# Depth of Reasoning and Information Revelation: An Experiment on the Distribution of $k$-Levels*

Volker Benndorf[†] Dorothea Kübler[‡] Hans-Theo Normann[§]

July 2017

**Abstract**

The level-$k$ model is a workhorse in behavioral game theory. For comparisons across experiments and predictions in future studies, it is crucial to assess the empirical distribution of $k$-levels. We present a revelation game suitable for this purpose. In a labor market context, workers can choose to reveal their productivity at a cost, and players' strategies reveal their level of reasoning in terms of a $k$-level. We find that the most frequently observed reasoning levels are $k = 2$ and $k = 3$. In our game roughly 30 percent of the players are $k \leq 1$ and 25 percent are $k \geq 4$. We compare our results to other experiments that identify level-$k$ distribution, foremost to the money request (or 11-20) game. Despite various differences to the 11-20 game, our revelation game suggests a very similar distribution of level-$k$ types.

JEL Classification numbers: C72, C90, C91

Keywords: level-$k$ reasoning, distribution of $k$-levels, unraveling, labor market, experiment

---

*We would like to thank three anonymous referees, an associate editor, and the editor for their helpful comments and suggestions. Dorothea Kübler gratefully acknowledges financial support through CRC TR 190.

[†]Goethe-Universität Frankfurt, Tel: +49 69 798 34806, Email: *benndorf@econ.uni-frankfurt.de*

[‡]WZB Social Science Center Berlin and Technical University Berlin, Tel: +49 30 25491 440, Email: *kuebler@wzb.eu.*

[§]Duesseldorf Institute for Competition Economics (DICE), Tel: +49 211 8115297, Email: *normann@dice.hhu.de.*

# 1 Introduction

The level-$k$ model introduced by Nagel (1995) and Stahl and Wilson (1995) is a workhorse in behavioral economics. In contrast to the standard model, level-$k$ players maintain wrong beliefs about other players. Specifically, a level-$k$ player believes all other players have a depth of reasoning of level-($k$-1) and she best-responds to that belief.[1] Players' beliefs may be heterogeneous, and this heterogeneity is what makes level-$k$ rationality attractive for behavioral economics. The level-$k$ model's original application was the behavior in experimental beauty contests,[2] but the model has also been applied to various other games.[3] The concept is generally well-suited for the analysis of experimental data.

For comparisons across experiments and for making predictions, it is crucial to assess the empirical distribution of $k$-levels. To date, this has mainly been done by observing an individual's choices in a number of different games and by assigning to her the $k$-level that best fits her choices (Costa-Gomes et al., 2001; Stahl and Wilson, 1995). Alternatively, additional data are used that hint at the reasoning process, for example, Mouse Lab data (Costa-Gomes et al., 2001), advice given to others (Burchardi and Penczynski, 2014), or the choice protocol approach (Agranov et al., 2015).

Arad and Rubinstein (2012) (henceforth called A&R) go one step further and

[1] The level-$k$ model is defined recursively and it requires an assumption for the behavior of level-0 players. The most common assumption is that level-0 players choose an arbitrary strategy, but there are also exceptions which for example assume that level-0 players choose a specific pure strategy (e.g., Arad and Rubinstein, 2012).

[2] See, for example, Bosch-Domènech et al. (2002), Costa-Gomes and Crawford (2006), Garza et al. (2012), Ho et al. (1998), Kocher and Sutter (2005), Nagel (1995), and Shapiro et al. (2014).

[3] This includes the centipede game (Kawagoe and Takizawa, 2012), auctions (Crawford and Iriberri, 2007a; Rasch et al., 2012), the "hide-and-seek" game (Crawford and Iriberri, 2007b) or contests (Bernard, 2010).

introduce the “11–20 money request game,” specifically designed to determine the precise individual level of reasoning.[4] Two players simultaneously request an amount of money (measured in integer Shekel) from the set $\{11, 12, ..., 20\}$. Each player receives the amount requested plus an additional 20 Shekels if and only if the own request is exactly one Shekel lower than the other player’s request. If level-0 players choose 20, any choice of $20 - x$ corresponds exactly to a $k$-level of $x$. A&R find that most participants exhibit levels of two or three; only roughly 20 percent of the players are $k \leq 1$, and the proportion of players with $k \geq 4$ is negligible in statistical terms.

A&R convincingly list several arguments in favor of their model. They write that (i) incentives in their game are not confounded by social preferences, (ii) there does not exist a pure-strategy Nash equilibrium, and (iii) there are no dominated strategies.

We present another experimental game suitable for the elicitation of $k$-levels which, however, violates the aforementioned arguments in favor of the money-request game. In contrast to A&R, rational play imposes an externality on others, so social preferences might have an impact. Moreover, there is a unique pure-strategy equilibrium, and at least one player has a dominated action in our game. We will investigate whether these properties have an effect on the inferred distribution of $k$-levels. While it has emerged from previous studies that the distribution of cognitive types depends on the exact game used and its specifications such as information about the opponent (see Agranov et al. (2012) for a recent study), it is exactly for this reason that it is important to understand the determinants and

[4]The money request game is also used by Lindner and Sutter (2013) who analyze the impact of time pressure on the depth of reasoning.

relevant contextual features in order to allow for educated out-of-game predictions.

When studying the unraveling of privacy (Benndorf et al., 2015), we noticed that the structure of the game we employed lends itself well to the measurement of the level of reasoning of subjects. In that paper, we propose the level-$k$ model as a prime candidate to explain the data. However, we are not able to report the actual distribution of $k$-levels. The novelty of the experiments in this paper is that we use the strategy method. When players state a complete strategy, this directly implies their $k$-levels. As a result, we use the same game as in Benndorf et al. (2015), except for the response mode which was spontaneous in the companion paper but is strategy elicitation in the experiments reported here.

We study a labor market where workers can choose to reveal information (their productivity) at a cost. In our main variant, only the worker with the highest productivity will reveal when $k = 1$; the worker with the second highest productivity will reveal when $k = 2$, and so on. Generally, lower-productivity workers will reveal only for higher $k$-levels. For $k \geq 5$ and in Nash equilibrium, there is complete unraveling of information.

Our results speak to the hypotheses put forward by A&R. On the one hand, we find evidence for social preferences affecting choices and their classification according to $k$-levels. About one in five participants in our data chooses not to reveal for any productivity levels. Within the level-$k$ model, this behavior would be classified as $k = 0$. However, this strategy is unlikely to be the result of random or uninformed behavior; instead it is consistent with other-regarding preferences including inequality aversion (Fehr and Schmidt, 1999) and surplus maximization (Charness and Rabin, 2002; Engelmann and Strobel, 2004). When we exclude these choices, our distribution of $k$-levels is virtually identical to the one found in A&R. This supports their claim that social preferences can interfere with the classification of choices. On the other hand, further properties that A&R mention as possible confounds for the identification of $k$-levels (unique pure-strategy equilibrium, dominated actions) can be dealt with in our setup.

