The nature of information and its effect on bidding behavior: laboratory evidence in a common value auction*

Isabelle Brocas
University of Southern California and CEPR

Juan D Carrillo
University of Southern California and CEPR

Manuel Castro
University of Southern California

March 18, 2011

Abstract

We study in the laboratory a series of first price sealed bid auctions of a common value good. Bidders face three types of information: private information, public information and common uncertainty. Auctions are characterized by the relative size of these three information elements. According to Nash Equilibrium theory, bids can be decomposed into two additive parts. For the private information, bidders should shade their bid. For the common uncertainty and public information, bidders should compete à la Bertrand and bid the expected and realized values respectively. We find that departures from equilibrium predictions occur not only with respect to private information but with respect to public information and common uncertainty as well. A cluster analysis shows that there is heterogeneous behavior with respect to each of these three information elements. Estimation of the Cognitive Hierarchy and Cursed Equilibrium models reveals that each model captures some important aspects of the behavior of subjects. However, the disparity of the estimated parameters as we vary the relative size of the three types of information suggests that their predictive power is limited.

*Support of the LUSK Center, the Office of the Provost at USC and the Microsoft Corporation is gratefully acknowledged.

Address for correspondence: Department of Economics, University of Southern California, Los Angeles, CA 90089, emails: <brocas@usc.edu>, <juandc@usc.edu>, <castrolo@usc.edu>.

1 Introduction

Common value auctions have been extensively studied in the laboratory. Two major findings are behavioral heterogeneity (Crawford and Iriberri, 2007) and the pervasiveness of the winner's curse (Kagel and Levin (1986, 2008)). Despite the existing literature, our knowledge of bidding behavior in those games is still incomplete. The goal of this paper is to improve such understanding. To this purpose we introduce two novel features in the design of an otherwise standard first-price sealed bid common value auction with two bidders. First, we assume that the value of the good is the sum of N independent components and that each bidder observes a subset of these components. As a result, there are three separable and clearly identified types of information in this game: private information (the components observed by only one bidder), public information (the components observed by both bidders) and common uncertainty (the components observed by no bidder). Second, we vary the number of components observed by each bidder and hold everything else constant. This changes the relative importance of each type of information and allows for a comparative statics analysis. More precisely we consider five combinations of information structures: two with private information and common uncertainty, two with private information and public information, and one with private information only.

Assuming risk neutrality, we show that the Nash Equilibrium (NE) bids in this auction can be decomposed into two additively separable parts. With respect to private information, bidders shade their bid like in a typical common value auction (see e.g. Milgrom and Weber (1982)). With respect to common uncertainty and public information, bidders compete à la Bertrand and bid the expected value and the realized value respectively.

We then analyze the results of the laboratory experiment. First, we perform descriptive statistics of aggregate bids and payoffs. Then, we run an OLS regression assuming that the bid of an individual is a linear function of public information and a polynomial function of private information, with the constant term capturing how the individual treats common uncertainty. Next, we conduct a cluster analysis. The variable we use is the average deviation of the subject's bids from the NE prediction for each combination of information structures. Since we have five different information structures, each individual is represented by five averages. Finally, we perform a structural estimation of two behavioral models, Cognitive Hierarchy (CH – Camerer et al. (2004)) and Cursed Equilibrium (CE – Eyster and Rabin (2005)).

By introducing different types of information and varying their relative importance we obtain some new interesting conclusions that we summarize below. First, the regression analysis in section 4 suggests that departures from the NE predictions occur not only with respect to private information but with respect to public information and common uncertainty as well. In particular, subjects overbid for low realizations of public information and underbid for high realizations, that is, their reaction to that information is smaller than what theory predicts. Also, subjects underestimate the common uncertainty element when the amount of information they possess is small. Second, the cluster analysis performed in section 5 reveals that the behavioral heterogeneity emphasized in the literature is not a trait that arises exclusively due to the difficulty in realizing the informational content of the rival's bid. Instead, it is partly a result of subjects treating the three types of information differently. For example, clusters 1 and 2 in our sample have different aggregate behavior when there is common uncertainty even though their behavior with respect to their private information is quite similar (and quite close to NE predictions). A key difference is that cluster 1 bids more than the expected value of the common uncertainty whereas cluster 2 does not. Third, the structural estimation of CE and CH in section 6 suggests, just like previous research, that both models capture interesting aspects of the subjects' behavior. In particular, CE captures in a parsimonious way the average bid of the data and the departure from NE with respect to private information whereas CH captures remarkably well the dispersion in the distribution of bids and the behavioral heterogeneity in the sample. However, in both models, the estimated parameters vary significantly (and unpredictably for the CH model) as we change the relative size of the three types of information. This is an important setback for the models, and one that deserves further consideration, since all the estimations use the same pool of individuals. Finally we provide two examples of the potential benefits of combining a cluster analysis with the estimation of a behavioral model. One relates to the close mapping between clusters and levels of thinking. We find that for five out of six clusters, the vast majority of subjects within a cluster are mapped into the same level of the CH model. This is quite remarkable since clustering is based on average deviations from NE, so the variable employed is only indirectly related to steps of reasoning. The other relates to the way in which clustering can inform models. Indeed, a crucial element in a CH model is the anchoring type. In our data, we find no subject of the so-called "random level 0" type, that is, a subject who bids randomly over the entire set of possible valuations $[0, 300]$. By contrast, the cluster analysis suggests that the behavior of the seven subjects in cluster 6 is consistent with a random bid over a "prudent" interval $[0, 150]$. This alternative definition for the anchoring type improves the fit of the CH model.

