEL Classification: D03, C90, Q30, L13, C61

1 Introduction

Existing models of nonrenewable resource markets assume that optimizing agents compete against other optimizing agents in environments characterized by Cournot or Stackelberg competition (see Groot et al., 2003; Loury, 1986; Newbery, 1981; Polasky, 1992; Salo and Tahvonen, 2001, and many others). This implies that agents will, in equilibrium, have a perfect understanding of what their opponents will do. In reality, however, agents may experience considerable uncertainty about what strategy the opponent will follow. For example, agents may not have all the information (such as cost functions, stock sizes, etc.) required to calculate an opponent's Nash equilibrium strategy (Pindyck, 1980; Reynolds, 2013). In such cases, they must form beliefs about the possible strategies chosen by the opponent, and maximize their expected profits given these beliefs.

Similarly, even with perfect information, agents may in some cases expect their opponent not to follow the Nash-Hotelling equilibrium. For example, the opponent may face political pressure to maximize current revenue and produce at maximal capacity, or otherwise be too focused on maximizing present revenue rather than discounted profits (see e.g., Farrow, 1985; Pindyck, 1981; Spiro, 2012). In such cases, agents must again form beliefs about the possible strategies chosen by the opponent, and maximize their expected profits given these beliefs.

We model the ensuing uncertainty about the opponent's chosen strategy using a level-k framework (see e.g., Camerer et al., 2004, 1998; Nagel, 1995; Runco, 2013a,b; Stahl and Wilson, 1994), a commonly used approach to model non-equilibrium opponents in behavioral economics. The framework starts by specifying the strategy for a level-0 player, who is argued to choose a random production trajectory from the set of all possible trajectories. Level-k players then best respond to the strategy of a level-$(k-1)$ opponent. We use a standard linear demand function that allows us to compute the Nash, collusive and Walrasian benchmarks. We then show that higher level strategies converge to the Nash-Hotelling equilibrium, while lower level strategies may closely approximate the Walras and collusive benchmark. We then fit the results to experimental data from Van Veldhuizen and Sonnemans (2014) to empirically estimate the distribution of types, and find that participants use more sophisticated strategies in latter parts of the experiment.

2 Model and Benchmarks

We assume that two players, player k and player $k-1$ are active in a nonrenewable resource market characterized by linear demand and Cournot competition. (The reason for giving players integer numbers as identifiers is that we later will assume player k plays the level k.) Both players are assumed to start with a fixed stock of resources (S_0) which they can allocate over a discrete number of time periods (T). The quantity of the resource extracted by player k in period t is denoted as q_k^t. Further, let the set of quantities player k allocates over the T periods be denoted by $g_k = (q_k^1, q_k^2, \ldots, q_k^T)$. We will refer to g_k as player k's trajectory in the remainder of the text.

Given two trajectories g_k, g_{k-1} of players k and $k-1$ and assuming a linear demand function, the profit for player k is defined as

$$\pi(g_k, g_{k-1}) = \sum_{t=1}^{T} \left(\beta^t \left(a - b\left(q_{k-1}^t + q_k^t\right)\right) q_k^t\right).$$

Here, a and b are parameters of the demand function and β is the discount factor. In line with the lab experiment, we will use symmetric firms, and take $S_0 = 170$, $T = 6$, $b = 1$, $a = 372$, and $\beta = \frac{1}{1+r} = \frac{1}{1.1}$.

Player k's problem is to choose a strategy that maximizes the sum of discounted profits, subject to the resource constraint:

$$\begin{aligned} \underset{q_k^1, q_k^2, \ldots, q_k^T \in [0, S_0]}{\text{maximize}} \quad & \pi(g_k, g_{k-1}) \\ \text{subject to} \quad & \sum_{t=1}^{T} q_k^t \leq S_0, \end{aligned} \tag{1}$$

Player $k-1$ solves an analogous problem. There are three relevant benchmarks to consider: the Nash equilibrium, collusion and Walras. For the Nash equilibrium, both players maximize their own profits given the strategy of the opponent. In the collusive case, the two players maximize joint profits, and in the Walras case, each player maximizes its profits while taking prices as given. Table 1 below plots the relevant trajectories; for a more detailed derivation see Van Veldhuizen and Sonnemans (2014).