# 2 Theory

## 2.1 The game

We consider a complete-information game with $n$ workers who have productivities $\theta_1 < \theta_2 < \cdots < \theta_n$. Workers simultaneously choose whether to *reveal* or to *conceal* their productivity. Let $I_i \in \{0, 1\}$ indicate whether worker $i$ has chosen to reveal her productivity, with $I_i = 0$ denoting revelation and $I_i = 1$ concealment. The cost of revelation is $c > 0$. Worker i's payoff is

$$\Pi_i = \begin{cases} \theta_i - c & \text{if } I_i = 1 \quad (\textit{reveal}) \\ \sum_{j=1}^{n}(1 - I_j)\theta_j / \sum_{j=1}^{n}(1 - I_j) & \text{if } I_i = 0 \quad (\textit{conceal}). \end{cases}$$

In words, if worker $i$ chooses to reveal, $i$ earns her productivity minus the revelation cost. If not, she receives the average productivity of all workers who have chosen not to reveal.

The Nash equilibria of this game depend on the productivities and the cost $c$. Multiple pure-strategy equilibria and mixed-strategy equilibria may exist. Bendorf et al. (2015) derive a simple and intuitive condition (necessary and sufficient) for a unique equilibrium in pure strategies. It holds in general that $I_1 = 0$ is a dominant action. For $c = 0$, all workers except for worker 1 will reveal.

## 2.2 Level-k analysis

Our experimental design is suitable to identify players' $k$-levels for two reasons. First, reveal decisions are generally monotonic in productivities, regardless of the $k$-level: given $k$, there will be a unique cutoff level of productivity for which workers switch from $I_i = 0$ to $I_i = 1$. Second, the parametrization of our labor markets ensures that different $k$-levels correspond to different switching points. Consider these two issues in turn.

First, the level-$k$ choices in this game follow a regular pattern in that they are monotonic in productivity. Consider two workers $i$ and $j$ with the same $k$-level ($k \geq 1$) but different productivities $\theta_j > \theta_i$ . We demonstrate that worker $i$ will not reveal provided worker $j$ finds it worthwhile to conceal, or vice versa, worker $j$ will not conceal when worker $i$ finds it worthwhile to reveal. Proposition 1 formalizes this:

**Proposition 1.** *Suppose workers reason at some common level* $k \geq 1$*. Then their choices satisfy* $1 \geq I_n^k \geq I_{n-1}^k \geq \cdots \geq I_2^k \geq I_1^k = 0$.

A proof can be found in the appendix. Since level-$k$ players by definition choose a best reply, the proof essentially shows the monotonicity of best responses for any $k$. If players make as-if decisions for all workers, they will switch (at most) once from concealment to revelation as the productivity worker they decide for increases. Proposition 1 implies that any monotone strategy can be associated with a k-level.

Second, with a set of appropriate parameters, these switching points may be exploited to distinguish between different $k$-levels where higher $k$-levels are associated with lower switching points. That means that more workers will switch to *reveal* as the $k$-level of the player increases. The process will stop (that is, choices of level-$k+1$ players are identical to choices of level-$k$ players) when a Nash equilibrium is reached.

### 2.3 Level-0 assumption

For many games, the level-$k$ predictions depends on the assumption for level-0 play. Our game is no exception but we can show that changing the level-0 assumption only shifts each level of reasoning by one.

Consider any symmetric mixed strategies as level-0 play. Let $p$ denote the probability that a level-0 player chooses *reveal*. Consider $p = 0$. Now, $k = 1$ workers will conceal if their productivity satisfies $\theta - c \leq \sum_{j=1}^{n} \theta_j / n$. In words, only workers whose reveal payoff is above the average productivity will reveal. Increasing $p \in (0, 1)$, the number of $k = 1$ players who conceal rises monotonically. If $p = 1$, level-1 players of any productivity level $\theta_i$ and for any parametrization will prefer to conceal.[5] Given that all $k = 1$ players choose to conceal when $p = 1$, the choices of $k = 2$ players when $p = 1$ are identical to those of $k = 1$ players when $p = 0$ is the level-0 assumption. In other words when moving from $p = 0$ to $p = 1$, we merely see a shift of the level-$k$ predictions but no qualitative differences.

For the rest of the paper, our level-0 assumption is $p = 0.5$. This assumption is commonly made in the literature and is consistent with random choices.

[5]The intuition is that, if player $i$ expects all other players $\neq i$ to reveal, she cannot benefit from revealing herself because $i$ is identified by very the fact that she is the only worker who conceals.

# 3 Experimental design and procedures

In the experiments, we use three different variants called *Market A*, *Market B*, and *Market C*. Each market represents different worker productivities but we set $n = 6$ and $c = 100$ throughout. The productivities and the unique Nash equilibrium for each market are summarized in Table 1. Table 1 also shows which level of reasoning is required for equilibrium play, and it displays the choices predicted for a given $k$-level and productivity.

| Market A | | | Market B | | | | | | | Market C | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Theta$ | $I^1$ | $I^{EQ}$ | $\Theta$ | $I^1$ | $I^2$ | $I^3$ | $I^4$ | $I^5$ | $I^{EQ}$ | $\Theta$ | $I^1$ | $I^2$ | $I^3$ | $I^{EQ}$ |
| 200 | 0 | 0 | 200 | 0 | 0 | 0 | 0 | 0 | 0 | 200 | 0 | 0 | 0 | 0 |
| 210 | 0 | 0 | 448 | 0 | 0 | 0 | 0 | 1 | 1 | 448 | 0 | 0 | 0 | 0 |
| 230 | 0 | 0 | 510 | 0 | 0 | 0 | 1 | 1 | 1 | 510 | 0 | 0 | 0 | 0 |
| 260 | 0 | 0 | 551 | 0 | 0 | 1 | 1 | 1 | 1 | 551 | 0 | 0 | 1 | 1 |
| 300 | 0 | 0 | 582 | 0 | 1 | 1 | 1 | 1 | 1 | 582 | 0 | 1 | 1 | 1 |
| 600 | 1 | 1 | 607 | 1 | 1 | 1 | 1 | 1 | 1 | 607 | 1 | 1 | 1 | 1 |

Table 1: Level-k and equilibrium strategies using the parameters from the experiment (the cost of revelation is $c = 100$ in all markets). The vector $I^k = \{I_1^k, ..., I_n^K\}$ refers to the strategy of a level-$k$ player given the assumption that level-0 players reveal with probability 0.5. $I^{EQ}$ refers to the equilibrium strategy.