Our analysis relates to two strands of the experimental literature: common value auctions and auctions with variable amounts of information. Kagel and Levin (1986) is the classical reference on common value auctions in the laboratory. They assume the value of the good is drawn from some distribution (typically uniform). Bidders receive a signal which is drawn from another distribution centered around the true realization. In our study, we model the value of the good as the sum of several independent signals, and each signal may or may not be observed by bidders. This is formally closer to the models by Avery and Kagel (1997) and Klemperer (1998).[1] As noted above, the novelty of our paper lies in explicitly modeling different types of information and varying their relative importance.

There is also an experimental literature on auctions with different amounts of information. Andreoni et al. (2007) study a series of private value auctions in which bidders know not only their own valuation but also the valuation of some other bidders. Naturally, the private value setting precludes any winner's curse problem. Mares and Shor (2008) analyze common value auctions with constant informational content but distributed among a varying number of bidders. The paper explores the trade-off competition vs. precision of estimates. Overall, both papers study how the amount of private information affects the strategy of bidders. They do not consider how the existence of other types of information may affect their strategy.

The paper proceeds as follows. The theoretical framework is developed in section 2 and the experimental setting is exposed in section 3. The aggregate analysis of the experimental data is discussed in section 4 and the cluster analysis is performed in section 5. Behavioral models are tested in section 6 and conclusions are presented in section 7. Proofs, tables and figures are relegated to the Appendix.

2 Theoretical model

Consider a single good made of N components (with N even and greater or equal than four). Each component $i \in \{1, \ldots, N\}$ has a value x_i independently drawn from a continuous distribution with positive density $g(x_i)$ on $[\underline{x}, \overline{x}]$ and cumulative distribution $G(x_i)$. The total value of the good is the same for every individual and equal to the sum of the components, $V = \sum_{i=1}^{N} x_i$.

Two risk neutral bidders, A and B indexed by j, bid for this good in a first price sealed bid auction with no reserve price. Before placing their bids, A observes the first r components of the good, $\{x_1, \ldots, x_r\}$, and B observes the last r components of the good, $\{x_{N-r+1}, \ldots, x_N\}$, where $r \in \{1, \ldots, N-1\}$.

[1] In those studies, each bidder has one private signal. The value of the good is the sum of signals for one bidder and the sum of signals plus a private value component for the other bidder. Therefore, when the private value component is zero, their model is equivalent to one of our five treatments (the one with only private information).

Note that each bidder observes exactly r components and does not observe exactly $N-r$ components. Each bidder knows which components are and are not observed by the other bidder. Also, bidders always have *private information*: by construction, some components are only observed by A (e.g., x_1) while other components are only observed by B (e.g., x_N). It is useful to split the auction into three cases. When $r < N/2$, there is *common uncertainty*: the components $\{x_{r+1}, \ldots, x_{N-r}\}$ are not observed by any bidder. When $r > N/2$, there is *public information*: the components $\{x_{N-r+1}, \ldots, x_r\}$ are observed by both bidders. Finally, when $r = N/2$, there is neither common uncertainty nor public information. The value of the good is the sum of the private information of both bidders.

As r increases, both bidders have more information about the value of the good. From $r = 1$ to $r = N/2$, private information increases and common uncertainty decreases. From $r = N/2$ to $r = N$, private information decreases and public information increases. It is useful for the rest of the analysis to introduce the following notations.

- $X_A^r = \sum_{i=1}^{\min\{r,N-r\}} x_i$: the sum of A's private information.
- $X_B^r = \sum_{i=\max\{N-r+1,r+1\}}^{N} x_i$: the sum of B's private information.
- $E[X_\emptyset^r] = \sum_{i=r+1}^{N-r} E[x_i]$: the expected common uncertainty when $r < N/2$.
- $X_{\text{Pub}}^r = \sum_{i=N-r+1}^{r} x_i$: the sum of public information when $r > N/2$.