Two results are worth highlighting. First, all three benchmarks have decreasing trajectories. This follows directly from the standard result that the shadow price of the resource increases over time at the rate of interest (Hotelling, 1931). Second, relative to Nash, the collusive quantities are smaller in early periods and larger in late periods; the converse is true for Walras, which is again a standard result (Solow, 1974).

3 Level-k Approach

We now move away from the symmetric benchmarks discussed in the previous section in order to derive the optimal strategies for players of different level k. As the first step, the level-k framework requires a level-0 strategy to be specified. A typical approach is to assume that a level-0 player uniformly randomizes across all possible actions. We take a similar approach by assuming that the level-0 player randomizes across all feasible trajectories (i.e., resource allocations over all periods).

	q^1	q^2	q^3	q^4	q^5	q^6
Nash	49.61	42.17	33.98	24.98	15.08	4.19
Walras	61.51	49.06	35.37	20.31	3.74	0.00
collusion	42.71	37.68	32.16	26.07	19.37	12.01
$k=0$ (average)	24.29	24.29	24.29	24.29	24.29	24.29
$k=1$	59.33	47.87	35.26	21.39	6.15	0.00
$k=2$	44.76	39.30	33.33	26.79	19.55	6.27
$k=3$	52.04	43.59	34.30	24.10	12.84	3.14
$k=4$	48.40	41.44	33.82	25.43	16.21	4.70
$k=10$	49.58	42.14	33.97	25.00	15.12	4.18

Table 1: The first three rows of give the Nash, Walras and collusive quantities for each period. The fourth row in the Table gives the *average* production quantity for a level-0 for each period. The remaining rows give the trajectories for different level-k players.

We assume that the resource constraint for level-0 is not required to hold with equality,[1] i.e., we only require that each trajectory $g_0 = (q_0^1, \dots, q_0^T)$ satisfies:

$$\sum_{t=1}^{T} q_0^t \leq S_0. \tag{2}$$

3.1 Level 0

We will start by considering the set of all possible trajectories of the level-0 player, denoted by G_0. For the purpose of this derivation and to facilitate a comparison with the experimental data, we consider the case where players can only choose discrete extraction quantities. We therefore discretize the interval of possible quantities $Q := [0, S_0]$ into $n_\alpha \in \mathbb{N} \setminus \{0, 1\}$ equidistant values with distance

$$\delta := \frac{S_0}{n_\alpha - 1}.$$

The discrete set of valid extraction quantities for a given period is then

$$\hat{Q} := \{\delta\alpha : \alpha \in \{0, 1, \dots, n_\alpha - 1\}\}. \tag{3}$$

We denote the corresponding set of possible trajectories for the level-0 player that consist of quantities in $\hat{Q}$ by $\hat{G}_0 \subseteq G_0$. Note that $\lim_{\delta \to 0} \hat{G}_0 = G_0$. In the experiment $\delta = 1$, and therefore $\hat{G}_0$ contains all integer extraction quantities in Q.

We denote the cardinality of $\hat{G}_0$ by $n_\ell := |\hat{G}_0|$. We can now number trajectories in $\hat{G}_0$:

$$\hat{G}_0 = \{g_{0,0}, g_{0,1}, \dots, g_{0,n_\ell - 1}\} \tag{4}$$

and add a corresponding index to the quantities defining each trajectory in $\hat{G}_0$:

$$g_{0,\ell} = (q_{0,\ell}^t)_{t=1,\dots,T} = \left(q_{0,\ell}^1, q_{0,\ell}^2, \dots, q_{0,\ell}^T\right) \tag{5}$$

[1]The results are very similar if the resource constraint is forced to hold with equality, results are available upon request.