The three markets are played in turn. Subjects begin with Market A, then turn to Market B in the second period, and to Market C in the third period before they start all over again with Market A in period four. In total, subjects play 15 periods, that is, five repetitions of each market. Note that we use a fixed-matching protocol in our experiments.

Table 1 indicates that the workers with the highest productivity reveal when $k \geq 1$. Workers who conceal in equilibrium play their equilibrium action for any $k \geq 1$. By contrast, revealing may require higher levels of reasoning.

Our main variant is Market B because it is suitable for eliciting $k$-levels up to $k = 5$. Markets A and C serve as control treatments. The primary purpose of these markets is to check whether the distribution of $k$-levels is constant across different markets.

Subjects were inexperienced with the game that we implemented. In order to determine the subjects' $k$-levels, we use the strategy-elicitation method and have our participants make as-if decisions for each of the six workers. If we had asked subjects for an actual decision for just one specific worker (as was the case in the experiments for Benndorf et al., 2015), the data would be less conclusive regarding the exact $k$-level. Consider as an example worker 4 in Market B: if she conceals, this implies $k < 3$; if she reveals, we learn $k \geq 3$, but an exact $k$-level cannot be concluded. Using the strategy method, the subject's switching point yields an exact $k$-level (up to $k = 5$): if, in Market B, a subject conceals as worker 1 to 3 and reveals as worker 4 to 6, we know that this participant reasons at level $k = 3$. Thus, the strategy method is a parsimonious way to elicit levels of reasoning, as an alternative to the more common approach where subjects take multiple decisions in different roles. Once all subjects have taken their decisions, a random computer draw determines which subject will act in the role of which worker. Then the six decisions are matched.

At the end of the period, subjects are presented with a summary of the results. The feedback contains information about the realization of the random draw, the corresponding decision of the subject, the payment to subjects who have not revealed their productivity, and the payment to the subject. The feedback does not contain specific information on the decisions of the other participants.

In total, 66 subjects participated in the experiment. This results in 11 independent observations when counting one group of six participants as one independent observation. The experiment was conducted at the TU-WZB lab in Berlin in January 2012. A session lasted approximately 90 minutes and subjects earned 10.99 EUR on average. The experiments were conducted using Fischbacher's (2007) z-Tree software. The ORSEE tool (Greiner, 2015) was used for online recruitment.

# 4 Results

## 4.1 Choices

Figure 1 displays the aggregate distribution of choices in Market B. Among the $2^6 = 64$ possible strategies, six strategies are chosen frequently: there are five monotone strategies (with a unique switching point from conceal to reveal for some higher productivity) which are consistent with level-$k$ thinking. Another frequent strategy is to conceal for all productivities, labeled "Conceal" in Figure 1. The remaining 58 strategies are rarely chosen. They include all non-monotone strategies. We interpret these as level-0 behavior as they can be considered as the result of players picking arbitrary strategies.[6] Players may also reveal too often for any $k > 0$, even though they submit a monotone strategy. These strategies are also counted as level-0 because these participants reveal for productivity levels where they should not reveal in equilibrium. Figure 1 displays the frequency of choices across all periods labeled level $k = 0$ to $k = 5$ and Conceal, accordingly.

[Figure 1 about here.]

[6]Non-monotone strategies are never a best response in this game. From the level-k perspective, level-0 play is the only explanation for this kind of behavior.

Consider the strategy where subjects never reveal, regardless of their productivity and denote it by "Conceal". From a level-$k$ perspective, such strategies should be considered as level-0 play. However, the Conceal strategy constitutes a sizable share of the choices (more than 20%), and it appears unlikely that so many choices are the result of randomization (which would suggest $0.5^6 \approx 1.6\%$ Conceal choices). The reasons for the popularity of this particular strategy are likely to be found outside the level-$k$ model: the Conceal strategy maximizes joint payoffs and leads to the only outcome where all players earn the same payoffs. Moreover, a calibrated Fehr-Schmidt model reveals that it is a Fehr-Schmidt equilibrium in about 56% of all cases (see Benndorf et al., 2015 for more details on inequality aversion in this context). Also, privacy concerns can lead people to choose Conceal.[7] It seems inappropriate to categorize this strategy as level-0. One advantage of our game is that this strategy, motivated by social preferences and/or privacy concerns, is easily identifiable and we can therefore consider it separately from the level-$k$ analysis.

## 4.2 Identification of levels of reasoning

In order to compare our level-$k$ results to A&R and other games, we need to make a minor adjustment regarding the data. The reason is that we assume that level-0 players randomize over their entire action set whereas A&R assume level-0 players choose 20 in their game.[8] It could accidentally be the case that our level-0 players

[7]Regarding privacy concerns, Benndorf et al. (2015) point out that such concerns may prevent subjects from revealing their productivity, as suggested by a comparison to an additional neutrally framed treatment they employ.

[8]If level-0 players are assumed to pick a pure strategy, such adjustment is unnecessary, and the distribution of $k$-levels can directly be learned from the distribution of choices as presented in Figure 1. However, such a level-0 assumption comes with a significant downside. If level-0

pick one of the five strategies that are associated with higher $k$-levels. There are 64 strategies, 58 of which are level-0 play. Since level-0 players are assumed to randomize over all strategies, the share of level-0 decisions should equal $\frac{64}{58}$ times the density of choices that directly qualify as level-0. The density of the other choices can then be derived by distributing the remaining mass according to the frequencies observed.

Table 2 summarizes the data for Market B (all periods). The column "Density unadjusted" refers to the choice data and is illustrated by Figure 1. "Density adjusted" shows the numbers after performing the aforementioned adjustment procedure. Finally, the column "Density without Conceal" normalizes the sum of level-$k$ choices to 1 when we drop the Conceal choices from the data.

| | | Density | | |
|---|---|---|---|---|
| | Freq. choices | unadjusted | adjusted | without Conceal |
| $k = 0$ | 42 | 0.127 | 0.140 | 0.178 |
| $k = 1$ | 34 | 0.103 | 0.101 | 0.128 |
| $k = 2$ | 64 | 0.194 | 0.191 | 0.241 |
| $k = 3$ | 55 | 0.167 | 0.164 | 0.208 |
| $k = 4$ | 47 | 0.142 | 0.140 | 0.177 |
| $k \geq 5$ | 18 | 0.055 | 0.054 | 0.068 |
| Conceal | 70 | 0.212 | 0.209 | – |
| Sum | 330 | 1.000 | 1.000 | 1.000 |

Table 2: The observed level-$k$ distribution of unadjusted and adjusted choices, as well without Conceal choices.