Therefore, for all r, $X_A^r, X_B^r \in [\underline{X}^r, \overline{X}^r]$ with $\underline{X}^r = \min\{r, N-r\} \cdot \underline{x}$ and $\overline{X}^r = \min\{r, N-r\} \cdot \overline{x}$. Since X_j^r is a sufficient statistic for the value of the components privately known by j, we can restrict our analysis to strategies that are a function of the total private information of a bidder rather than a function of each component privately observed (see Milgrom and Weber, 1982). When $r \leq N/2$, the private information of bidders A and B are independent random variables with cumulative distribution $F^r(\cdot)$ and density $f^r(\cdot)$. When $r \geq N/2$, the private information of bidders A and B are independent random variables with cumulative distribution $F^{N-r}(\cdot)$ and density $f^{N-r}(\cdot)$. Given each component x_i has distribution $G(x_i)$ and components are independent, we have

$$F^r(X_A^r) = \int_{\underline{x}}^{\overline{x}} \ldots \int_{\underline{x}}^{\overline{x}} G(X_A^r - x_1 - \cdots - x_{r-1}) g(x_1) \ldots g(x_{r-1}) dx_1 \ldots dx_{r-1}.$$

and analogously for $F^r(X_B^r)$. Proposition 1 characterizes the optimal bidding strategies and equilibrium utilities in this auction as a function of r.

Proposition 1. (Nash equilibrium) *The optimal bidding strategy of agent j is:*

- $b^r(X_j^r) = E[X_\emptyset] + 2\left(X_j^r - \frac{\int_{\underline{X}^r}^{X_j^r} F^r(S)dS}{F^r(X_j^r)}\right)$ *when* $r < N/2$,

- $b^r(X_j^r) = 2\left(X_j^r - \frac{\int_{\underline{X}^r}^{X_j^r} F^r(S)dS}{F^r(X_j^r)}\right)$ *when* $r = N/2$,
- $b^r(X_j^r) = X_{\text{Pub}}^r + 2\left(X_j^r - \frac{\int_{\underline{X}^{N-r}}^{X_j^r} F^{N-r}(S)dS}{F^{N-r}(X_j^r)}\right)$ *when* $r > N/2$.

and the equilibrium expected utility of agent j *is:*

- $U_j^r(X_j^r) = \int_{\underline{X}^r}^{X_j^r} F^r(s)ds$ *when* $r \leq N/2$,
- $U_j^r(X_j^r) = \int_{\underline{X}^{N-r}}^{X_j^r} F^{N-r}(s)ds$ *when* $r \geq N/2$.

The optimal bidding function can be split in two parts. The first part reflects common uncertainty when $r < N/2$ and public information when $r > N/2$, while the second part reflects private information for all r. For the first part, agents compete à la Bertrand and end up bidding the expected value of the common uncertainty or the realized value of the public information. For the second part, agents trade-off price vs. likelihood of getting the good and shade their bids accordingly.

To determine the optimal shading, recall that in a private value auction, the symmetric Nash equilibrium bidding strategy for an agent with valuation X drawn from distribution $F(\cdot)$ is $b^*(X) = X - \frac{\int_{\underline{X}}^{X} F(S)dS}{F(X)}$. In a common value auction, and as seen in Proposition 1, each agent bids $2b^*$ to account for the private information components. Hence, agent A bids as if the value of the private information of agent B were equal to the value of his own private information. The reason is simply that, in a symmetric equilibrium, the winner is the bidder with the highest signal. If A estimates the value of B's private information to be higher than his own, he may overpay in case of winning. If he estimates this value to be lower than his own, he may lose the good at a price lower than what he was actually willing to pay for it.

Finally, agents do not obtain any rent from common uncertainty or public information. Therefore, their equilibrium expected utility is the same as in a standard common value auction.

3 Experimental setting

We conducted 6 sessions with 8 or 10 subjects per session for a total of 52 subjects. The subjects were students at the California Institute of Technology who were recruited by email solicitation, and all sessions were conducted at the Social Science Experimental Laboratory (SSEL). All interaction between subjects was computerized using an extension of the open source software package Multistage Games.[2] No subject participated

[2]Documentation and instructions for downloading the software can be found at http://multistage.ssel.caltech.edu.

in more than one session.

In each session, subjects made decisions over 15 paid matches, with each match being divided into 5 rounds. At the beginning of a match, subjects were randomly matched into pairs and randomly assigned a role as bidder A or bidder B. Pairs and roles remained fixed for the 5 rounds of a match. At the end of the match, subjects were randomly rematched into new pairs and reassigned new roles.