In line with the literature (see e.g., Camerer et al., 2004, 1998; Nagel, 1995) we assume that the level-0 player's strategy is to pick one of the trajectories in $\hat{G}_0$ at random, where each trajectory is chosen with probability $1/n_\ell$.

3.2 Level 1

The next step is to derive the strategy of the level-1 player. The level-1 player's goal is to maximize the expected sum of his discounted profits π, conditional on his opponent playing the level-0 strategy. Player 1's objective Π_1 can be expressed as

$$\Pi_1 = \frac{1}{n_\ell} \sum_{\ell \in \hat{G}_0} \pi(g_1, g_{0,\ell}). \tag{6}$$

Then the optimization problem for level 1 reads as follows:

$$\begin{aligned} & \underset{q_1^1, q_1^2, \ldots, q_1^T \in [0, S_0]}{\text{maximize}} && \Pi_1, \\ & \text{subject to} && \sum_{t=1}^{T} q_1^t \leq S_0. \end{aligned} \tag{7}$$

Averaging Problem

Writing out Π_1 explicitly, using the definition of $\pi(g_1, g_0)$ given in Equation (2), and changing the order of the two sums (the averaging over strategies and the sum over time periods), we can write

$$\begin{aligned} \Pi_1 &= \frac{1}{n_\ell} \sum_{\ell=1}^{n_\ell} \sum_{t=1}^{T} \left(q_1^t \cdot \beta^t \cdot \left(a - b \left(q_1^t + q_{0,\ell}^t \right) \right) \right) \\ &= \sum_{t=1}^{T} (q_1^t \cdot \beta^t \cdot (a - b(q_1^t + \underbrace{\frac{1}{n_\ell} \sum_{\ell=1}^{n_\ell} q_{0,\ell}^t}_{q_{\text{eff}}^t}))). \end{aligned} \tag{8}$$

This allows us to accumulate the averaging process in an *effective quantity* q_{eff}^t that expresses the average production of the level-0 player in period t.

Since the level-0 player is assumed to choose among the trajectories $\hat{G}_0$ in a uniformly random fashion, the averaging process of all the strategies is symmetric with respect to the time index. It follows, that the value of q_{eff}^t is independent of the period:

$$q_{\text{eff}}^t = \frac{1}{n_\ell} \sum_{\ell=1}^{n_\ell} q_{0,\ell}^t =: q_{\text{eff}}, \quad \forall \ \ t \in \{1, \ldots, T\}. \tag{9}$$

Proposition 1. *The value of q_{eff} depends only on the number of periods and the total available resource per player:*

$$q_{eff} = \frac{S_0}{T+1}. \tag{10}$$

Proof. We define the set $\hat{G}_0^\alpha \subset \hat{G}_0$, that contains all trajectories with a total resource extraction of $\delta\alpha$: $\hat{G}_0^\alpha := \{g_{0,\ell} : \sum_t q_{0,\ell}^t = \delta\alpha\}$.

The cardinality of $\hat{G}_0^\alpha$ is denoted by $n_{\hat{G}_0^\alpha} := |\hat{G}_0^\alpha|$. Its value can be expressed using a binomial coefficient:

$$n_{\hat{G}_0^\alpha} = \binom{\alpha + T - 1}{T - 1}. \tag{11}$$

Using the combinatorial identity:

$$\sum_{l=0}^{m} \binom{l+n}{n} = \binom{n+m+1}{n+1}, \tag{12}$$

we can calculate the number of all possible strategies

$$n_\ell = \sum_{\alpha=0}^{n_\alpha - 1} n_{\hat{G}_0^\alpha} = \sum_{\alpha=0}^{n_\alpha - 1} \binom{\alpha + T - 1}{T - 1} = \binom{T + n_\alpha - 1}{T}. \tag{13}$$