Figure 2 illustrates our main result. It reports the cumulative density function of the $k$-levels observed in A&R's "Basic" treatment and our Market B. We display the classification for our game in the first and last period separately in order to

players were, for example, assumed to conceal with probability one, the decisions that are labeled "Conceal" in the figure would qualify as level-0 while the decisions that are labeled with "k=0" could not be explained using the level-$k$ model.

check for potential learning effects. In the left panel, we exclude the Conceal choices and find that there are virtually no differences between A&R's and our two distributions. A minor discrepancy is that we have slightly more probability mass on $k = 0$, and there is more level-4 play in our game. Kolmogorov-Smirnov tests (one for each period and one for all periods at once) reveal, however, that we cannot reject the null hypothesis that the samples from Market B and A&R's Basic are drawn from the same distribution at any conventional significance level (all $D < 0.181$ whereas the threshold for $p = 0.1$ would be $D = 0.202$). The right panel of the figure includes the Conceal decisions as level-0 play. It emerges that the distributions from the two experiments now differ considerably. This can be interpreted as evidence that social preferences can confound the distribution of k-levels, as suggested by A&R.

[Figure 2 about here.]

A second result is that there is very little learning in our game. The choices in period one and period five of Market B do not differ much. To formally test this, we conduct a Wilcoxon signed-rank test using group mean and median $k$-levels. We find that neither test is significant ($p = 0.464$ and $p = 0.344$, two-tailed, for mean and median, respectively). This result is consistent with repeated versions of the money request game where not much learning is observed either (see Lindner and Sutter, 2013).

### 4.3 Level-$k$ distributions in other experiments

Our level-$k$ distribution can also be compared to the evidence from other games. In $3\times3$ normal-form games, Stahl and Wilson (1994) identify about one quarter of players as level-1, about two quarters as level-2, and about one quarter as naive Nash players, a result that finds support in Costa-Gomes, Crawford, and Broseta (2001). In our game, we also find more level-2 than level-1 play, at 24% and 13% respectively. However, in our game and in A&R, the proportion of play consistent with $k > 2$ is larger than in the 3×3 normal-form games.

In contrast to 3×3 normal-form games, the money request game, and our revelation game, the proportion of level-1 play is higher relative to level-2 play in guessing games. Costa-Gomes and Crawford (2006) find that almost half of the players are level-1 in these games, slightly more than a quarter are level-2 and slightly less than a quarter are naive Nash while the number of level-3 players and higher is very small.

As already pointed out by A&R, the money request game induces higher levels of reasoning than 3×3 normal-form and guessing games. Our revelation game shares this property with the money request game.

## 4.4 Markets A and C and A&R's other variants

Markets A and C serve as control treatments for Market B. Will behavior change when fewer steps of reasoning are required for equilibrium play? The data from our controls are generally consistent with the data from our main treatment. For instance, the share of subjects reasoning at a level $k \geq 1$ is remarkably stable across markets. We find no significant differences in the share of subjects reasoning at $k \geq 1$ between the first period of Market B and Market A or Market C (Friedman test, $p = 0.529$). Market C has a higher share of level-2 players than Market B, though, also compared to A&R.

| | Arad and Rubinstein | | | | This study | |
|---|---|---|---|---|---|---|
| Level | Basic | Cycle | Costless | Market A | Market B | Market C |
| $k = 0$ | 0.065 | 0.125 | 0.151 | 0.185 | 0.178 | 0.200 |
| $k = 1$ | 0.120 | 0.472 | 0.396 | 0.815* | 0.128 | 0.160 |
| $k = 2$ | 0.296 | 0.222 | 0.208 | | 0.241 | 0.477 |
| $k = 3$ | 0.324 | 0.097 | 0.094 | | 0.208 | 0.164* |
| $k = 4$ | 0.065 | 0.042 | 0.038 | | 0.177 | |
| $k \geq 5$ | 0.130* | 0.042* | 0.113* | | 0.068* | |
| $k \geq 1$ | 0.935 | 0.875 | 0.849 | 0.815 | 0.822 | 0.800 |

Table 3: Comparisons of $k$-levels elicited in this study (without choices of the Conceal strategy) and in Arad and Rubinstein (2012). Entries marked with * also comprise the choices of all higher levels and of Nash types.

[Figure 3 about here.]

A&R also conduct other variants of the money request game as robustness checks. The "Cycle" variant is identical to "Basic" (described in our introduction) except that a player choosing 20 will receive the bonus if the other player chooses 11. The difference between "Costless" and "Basic" is that all players who do not choose 20 will receive a risk-free payoff of 17 instead of the number they have chosen. Players will still receive the bonus if they choose a number that is one less than the number chosen by the other player. Compared to "Basic," "Cycle" and "Costless" have substantial mass on level-1 play, as detailed in Table 3. Figure 3 illustrates that the discrepancies between A&R's three variants are at least as substantial as the difference to our Market B, even if we include the Conceal players.

# 5 Conclusion

Our experiments, like those of Arad and Rubinstein (2012), aim at deriving a distribution of the players' depths of reasoning, based on the level-$k$ model introduced by Nagel (1995) and Stahl and Wilson (1995). Such distributions are essential for making predictions in other games.

Even though our game differs in many respects from the money request game of Arad and Rubinstein (2012), we observe almost the same distribution of $k$-levels once we isolate the choices where players conceal for any productivity. Rather than categorizing these choices as $k = 0$, we argue that such behavior results from other-regarding preferences or privacy concerns. The fact that the distribution of $k$-levels in our main market variant is practically identical to the distribution in Arad and Rubinstein (2012), may indicate some degree of irrelevance of a number of features distinguishing the two games. First of all, since the Nash equilibrium strategy only coincides with the strategy indicating $k = 5$, which is rarely observed, its confounding effect is relatively unimportant in our game. Furthermore, neither the existence of a pure-strategy equilibrium nor of players with dominant strategies impact the distribution of cognitive types. On the other hand, the distributions of $k$-levels in our revelation game and in the money request game are different from those observed in $3\times3$ normal-form games and guessing games. A systematic investigation of what determines the distribution of $k$-levels across games seems a worthwhile future endeavour.

# References

Agranov, Marina, Elizabeth Potamites, Andrew Schotter, and Chloe Tergiman (2012). "Beliefs and endogenous cognitive levels: An experimental study". In: *Games and Economic Behavior* 75.2, pp. 449–463.

Agranov, Marina, Andrew Caplin, and Chloe Tergiman (2015). "Naive play and the process of choice in guessing games". In: *Journal of the Economic Science Association*, pp. 1–12.