The game closely followed the setting described in section 2. Subjects in a pair had to bid in a first price sealed bid auction for a good made of $N = 6$ components. Each component $i \in \{1, \dots, 6\}$ contained x_i tokens drawn from a uniform distribution in $[0, 50]$ (so $G(x_i) = x_i/50$).[3] The total value of the good was common to both bidders and equal to the sum of the six components. Visually, each component was represented by a box in the computer screen. The number of tokens inside each of the six boxes was drawn at the beginning of the match. Subjects could see the six boxes but not their content.

The match was then divided into 5 rounds. Round 1 corresponded to $r = 1$ in the theory section. Subject A observed x_1 (the content of box 1) and subject B observed x_6 (the content of box 6). Neither subject observed x_2 to x_5 (the content of boxes 2 to 5). Given this information, both participants submitted a bid for *the entire good* of value $V = \sum_{i=1}^{6} x_i$. Subjects could not see the bid of the other subject, instead they moved to round 2. Round 2 corresponded to $r = 2$ in the theory section. Subjects A and B could still observe x_1 and x_6 respectively, but now they could also see the content of a second box (x_2 for bidder A and x_5 for bidder B), and placed a new bid again for the entire good V. This process continued until round 5, $r = 5$ in the theory section, where subject A observed x_1 to x_5 and subject B observed x_2 to x_6.[4] At the end of the fifth round, the value V of the item and the five bids of each subject were displayed on the computer screen. One of the rounds was randomly selected by the computer, and subjects were paid for their performance in that round (this procedure allowed us to have a higher conversion rate than if we had paid subjects for every round without, in principle, affecting their strategy). Payoffs were computed according to the standard rules of a first price auction without reserve price: the highest bidder would get the item and pay his bid, while the lowest bidder would get nothing and pay nothing.

Summing up, the only variable that changed between rounds was the amount of information each bidder had, which increased from r to $r+1$. By contrast, the opponent,

[3]To simplify computations, we restricted x_i to integer values.

[4]We did not consider the least interesting cases, $r = 0$ and $r = 6$. In the former case, there is no information so risk-neutral subjects should bid the expected value. In the latter case, there is full information so subjects should bid the realized value.

role and total value of the item remained the same for the entire match. Naturally, it was crucial not to disclose the bids of the opponent between rounds since they contained information which could have been used as a signaling device, making the theoretical analysis substantially more complicated. Finally, note that as we moved from one round to the next, $F^r(\cdot)$ the distribution function of the private information of a bidder changed.[5] Because the entire distribution (and not only the realized values) affects the optimal bid (see Proposition 1), it *is not* possible to separate the optimal bid in round $r+1$ into the optimal bid in round r plus the bid due to x_{r+1}, the content of box $r+1$.

All participants started the experiment with an endowment of 400 tokens to which the payoffs of each match were added or subtracted. At any moment, the participants could not bid more than their current stock of tokens, resulting in a potential selection effect due to liquidity constraints. However, this constraint was never binding in the experiment. Indeed, the stock of tokens of all participants was always greater than 300, the maximum value of the item and therefore the maximum possible bid.[6]

At the beginning of each session, instructions were read by the experimenter standing on a stage in the front of the experiment room. The experimenter fully explained the rules and how to operate the computer interface. After the instructions were finished, one practice match was conducted, for which subjects received no payment. After the practice match, there was an interactive computerized comprehension quiz that all subjects had to answer correctly before proceeding to the paid rounds. Then, the 52 subjects participated in 15 paid matches each of them divided into 5 rounds for a total of 75 bids per subject. Opponents, roles and values in the boxes were randomly reassigned at the beginning of each match and held constant between rounds of a match. In the end, subjects were paid, in cash, in private, their accumulated earnings, which was equal to their initial endowment plus the payoffs of all matches. The conversion rate was $1.00 for 25 tokens, so each good was worth between 0 and $12. Sessions averaged one hour in length, and subjects earnings averaged $27.

4 Aggregate analysis

For our particular case with 6 boxes and values drawn from a uniform distribution in $[0, 50]$ ($N = 6$ and $G(x_i) = x_i/50$), Proposition 1 implies that the expected bid $E[b^r]$ is U-shaped across rounds: decreasing over rounds when $r \leq N/2$ and increasing over

[5]In rounds 1 and 5, $F^1(X^1)$ is uniform in [0,50]. In rounds 2 and 4, $F^2(X^2)$ is triangular in [0,100], etc.

[6]We constrained the bids to be between 0 and 300: the minimum and maximum possible values of the good before any information is revealed.

rounds when $r \geq N/2$.[7] Indeed, bid shading is increasing in the amount of private information that both bidders have, whereas for the public information and common uncertainty elements agents always bid the realized and expected values respectively. Also, since we focus on symmetric equilibria, when both bidders reduce their bids, their ex-ante expected utility increases. This happens because the ex-ante expected value of the good is constant. Therefore, the expected utility is hump-shaped across rounds: increasing over rounds when $r \leq N/2$ and decreasing over rounds when $r \geq N/2$. In the next sections we compare these properties of equilibrium behavior with the empirical counterpart.