We now rewrite q_{eff} in terms of the sum over all elements of the matrix $q_{0,\ell}^t$:

$$q_{\text{eff}} = \frac{1}{n_\ell}\frac{1}{T}\sum_{t=1}^{T}\sum_{\ell=1}^{n_\ell} q_{0,\ell}^t = \frac{1}{n_\ell}\frac{1}{T}\sum_{\alpha=0}^{n_\alpha - 1} n_{\hat{G}_0^\alpha}\cdot\delta\alpha = \frac{1}{n_\ell}\frac{1}{T}\sum_{\alpha=1}^{n_\alpha - 1} n_{\hat{G}_0^\alpha}\cdot\delta\alpha. \tag{14}$$

Using (11), we arrive at the following expression for q_{eff}:

$$\begin{aligned}
q_{\text{eff}} &= \frac{\delta}{n_\ell}\sum_{\alpha=1}^{n_\alpha - 1}\frac{\alpha}{T}\frac{(\alpha+T-1)!}{(T-1)!\alpha!} \\
&= \frac{\delta}{n_\ell}\sum_{\alpha=1}^{n_\alpha - 1}\frac{(\alpha+T-1)!}{(T)!(\alpha-1)!} \\
&= \frac{\delta}{n_\ell}\sum_{\alpha=0}^{n_\alpha - 2}\frac{(\alpha+T)!}{(T)!(\alpha)!} \\
&\overset{(12)}{=} \frac{\delta}{n_\ell}\binom{T+n_\alpha-1}{T+1} \\
&\overset{(13)}{=} \delta\frac{T!(n_\alpha-1)!}{(T+n_\alpha-1)!}\frac{(T+n_\alpha-1)!}{(T+1)!(n_\alpha-2)!} \\
&= \frac{\delta(n_\alpha-1)}{T+1} = \frac{S_0}{T+1}.
\end{aligned} \tag{15}$$

□

We emphasize that we obtained an expression for q_{eff} that does not depend on δ, the step size of the discretization. It follows, that we can now take the limit $\delta \to 0, n_\alpha \to \infty$ and recover the same result for the case of continuous allocation quantities ($\hat{G}_0 \equiv G_0$).

Solving the Level-1 Problem

In the previous section we have shown that the problem (7) that requires the evaluation of a potentially large sum can be reduced to a much simpler problem. In particular, we have shown that

$$\Pi_1 = \frac{1}{n_\ell} \sum_{\ell \in \hat{G}_0} \pi(g_1, g_{0,\ell}) = \pi\left(g_1, g_{0,\text{eff}}\right) \tag{16}$$

where

$$g_{0,\text{eff}} = (q_{\text{eff}}, q_{\text{eff}}, \ldots, q_{\text{eff}}). \tag{17}$$

Using this definition of Π_1, the problem (7) is computationally equivalent to the original problem (1) where the opponent is known to play a particular trajectory. We can now formulate this problem as a nonlinear program with one decision variable for each of the 6 periods, and find the globally best strategy by numerically solving the problem. We use the mixed-integer nonlinear programming (MINLP) solver SCIP (Achterberg, 2009) to solve the problem to global optimality.[2]

3.3 Solving for $k > 1$

We can now find solutions up to an arbitrary level k using an iterative process. For each level k, we solve (1) using the trajectory computed in the last level for the $k-1$ player. We implemented this as a shell script that solves each level with an individual call to our solver SCIP, using the results of the previous level as input.