Arad, Ayala and Ariel Rubinstein (2012). "The 11-20 money request game: A level-k reasoning study". In: *American Economic Review* 102.7, pp. 3561–3573.

Benndorf, Volker, Dorothea Kübler, and Hans-Theo Normann (2015). "Privacy Concerns, Voluntary Disclosure of Information, and Unraveling: An Experiment". In: *European Economic Review* 75, pp. 43–59.

Bernard, Mark (2010). "Level-k reasoning in contests". In: *Economics Letters* 108.2, pp. 149–152.

Bosch-Domènech, Antoni, José G. Montalvo, Rosemarie Nagel, and Albert Satorra (2002). "One, two, (three), infinity, ... : Newspaper and lab beauty-contest experiments". In: *American Economic Review* 92.5, pp. 1687–1701.

Burchardi, Konrad B. and Stefan P. Penczynski (2014). "Out of your mind: Eliciting individual reasoning in one shot games". In: *Games and Economic Behavior* 84, pp. 39–57.

Charness, Gary and Matthew Rabin (2002). "Understanding Social Preferences with Simple Tests". In: *Quarterly journal of Economics* 117.3, pp. 817–869.

Costa-Gomes, Miguel A. and Vincent P. Crawford (2006). “Cognition and behavior in two-person guessing games: An experimental study”. In: *American Economic Review* 96.5, pp. 1737–1768.

Costa-Gomes, Miguel A., Vincent P. Crawford, and Bruno Broseta (2001). “Cognition and Behavior in Normal-Form Games: An Experimental Study”. In: *Econometrica* 69.5, pp. 1193–1235.

Crawford, Vincent P. and Nagore Iriberri (2007a). “Fatal attraction: Salience, Naïveté, and sophistication in experimental ‘hide-and-seek’ games”. In: *American Economic Review* 97.5, pp. 1731–1750.

— (2007b). “Level-k Auctions: Can a nonequilibrium model of strategic thinking explain the winner’s curse and overbidding in private-value auctions?” In: *Econometrica* 75.6, pp. 1721–1770.

Engelmann, Dirk and Martin Strobel (2004). “Inequality aversion, efficiency, and maximin preferences in simple distribution experiments”. In: *American Economic Review* 94.4, pp. 857–869.

Fehr, Ernst and Klaus M. Schmidt (1999). “A Theory of Fairness, Competition, and Cooperation”. In: *Quarterly Journal of Economics* 114.3, pp. 817–868.

Fischbacher, Urs (2007). “Z-Tree: Zurich toolbox for ready-made economic experiments”. In: *Experimental Economics* 10.2, pp. 171–178.

Garza, Pablo Brañas, Teresa García-Muñoz, and Roberto Hernán González (2012). “Cognitive effort in the Beauty Contest Game”. In: *Journal of Economic Behavior & Organization* 83.2, pp. 254–260.

Greiner, Ben (2015). “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE”. In: *Journal of the Economic Science Association* 1.1, pp. 114–125.

Ho, Teck-Hua, Colin F. Camerer, and Keith Weigelt (1998). “Iterated Dominance and Iterated Best Response in Experimental ‘p-Beauty Contests’”. In: *American Economic Review* 88.4, pp. 947–969.

Kawagoe, Toshiji and Hirokazu Takizawa (2012). “Level-k analysis of experimental centipede games”. In: *Journal of Economic Behavior & Organization* 82.2-3, pp. 548–566.

Kocher, Martin G. and Matthias Sutter (2005). “The decision maker matters: Individual versus group behaviour in experimental beauty-contest games”. In: *Economic Journal* 115.500, pp. 200–223.

Lindner, Florian and Matthias Sutter (2013). “Level-k reasoning and time pressure in the 11-20 money request game”. In: *Economics Letters* 120.3, pp. 542–545.

Nagel, Rosemarie (1995). “Unraveling in Guessing Games: An Experimental Study”. In: *American Economic Review* 85.5, pp. 1313–26.

Rasch, Alexander, Achim Wambach, Peter Werner, and Philippe Gillen (2012). “Level-k Thinking in a Multi-Unit Dutch Auction—An Experimental Investigation”.

Shapiro, Dmitry, Xianwen Shi, and Artie Zillante (2014). “Level-k reasoning in a generalized beauty contest”. In: *Games and Economic Behavior* 86, pp. 308–329.

Stahl, Dale O and Paul W Wilson (1995). “On Players’ Models of Other Players: Theory and Experimental Evidence”. In: *Games and Economic Behavior* 10.1, pp. 218–254.

## Acknowledgments

We are grateful to Nina Bonge who programmed and helped conduct the experiments as well as to three anonymous referees and the editor. Financial support by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR190 is gratefully acknowledged.

# Appendix

# Proof of Proposition 1

We first establish in two lemmas stating that, given any pure (Lemma 1) or symmetric mixed (Lemma 2) strategy profile, the best replies against that strategy profile are monotonic in worker productivity. Let $\xi = (I_1, I_2, ..., I_{n-1}, I_n)$ denote a pure-strategy profile and let $I_i^{br}(\xi)$ indicate worker $i$'s best reply against $\xi$. We have

**Lemma 1.** *Given any pure-strategy profile $\xi$, we have $1 = I_n^{br}(\xi) \geq I_{n-1}^{br}(\xi) \geq \cdots \geq I_2^{br}(\xi) \geq I_1^{br}(\xi) = 0$.*

*Proof.* Upfront, note $I_1^{br}(\xi) = 0$ since $I_1 = 0$ is a strictly dominant strategy: by concealing, worker 1 earns at least $\theta_1$ whereas, by revealing, she earns $\theta_1 - c < \theta_1$. Hence, concealing is strictly dominant and $I_1^{br}(\xi) = 0$. Further, we must have $I_n^{br}(\xi) = 1$ because otherwise we get the trivial case where all players conceal and level-$k$ analysis is bland.

We now prove, for some arbitrary profile $\xi'$, that $I_i^{br}(\xi') = 0 \wedge I_j^{br}(\xi') = 1$ only if $\theta_i \leq \theta_j$. In profile $\xi'$, workers $i$ and $j$ either conceal or reveal, implying four cases, namely (i) $I'_i = I'_j = 1$ (both reveal), (ii) $I'_i = I'_j = 0$ (both conceal), (iii) $I'_i = 0$ and $I'_j = 1$ ($i$ conceals, $j$ reveals), and (iv) $I_i = 1'$ and $I'_j = 0$ ($i$ reveals and $j$ conceals). Accordingly, we prove $I_i^{br}(\xi') = 0 \wedge I_j^{br}(\xi') = 1$ only if $\theta_i \leq \theta_j$ for these four cases. In terms of notation, let $I' = \sum_{l \neq i,j}(1 - I_l)$ be the number of concealing players in $\xi'$ other than $i$ and $j$, and let $\theta' = \sum_{l \neq i,j}(1 - I_l)\theta_l$ be the sum of concealed productivities of players other than $i$ and $j$. Consider the four cases in turn.