4.1 Aggregate bids and payoffs

The first cut at the data consists in an aggregate analysis of the data per round in order to compare actual behavior with the NE predictions derived in Proposition 1. Figure 1 shows the difference between actual bids and NE predictions in each round. For each observation, we compute the NE bid and subtract it from the corresponding observation. The line in the middle is the median of this statistic, whereas the top and bottom lines are the 75th and 25th percentiles. The notches are the 95% confidence interval for the median. We can make two main observations. First, deviations from NE predictions exhibit a hump shaped pattern across rounds: they increase from round 1 to round 3 and decrease from round 3 to round 5. Deviations of the median in rounds r and $r+1$ are significantly different from each other at the 95% level for all r, and they are all significantly different from zero. There is underbidding in round 1 and overbidding in rounds 2 to 5. Second, the dispersion in the data decreases over rounds, that is, it is inversely related to the total amount of information (private and public).

Table 1 displays the average actual bids per round and the NE predictions. It also displays the best response to the empirical distribution (BR) for comparison. To construct this table, we computed the NE bid and the best response to the empirical distribution for each observation. Ideally, we would want to compute the empirical distribution of the bids for each possible value of private information. Then, we would be able to calculate the expected gain of each possible bid for each possible value of private information. The best response to the empirical distribution would be the bid that maximizes the expected gain. However, this procedure would require a massive amount of data. Thus, we decided to divide the values of private information into 52

[7] The result is true at the boundaries for any distribution ($E[b^0] > E[b^1]$ and $E[b^{N-1}] < E[b^N]$ for all $G(x_i)$). We conjecture that it should hold for all r when x_i is in a certain class of distributions but we have been unable to determine the properties of that class.

bins, each with 15 observations.[8] We used the average of private information in each bin as the value of the opponent's private information in that bin and then computed the best response to the empirical distribution using the method described above.

As noted above, the average NE bid is U-shaped across rounds. In the data, the average bid is increasing across rounds. Subjects tend to become more confident and therefore bid more aggressively as their total information (whether it is private or public) increases. Note, however, that the difference between the average bid and the average NE prediction is relatively small in percentage terms (between 1% and 13%). The best response to the empirical distribution is hump-shaped across rounds, which is the opposite pattern of the NE predictions. Typically, it is optimal to underbid significantly in rounds 1 and 5 and overbid in round 3. Subjects deviate more from the best response in late rounds, with a maximum difference of 17% in round 5.

Table 2 displays the average gain.[9] Notice that the loser of the auction makes zero profits whereas the winner can make positive or negative profits. Gains at NE were computed as if all subjects in the population were bidding the NE. The table shows that the average gain in our sample decreases across rounds, instead of being hump-shaped. This results from the fact that actual bids are increasing across rounds. Despite the small reported differences in bids, the percentage difference in gains is significant. By underbidding in round 1, subjects increase their payoff by 19% whereas by overbidding in rounds 2 to 5 they decrease their payoff between 15% and 38%. Moreover if subjects best responded to the empirical distribution their gains could have increased by as much as 50% in some rounds. We summarize the findings of this section in the following result.

Result 1. *Average empirical bids and gains depart from NE predictions: bids are increasing across rounds instead of U-shaped and gains are decreasing across rounds instead of hump-shaped. Overbidding in rounds 2 to 5 imply losses of up to 38%.*

4.2 Aggregate bids conditional on private information

In this section, we analyze bids as a function of private information only. We allocate the values of private information into 5 bins containing the same number of observations. For each bin, we compute the average empirical bid, the Nash Equilibrium bid (NE) and the Best Response to the empirical distribution (BR).[10] Figure 2 displays

[8] To have the same precision when estimating the empirical distribution in each bin, we decided to have the same number of observations per bin. This implies that bins have different lengths.

[9] Remember that we only paid subjects for one randomly drawn round per match. However, when we mention 'gain' we refer to the profits subjects would have made if they were paid for all rounds.

[10] Again, we decided to have the same number of observations per bin instead of equal length bins.

for each round the bids as a function of private information. Interestingly, the actual bidding function is upward slopped but flatter than the NE bidding function. Subjects overbid for low values of private information and underbid for high values of private information in all rounds except round 1 where they always underbid. As noted in Table 1 for the averages, in rounds 4 and 5 the actual bidding functions are above the best response functions whereas in rounds 1, 2 and 3 they are remarkably close to each other.