4 Computational Results

Table 1 displays the computational results for the continuous case.[3] The level-0 player extracts on average a quantity of 24.29 in each period. Given that the level-0 player on average moves its production towards the later periods relative to Nash, the level-1 player responds by moving production to earlier periods. In fact, the level-1 strategy is much closer to Walras than to the Nash equilibrium. Faced with the high initial extraction of the level-1 player, the level-2 player then responds by moving production back to later periods, and so on. This alternating process quickly converges to the Nash equilibrium quantities; for example, for the level-10 player, the distance to Nash is already less than 0.04 for each period. At $k = 33$, the solution has converged with a tolerance of 10^{-4}, i.e.

$$|q_{32}^t - q_{33}^t| < 10^{-4}, \quad \forall \ \ t \in \{1, \ldots, T\}. \tag{18}$$

5 Experimental Data

The model studied in this paper was implemented in a laboratory experiment by Van Veldhuizen and Sonnemans (2014). In their experiment, participants went through 10 repetitions of the same 6 period set-up. We refer to their paper for more details on the experimental design and instructions. However, it is important to note that participants were rematched to a different opponent for every repetition. Their first period quantity could therefore not be based on any knowledge of the behavior of their opponent, and can hence serve as a good proxy for their level in the cognitive hierarchy.

We classify all participants in the experiment by the level that most closely corresponds to their first period choice of quantity, as per Table 1. We classify quantities that are either three units smaller than the collusive quantity or three units larger than the Walras quantity as level-0

[2] A phase space analysis of the level-1 problem can be found in Fügenschuh et al. (2013).
[3] The integer results are similar and available on request

players, because, within the level-k framework, these quantities can only be generated by pure randomization (level-0) and not by any higher level strategy.[4]

Table 2 below gives the results. Most notably, the number of participants classified as level-0 decreased drastically from round 1 (24 participants) to round 10 (5 participants). The number of participants classified as level 1, 2 and particularly level 3 increased correspondingly. This suggests that participants learned to make higher level decisions over successive repetitions of the experiment.

Quantity q_k^1	Type	Round 1	Round 10
< 40	level-0	9	3
$40 - 43$	collusive	11	6
$44 - 46$	level-2	7	9
$47 - 49$	level-4	1	3
50	Nash	11	12
$51 - 55$	level-3	2	13
$56 - 60$	level-1	7	8
$61 - 65$	Walras	1	8
> 65	level-0	15	2

Table 2: This table shows the number of participants who chose a particular extraction quantity in the first period of respectively the first and last (10th) repetition (or round) of the experiment. The corresponding levels follow from the analysis above.

6 Discussion

Existing models of non-renewable resources assume sophisticated agents compete with other sophisticated agents. This study instead uses a level-k approach to examine situations where the focal agent is uncertain about the strategy of his opponent, or predicts the opponent will act in a non-sophisticated manner. We modelled the uncertainty about the opponent's chosen strategy using a level-k framework.

Interestingly, the level-1 player's optimal strategy is quite close to the Walras benchmark. Intuitively, this player best responds to a random opponent, who on average under-extracts in early periods (relative to Nash). As a result, the level-1 player's best response is to instead over-extract in early periods. Similarly, the level-2 player best responds to level-1, and is therefore closer to the collusive quantity than to Nash. Thus, for producers who expect their opponents to be randomizers (or best respond to randomizers), it is not optimal to choose the Nash equilibrium strategy. Instead, it is optimal to choose a strategy that closely approximates respectively collusion or Walras. Only higher levels of rationality will more closely approximate the Nash equilibrium.

We then applied these computational results to data from a laboratory experiment, which allowed us to classify participants by the level of reasoning implied by their choices in the first period. For the early part of the experiment, many participants were classified as level-0. However, by the time they had garnered some experience in the game, their implied level of rationality increased. This seems intuitive and in line with the idea that repeated exposure to the same problem increases the quality of participants' responses.