*Case (i).* Assume some $\xi'$ with $I'_i = I'_j = 1$. We obtain $I_i^{br}(\xi') = 0$ and $I_j^{br}(\xi') = 1$ if and only if $\theta_i - c \leq (\theta_i + \theta')/(1+I')$ and $\theta_j - c \geq (\theta_j + \theta')/(1+I')$. Solving both equations with respect to $\theta'$ yields $(1+I')(\theta_i - c) - \theta_i \leq \theta' \leq (1+I')(\theta_j - c) - \theta_j$ which reduces to $\theta_i \leq \theta_j$. *Case (ii).* Take some $\xi'$ with $I'_i = I'_j = 0$. Here, $I_i^{br}(\xi') = 0$ and $I_j^{br}(\xi') = 1$ if and only if $\theta_i - c \leq (\theta_i + \theta_j + \theta')/(2+I') \leq \theta_j - c$ which reduces to $\theta_i \leq \theta_j$. *Case (iii).* Consider $I'_i = 0$ and $I'_j = 1$. In this case, $I_i^{br}(\xi') = 0$ and $I_j^{br}(\xi') = 1$ if and only if $\theta_i - c \leq (\theta_i + \theta')/(1+I')$ and $\theta_j - c \geq (\theta_i + \theta_j + \theta')/(2+I')$. Solving both equations with respect to $\theta'$ yields $-c + (\theta_i - c)I' \leq \theta' \leq \theta_j - \theta_i - 2c + (\theta_j - c)I'$. Simplify this to $0 \leq -c + (\theta_j - \theta_i)(1+I')$ which holds only if $\theta_i \leq \theta_j$. *Case (iv).* Assume $I_i = 1'$ and $I'_j = 0$. We obtain $I_i^{br}(\xi') = 0$ and $I_j^{br}(\xi') = 1$ if and only if $\theta_i - c \leq (\theta_i + \theta_j + \theta')/(2+I')$ and $\theta_j - c \geq (\theta_j + \theta')/(1+I')$. Solving both equations with respect to $\theta'$ yields $(\theta_i - c)(1+I') - c \leq \theta' \leq (\theta_j - c)(1+I')$ which reduces to $0 \leq -c + (\theta_j - \theta_i)(1+I')$ which, in turn, holds only for $\theta_i \leq \theta_j$.

Since $I_i^{br}(\xi') = 0 < 1 = I_j^{br}(\xi')$ only if $\theta_i \leq \theta_j$ and since $\theta_i \neq \theta_j, i \neq j$, we cannot have $I_{i+1}^{br}(\xi') < I_i^{br}(\xi')$. Thus $1 \geq I_n^{br}(\xi') \geq I_{n-1}^{br}(\xi') \geq \cdots \geq I_2^{br}(\xi') \geq I_1^{br}(\xi') = 0$ as claimed. □

Consider now a symmetric mixed-strategy profile $\xi = (p, p, ..., p, p)$. Here, $p$ denotes the probability of revelation.

**Lemma 2.** *Given any symmetric mixed-strategy profile $\xi$, we have $1 \geq I_n^{br}(\xi) \geq I_{n-1}^{br}(\xi) \geq \cdots \geq I_2^{br}(\xi) \geq I_1^{br}(\xi) = 0$.*

*Proof.* For the purpose of this proof, let $\Phi_i$ denote worker $i$'s expected conceal payoff. For worker $i$ to (weakly) prefer conceal and worker $j$ to (weakly) prefer reveal, we need $\theta_i - c \leq \Phi_i$ and $\theta_j - c \geq \Phi_j$. To prove the lemma, we thus need to show $\theta_i - \theta_j \leq \Phi_i - \Phi_j$.

When computing $i$ and $j$'s conceal payoffs $\Phi_i$ and $\Phi_j$, we need to consider all possible pure-strategy outcomes that may occur as a result of the randomization of the $n-2$ players $\neq i, j$. This set of possible pure-strategy outcomes is $(I_l | l \neq i, j) \in \{0,1\}^{n-2}$ and the cardinality of this set is $2^{n-2}$. Let $m$ denote a typical element (that is, one of the $2^{n-2}$ pure-strategy outcomes). Somewhat abusing notation, let $I_{l,m}$ indicate player $l$'s, $l \neq i, j$, realized action in pure-strategy outcome $m$. Further, we use $I'_m := \sum_{l \neq i,j}(1 - I_{l,m})$ as the number of the $n-2$ other players who conceal in strategy combination $m$, and $\theta'_m := \sum_{l \neq i,j}(1 - I_{l,m})\theta_l$ denotes the sum of the productivities of the other $n-2$ players who conceal in pure-strategy outcome $m$. Finally, the probability that pure-strategy outcome $m$ is chosen is $p'_m = \prod_{l \neq i,j} p^{I_{l,m}}(1-p)^{1-I_{l,m}}$.

Now, the expected utilities of player $i$ and $j$ from concealing are:

$$\Phi_i = \sum_{m=1}^{2^{n-2}} \left[ p \frac{\theta_i + \theta'_m}{1 + I'_m} + (1-p) \frac{\theta_i + \theta_j + \theta'_m}{2 + I'_m} \right] p'_m$$

$$\Phi_j = \sum_{m=1}^{2^{n-2}} \left[ p \frac{\theta_j + \theta'_m}{1 + I'_m} + (1-p) \frac{\theta_j + \theta_i + \theta'_m}{2 + I'_m} \right] p'_m$$

We obtain

$$\Phi_i - \Phi_j = (\theta_i - \theta_j) \sum_{m=1}^{2^{n-2}} \frac{p}{1 + I'_m} p'_m$$

Since the terms in the sum add up to strictly less than 1 due to $\sum_{m=1}^{2^{n-2}} p'_m = 1$, the condition $\theta_i - \theta_j \leq \Phi_i - \Phi_j$ is satisfied only if $\theta_i \leq \theta_j$. □

We are now ready to complete the proof.

**Proposition 1.** *Suppose workers reason at some common level* $k \geq 1$*. Then their choices satisfy* $1 = I_n^k \geq I_{n-1}^k \geq \cdots \geq I_2^k \geq I_1^k = 0$.