Result 2. *The empirical bidding function is less responsive to private information than predicted by NE theory.*

4.3 Analysis of bidding strategies

The first cut at the data has allowed us to find bidding differences at the aggregate level between NE predictions and the actual data. In particular, Result 1 implies that the average expected bids are too low in round 1 and too high thereafter. Also, Result 2 indicates that subjects overbid for low values of private information and underbid for high values of private information. Taken together, the results suggest that bidding behavior is heterogenous across rounds and that bidders have different attitudes towards private information, public information and common uncertainty. Said differently, the type and quantity of information is likely to play a role when bids are computed.

In this section, we take a closer look at bidding functions across rounds to understand better how subjects react to the three types of information. To do so, we compute an appropriate parameterized approximation of the NE bidding function and estimate it.

Note that the NE bidding function is a linear function of public information and a polynomial function of private information. In rounds 1 and 5, the polynomial is of degree 1, whereas in rounds 2, 3 and 4 it is the ratio of two higher order polynomials. A polynomial approximation is therefore a good candidate for our analysis. For each round, we compute a cubic approximation of the NE bidding function. Namely, we set the following relationship between the Nash Equilibrium bid, b^r, and the various types of information:

$$b^r = \alpha_0 + \alpha_1 Priv^r + \alpha_2 (Priv^r)^2 + \alpha_3 (Priv^r)^3 + \alpha_4 Pub^r + \eta(Priv^r)$$

In this equation, superscript r denotes the round, $Priv$ is the variable of private information (tokens observed by only one bidder) and Pub is the variable of public information (tokens observed by both bidders). The constant term α_0 is the coefficient of common uncertainty. At NE, α_0 is the expected number of tokens in the boxes that nobody observes. These would be 100 in round 1, 50 in round 2 and 0 in the remaining rounds under the exact polynomial expression for $Priv$. The coefficient of public information, α_4, appears only in rounds 4 and 5. At NE, it is equal to 1 since both bidders compete à la Bertrand. Also, α_1, α_2 and α_3 are the coefficients for the cubic approximation of private information. Finally, η is the error of the cubic approximation for each level of private information. It is different from zero in rounds 2, 3 and 4. The α-coefficients are reported in Table 3 for each round.[11]

We then use the bids of our data, b_o^r, to run the following regression:

$$b_o^r = \beta_0 + \beta_1 Priv_o^r + \beta_2 (Priv_o^r)^2 + \beta_3 (Priv_o^r)^3 + \beta_4 Pub_o^r + \varepsilon_o$$

where superscript r denotes the round and subscript o the observation. The α-coefficients have been replaced by β-coefficients with the same interpretation. For each round we compute the coefficients of the Feasible Generalized Least Squares (FGLS) Random Effects regression for each of the variables and the t-test to check if each coefficient is significantly different from the coefficient predicted by NE (α parameters above).[12] We also perform a global significance F-test to inspect if, overall, the data bidding function is different from the NE bidding function. Table 3 and Figure 3 display the results of this exercise. Our main findings are the following.

First, the proportion of the bid driven by common uncertainty is smaller than the NE prediction in round 1, similar in rounds 2 and 5 and higher in rounds 3 an 4. This result follows from a comparison between the estimated β_0-coefficients and the predicted NE α_0-coefficients. Therefore, deviations from NE regarding common uncertainty are hump-shaped. Subjects underestimate common uncertainty when they have little total information, and they tend to overbid with respect to it as private information grows.

Second, subjects react less to public information than what theory predicts. Indeed, β_4 is close to but smaller than the corresponding NE prediction. This result in overbidding for low values of public information and underbidding for high values of public information.

Third, the difference between subjects' reaction to private information and the NE prediction is round dependent. Subjects underbid for all values of private information

[11] We opted for a cubic rather than quadratic approximation, because the latter performs badly for extreme values of private information. For example, in round 3, the quadratic approximation of NE has a constant term of -20 instead of the theoretical prediction of 0.

[12] The FGLS estimator makes use of the panel data structure to get more precise coefficients. We also performed the Hausman test and we could reject the presence of unobserved fixed effects. The variance estimator is clustered by subject.

in round 1 and overbid for all values of private information in round 3. In rounds 2, 4 and 5, they overbid for low values of private information and underbid for high values of private information. In all cases and as noted in Result 2, subjects react less to changes in private information than what NE predicts. This is easier to visualize with the two- and three-dimensional plots of Figure 3 than with the coefficients of the cubic polynomial approximation in Table 3.

Taking these results together, we can draw a link between the attitude towards information *per round* and the aggregate results. Subjects place relatively low bids in round 1 because they underestimate both common uncertainty and private information. They place relatively high bids in round 3 because they overestimate their private information (that is, they succumb to the winner's curse). Therefore, in round 1 subjects have a higher payoff than they would have if everyone was playing the NE and the converse happens in round 3. Last, although subjects bid on average close to NE in rounds 2 and 5, the dispersion in those bids implies that the average gains are substantially smaller than what they would be if everyone was playing the NE. Indeed, for bidders with high values of private information, NE predicts that they should win the auction with high probability and make a large profit. Because these subjects are not bidding much more than the bidders with low values, they decrease their likelihood of winning the auction and thus forego a large gain. We summarize these findings in the following result.