Finally, the results also illustrate the applicability of numerical methods in solving iterated problems such as ours. For continuous cases, analytic methods are able to compute the Nash

[4] We use the three unit wiggle room in order not to immediately classify any deviation from Walras or collusion as level-0 behavior.

equilibrium either directly or as the limit of a level-k model where k converges to infinity. However we are not aware of analytical methods that are able to directly compute the strategy for a given finite value of k, especially when the set of possible production quantities is discrete, rather than continuous. By contrast, our paper illustrates that numerical solvers are able to provide the strategy for up to any level k. Numerical tools seem particularly well-suited in cases where the exact parameters of interest are known, such as the experimental data set analyzed in this paper.

Acknowledgement

We gratefully acknowledge the funding of this work by the German Research Association (Deutsche Forschungsgemeinschaft, DFG), Collaborative Research Center CRC1026 (Sonderforschungsbereich SFB1026). A. Fügenschuh conducted parts of this research under a Konrad-Zuse-Fellowship.

References

Achterberg T. 2009. SCIP: Solving constraint integer programs. *Mathematical Programming Computation*, 1(1):1–41.

Camerer CF, Ho TH, and Chong JK. 2004. A Cognitive Hierarchy Model of Games. *The Quarterly Journal of Economics*, 119(3):861–898. ISSN 0033-5533. doi:10.1162/0033553041502225.

Camerer CF, Ho TH, and Weigelt K. 1998. Iterated Dominance and Iterated Best Response in Experimental P-Beauty Contests. *American Economic Review*, 88(4):947–969.

Farrow S. 1985. Testing the efficiency of extraction from a stock resource. *Journal of Political Economy*, 93(3):452–87.

Fügenschuh A, van Veldhuizen R, and Vierhaus I. 2013. Production planning for non-cooperating companies with nonlinear optimization. In *11th Global Conference on Sustainable Manufacturing : Proceedings*, pages 536 – 541.

Groot F, Withagen C, and De Zeeuw A. 2003. Strong time-consistency in the cartel-versus-fringe model. *Journal of Economic Dynamics and Control*, 28(2):287–306.

Hotelling H. 1931. The economics of exhaustible resources. *Journal of Political Economy*, 39:137–75.

Loury GC. 1986. A theory of 'oil'igopoly: Cournot equilibrium in exhaustible resource markets with fixed supplies. *International Economic Review*, 27(2):285–301.

Nagel R. 1995. Unraveling in Guessing Games: an Experimental Study. *American Economic Review*, 85(5):1313–1326.

Newbery DM. 1981. Oil prices, cartels, and the problem of dynamic consistency. *Economic Journal*, 91:617–46.

Pindyck RS. 1980. Uncertainty and exhaustible resource markets. *Journal of Political Economy*, 88(6):1203–25.

Pindyck RS. 1981. Models of resource markets and the explanation of resource price behavior. *Energy Economics*, 3:130–9.

Polasky S. 1992. Do oil producers act as 'oil'igopolists? *Journal of Environmental Economics and Management*, 23(3):216–47.

Reynolds DB. 2013. Uncertainty in exhaustible natural resource economics: The irreversible sunk costs of Hotelling. *Resources Policy*, 38(4):532–541. ISSN 03014207. doi: 10.1016/j.resourpol.2013.09.002.

Runco M. 2013a. Bounded Rationality in a Cournot Duopoly Game. *Working paper*.

Runco M. 2013b. What Model Best Describes Initial Choices in a Cournot Duopoly Experiment? *Working Paper*.

Salo S and Tahvonen O. 2001. Oligopoly equilibria in nonrenewable resource markets. *Journal of Economic Dynamics and Control*, 25:671–702.

Solow R. 1974. The economics of resources or the resources of economics. *American Economic Review*, 64(2):1–14.

Spiro D. 2012. Resource prices and planning horizons. *Working Paper*.

Stahl DO and Wilson PW. 1994. Experimental evidence on players' models of other players. doi:10.1016/0167-2681(94)90103-1.

Van Veldhuizen R and Sonnemans J. 2014. Nonrenewable Resources, Strategic Behavior and the Hotelling Rule. *Working Paper*, pages 1–41.