*Proof.* Lemmas 1 and 2 establish that, for any pure or symmetric mixed strategy profile, best responses are monotonic in productivities. Since level-$k$ players always choose a best reply for any $k \geq 1$, this implies monotonicity in worker productivity for all level-$k$ decisions. □

# Instructions

Welcome to this experiment on economic decision making. Please read these instructions carefully. This experiment is conducted anonymously, that is, you will not get to know with which of the other participants you interacted or which participant acted in which role. Please note that you should not talk to the other participants during the whole experiment. If you have any questions please raise your hand and we will come to your place.

In this experiment all participants act as different workers. The workers in this experiment are different with respect to their health condition. The health condition of a worker determines her productivity and therefore the revenue of a fictional employer (played by the computer). Furthermore there are three different labor markets which are played on a rotating basis: Labor Market A, Labor Market B and Labor Market C. At the beginning of each period you will see the market played in that period on your monitor. On each labor market are six different workers with different health conditions.

| | Labor market A | Labor market B | Labor market C |
|---|---|---|---|
| Worker 1 | 200 | 200 | 200 |
| Worker 2 | 210 | 448 | 280 |
| Worker 3 | 230 | 510 | 360 |
| Worker 4 | 260 | 551 | 440 |
| Worker 5 | 300 | 582 | 520 |
| Worker 6 | 600 | 607 | 600 |
| Average | 300 | 483 | 400 |

Table 4: Health conditions of the workers 1-6 in the different labor markets.

The table above summarizes the different workers and their health conditions. Suppose market B is played in this period. If a fictional employer (played by the computer) hires for example worker 3, this worker will generate a revenue of 510 ECU for the employer. Worker 1 will only generate revenues of 200 ECU due to his poorer health. The same workers generate in a period, in which market C is played, the revenues of 200 (worker 1) or 360 ECU (worker 3). The health of any worker is of course completely fictional and randomly determined by the computer.

The experiment extends over 15 periods. At the beginning of each period the current market will be displayed on your monitor. In each period you need to make a decision for all six workers. Once all participants have made their six decisions, it will be randomly determined which of these six decisions is relevant for your payment in this period. That means, at the beginning of the experiment you will be randomly sorted in groups of six participants. Once all members of a group have made their six decisions the computer will randomly determine which group member represent which worker in that period. This will be the worker whose payoff you receive in that period. Even though you decide for all six workers, at the end of the period you will only receive a wage payment of a randomly chosen worker. This random draw will be such that there is always exactly one worker 1, one worker 2, one worker 3 and so on in each group. In other words, in each group there is always exactly one worker of either health condition. As mentioned before, at the beginning of each period you will get to know from your monitor, which of the three labor markets (A, B or C) is played in this period.

Your task in the experiment: In each period all participants need to make the following decision for every worker. You can choose whether the worker should buy a health certificate for a fee of 100 ECU. The health certificate reveals the worker's health condition and affects her payment in that period as follows:

1. If a worker purchases a health certificate, her payment will correspond to her health condition minus the fee of 100 ECU.

2. If a worker does not purchase a health certificate, her payment will correspond to the average health condition of all workers who do not have a health certificate.

All participants make their decisions independently. They do not know whether or not the other participants purchase any certificates. Moreover, at the time of your decision you will unaware of the resulting market wage. You will not get this information until the very end of the period. Once all workers have reached their decisions you will get detailed information about the result of this period on your monitor. The next period begins as soon as all participants have read the summary and clicked on "continue" Here is an example of the decision screen for Market A:

[Figure 4 about here.]

An example: Suppose labor market B is played in this period. At the beginning, when no worker has revealed his health yet, the average health of all workers without health certificate is:

$$\frac{200 + 448 + 510 + 551 + 582 + 607}{6} = \frac{2898}{6} = 483$$

The market wage equals 483 ECU in this case. Now each participant decides in the role of the single workers, whether he wants to reveal the respective health or not. In the table above you can also see the average health conditions for the markets A and C. Once all participants have made their decision, all will receive detailed information on the result.

Assume that the participants in the role of worker 3 and 5 in this period have decided to reveal the health conditions of those workers. In this case worker 3 receives a wage payment of 510 ECU and worker 5 receives a wage payment of 582 ECU. Both have revealed their health, hence both have to pay the fee of 100 ECU. Thus, the participant in the role of worker 3 earns 510 ECU - 100 ECU= 410 ECU and the participant in the role of worker 5 earns 482 ECU. Suppose the other participants have decided, that the remaining workers (1, 2, 4 and 6) shall not get a health certificate. In this case no fee has to be paid and the workers receive the market wage as wage payment. In this case the average health of all workers without health certificate will be: $\frac{200+448+551+607}{4} = \frac{1806}{4} = 451.5$ ECU. This is the market wage the participants in the role of the workers 1, 2, 4 and 6 will receive. Note that it is not possible that you have to more than one fee per period. For instance, if you decide that four different workers should get a health certificate in a period, there will still only be one worker that will determine your payoff and you have to pay for one health certificate if any. No fees will be charged for the other three health certificates since the corresponding decisions will remain payoff-irrelevant in that period. Even though you have to make a decision for all six workers, you will only be paid for one decision. However, you will not get to know the payoff relevant worker and the market wage until you have made all your decision.

As already mentioned, the experiment will take 15 periods in total. At the end, your earnings will be converted into Euro at a rate of: 500 ECU = 1 Euro. Furthermore we will round up the payoffs to the next 50-cent-amount. Please wait inside your cubicle until we call you for collecting your payment. After the experiment, please bring also all the documents you received from us. If you have any further questions, please raise your hand now!

# List of Figures

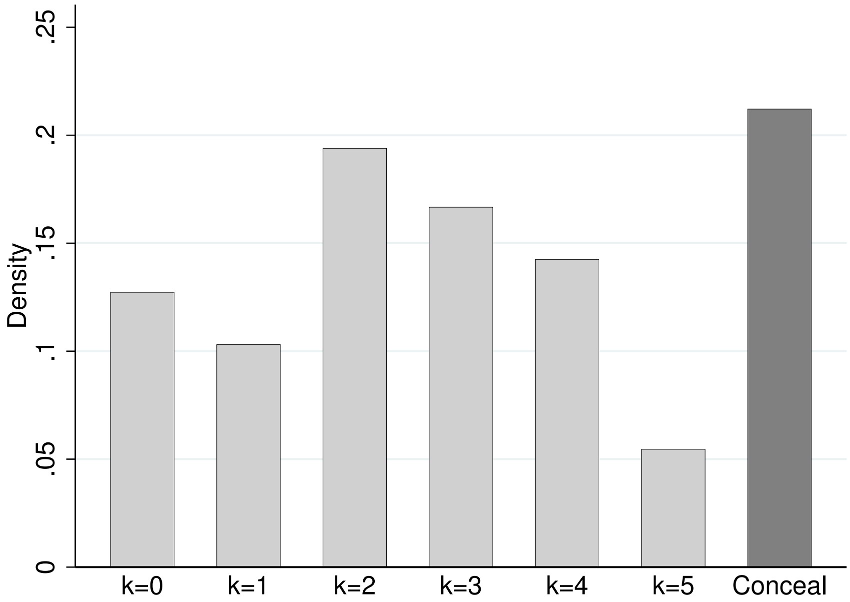


Figure 1: Distribution of choices in Market B (all periods).