Result 3. *The reaction of subjects to both private and public information is smaller than predicted by theory. The difference in the reaction to common uncertainty relative to the theory is hump-shaped across rounds.*

Overall, and as we can see from the F-test, the actual bidding functions are significantly different from the NE bidding functions, with the exception of round 5. Moreover, there is a large dispersion in the first rounds, hence the small R^2.

5 Cluster Analysis

In order to understand finer aspects of bidding strategies and uncover the reasons for the observed dispersion, we search for trends at a disaggregate level. One possible approach would be to do a subject-by-subject analysis (as in Costa-Gomes et al. (2001) for example). Even though this is in theory the most informative strategy, the reduced number of observations for each individual in each round would prevent us from making any confident assessment. We therefore take an alternative approach, which is to search for clusters of subjects (as in Camerer and Ho (1999) and Brocas et al. (2009)). It is an intermediate approach, as the aggregate approach implicitly assumes a single cluster is of interest while the subject-specific approach requires each subject to be in a singleton cluster. One advantage of the method is to provide an implicit measure of how well these two extreme cases capture the observed behavior.

5.1 Method

To find the clusters, we use the average deviation from NE for each of the 52 subjects in each round. Each subject is thus represented by five averages. There is a wide array of heuristic clustering methods that are commonly used but they usually require the number of clusters and the clustering criterion to be set ex-ante rather than endogenously optimized. By contrast, mixture models treat each cluster as a component probability distribution. Thus, the choice between different numbers of clusters and different models can be made using Bayesian statistical methods (Fraley and Raftery, 2002). Note that popular heuristic approaches such as 'k means clustering' are equivalent to mixture models where a particular covariance structure is assumed.

Given that we do not want to presuppose a particular cluster structure, we follow the second clustering strategy and implement a model-based clustering analysis. We consider a maximum of ten components/clusters and assume a diagonal covariance matrix. This implies that the correlation between the dimensions is zero and there is no restriction on the variance. We first fix the number of clusters from 1 to 10 and, for each of the models, we estimate the covariance matrix as well as the clustering that maximizes the likelihood function. We use random clustering as an initial guess.[13] Overall, for any possible number of clusters we obtain a clustering and the covariance matrices for each cluster, and we compute the corresponding Bayesian Information Criterion (BIC). Given this information, the optimal (endogenous) number of clusters is the one for which BIC is maximized. For our data, the BIC is maximized for 6 clusters with diagonal covariance matrix. These clusters, labeled 1 to 6, contain 11, 13, 6, 6, 9 and 7 subjects respectively. Notice that the clusters are comparable in size, although this does not necessarily need to be the case.

5.2 Bidding behavior by cluster

In order to determine the properties of our clusters, we perform the same regression analysis as in section 4.3 for each cluster separately. The results can be found in Table 4. Figure 4 is the equivalent of Figure 1 for each cluster (deviations from NE predictions). Finally, Table 5 reports the average gains in each cluster.

[13] We ran the model several times and the results are consistent.

Clusters 1 and 2 bid, on average, close to NE. A closer look at their bidding strategies reveals that both clusters react less to increases in private and public information than what NE predicts. Still a few differences can be noted. Subjects in cluster 1 *slightly overbid* in every round while subjects in cluster 2 *slightly underbid* in every round. This is in part due to the fact that in rounds 1 and 2 cluster 1 bids above the expected value of the common uncertainty whereas cluster 2 does not. The bidding function of the latter is also relatively closer to NE when there is more private information. Even though they bid close to Nash, these subjects do not obtain the highest payoffs. This is true because their strategy departs from the best response to the empirical distribution. Indeed, it is optimal to underbid significantly in rounds 1 and 5 and to bid close to Nash in rounds 2, 3 and 4.

Cluster 3 is composed of a relatively heterogeneous group of subjects who *underbid significantly in every round* (the heterogeneity can be seen in the large variance of the data). There is no pattern regarding common uncertainty and subjects do not react to private or public information. Nevertheless, they obtain relatively high payoffs because their underbidding strategy coincides often with the best response to the empirical distribution.

Subjects in Cluster 4 *overbid significantly in every round* and are extremely homogeneous. Typically, those subjects have a clear biased attitude towards common uncertainty, which they overestimate, but they react as Nash players with respect to the two other types of information.[14] Their average gains are lowest given their substantial overbidding and the optimality of underbidding.