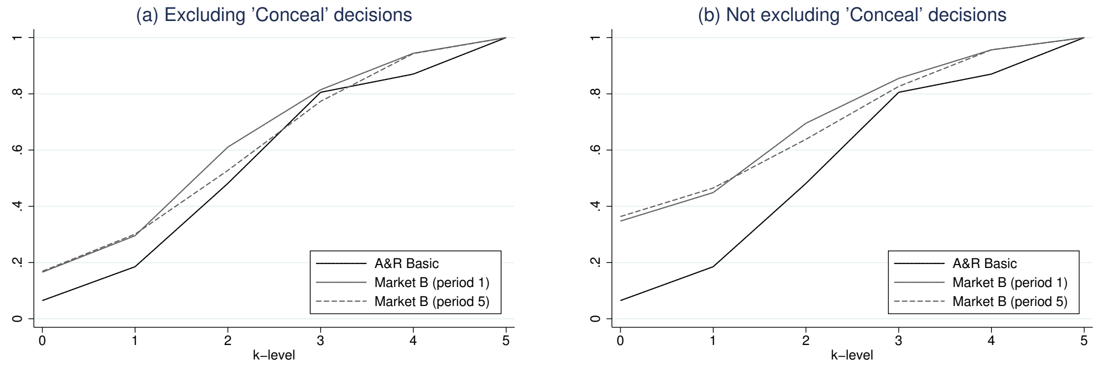


Figure 2: Cumulative distribution functions of k-levels: Arad and Rubinstein's Basic vs. Market B (first period) and Market B (period 5), with and without Conceal decisions.

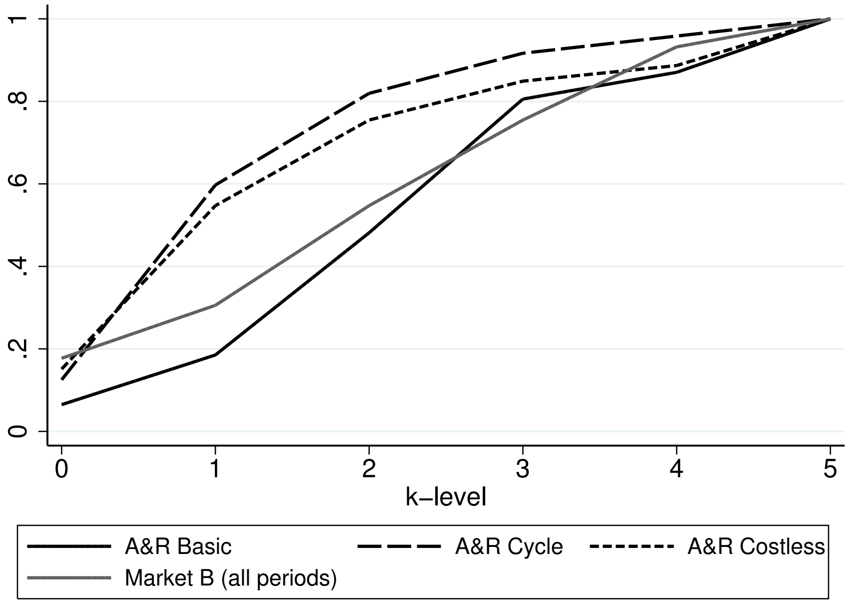


Figure 3: Cumulative distribution functions of k-levels: All A&R treatments vs. Market B, all periods and without Conceal choices.

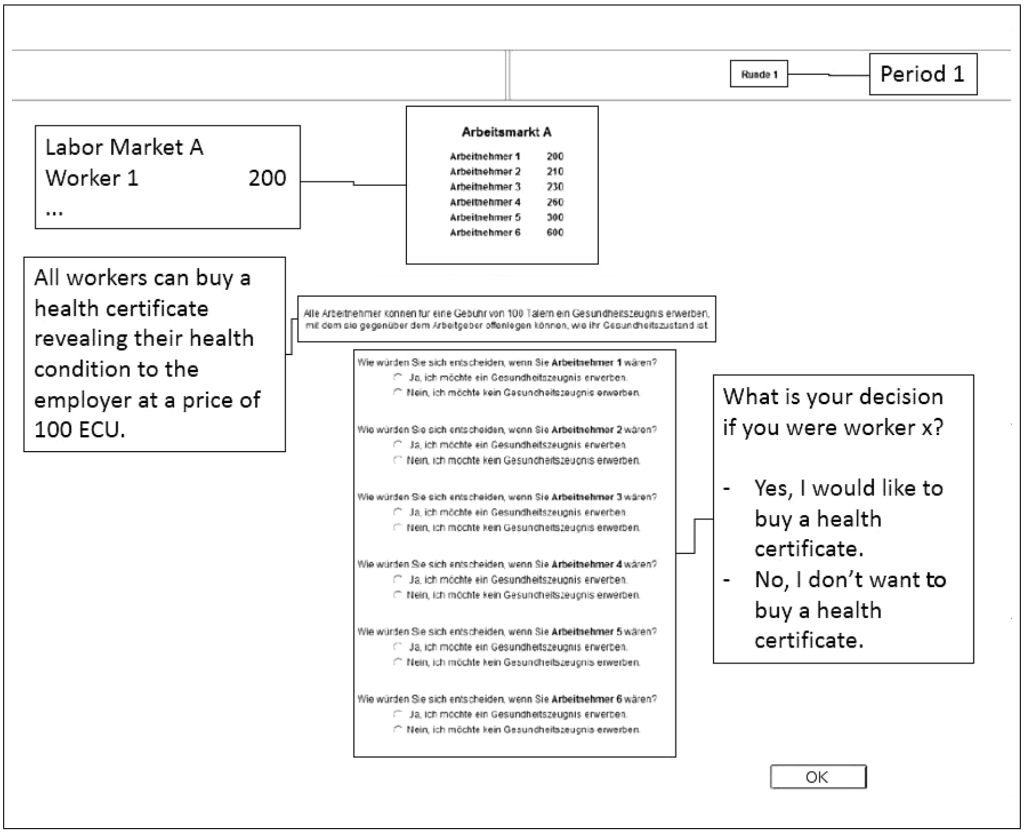


Figure 4: Screenshot of experimental software.