Cluster 5 is different from the other clusters because the *bidding behavior is round-dependent*. Indeed, these subjects underbid in round 1, bid close to NE in rounds 2, 4 and 5 and overbid in round 3. The bidding function of these subjects with respect to private information is very close to NE in rounds 1, 4 and 5. In rounds 2 and 3, the slope of private information is higher than the NE predictions, that is, they *react more to private information* than what NE predicts. As such, their attitude towards private information is also different from what we observed in other clusters. These subjects obtain the highest gains both overall and per round since their bidding strategy is close to the best response to the empirical distribution.[15]

[14]The slope of private information is close to NE for low/medium values of private information and higher than NE for high values of private information. The slope of public information is very close to the NE predictions.

[15]It would have been interesting to investigate whether these subjects learn over time to best respond to the empirical behavior of the population. We could not perform such analysis due to the reduced number of observations (9 subjects).

Last, subjects in cluster 6 *underbid substantially* (most notably in round 1) and are extremely heterogeneous. Compared to cluster 3, their bids are significantly lower and they are also more dispersed. Subjects in this cluster lose the auction most of the time and therefore obtain small payoffs.

Result 4. *Subjects in different clusters treat each type of information differently. It can be summarized as:*

Cluster	*Overall*	*Private Information*	*Public Information*	*Common Uncertainty*	*Heterogeneity*
1	*Slight overbid*	*Close to NE*	*React less than NE*	*Above NE*	*Small*
2	*Slight underbid*	*Close to NE*	*React less than NE*	*Below NE*	*Small*
3	*Underbid*	*No reaction*	*No reaction*	*No pattern*	*Large*
4	*Overbid*	*Equal to NE*	*Equal to NE*	*Above NE*	*Very small*
5	*Close to NE*	*Round dependent*	*Round dependent*	*Round dependent*	*Small*
6	*Strong underbid*	*No pattern*	*No pattern*	*Below NE*	*Large*

As in previous studies (e.g., Crawford and Iriberri (2007), Brocas et al. (2009)), we find clusters of subjects characterized by substantial homogeneity within groups and heterogeneity across groups. More interestingly, the treatment of private information, public information and common uncertainty differs across clusters: the type of information is a crucial explanatory variable for bidding behavior even though *it is not a cluster variable*. Also, subjects sometimes treat the same type of information differently across rounds. That is, the amount of a given type of information also impacts bidding behavior.

The heterogeneity across clusters suggests that subjects may have different cognitive abilities. This is most apparent for clusters 1, 2, 4 and 6. Subjects who strongly underbid (cluster 6) or systematically overbid (cluster 4) are likely to use simple heuristics. They fail to realize the link between bids and information and they end up collecting the smallest payoffs. Subjects who play relatively close to NE (clusters 1 and 2) look sophisticated enough to approximate NE behavior for all types of information but, at the same time, they overestimate the ability of their opponents to play NE. As a consequence, they do not obtain the highest gains. The reasoning made by clusters 3 and 5, who obtain the highest payoffs, is more difficult to grasp. The bidding pattern of cluster 3 suggests they may be using a 'lucky' heuristic that turns out to work well in this type of auction. As for cluster 5, their ability to bid close to NE and to depart from it at the correct time is intriguing. Perhaps they do realize the limitations of their opponents and take advantage of this knowledge. As we will see below, a combination of clustering and model estimation can help understand the differences in behavior.

6 Behavioral Models

As seen in section 4, there are departures from NE predictions in two important dimensions: across rounds and across types of information. This suggests that subjects compute different strategies than NE. Several leading behavioral models have been proposed to explain departures in settings with private information (see e.g. Crawford and Iriberri (2007) and Carrillo and Palfrey (2009) for estimations of several different models). In this section, we explore two of them, namely the Cursed Equilibrium (CE) and Cognitive Hierarchy (CH) models. We do not report results based on the Quantal Response Equilibrium model (QRE – McKelvey and Palfrey (1995)) for empirical and technical reasons. First, QRE performs poorly with our data. Indeed, for our parameter specification QRE predicts aggregate underbidding whereas we empirically observe aggregate overbidding with the exception of round 1. Second, the strategy space is very large, with 301 possible bids for each value of private and public information. Therefore, solving the system of equations required to calculate QRE exceeds the limits of a normal computer. To tackle this issue we would have to aggregate the possible bids and values of private and public information into bins, a procedure that would disregard a significant amount of information. A brief discussion of the results of the QRE model is reported in Appendix A.2. Finally, there is another recent behavioral theory for games with private information, the Behavioral Equilibrium model (BE – Esponda (2008)). This equilibrium concept is related to CE as it also assumes an imperfect account of the link between rival's bid and rival's information. However, the BE model relates the behavior of players with the information revealed in the end of each game. In our setting, BE predictions are equivalent to NE predictions because in all auctions the two highest bids and the value of the good are revealed to all bidders. Thus bidders have all the required information to make accurate forecasts about both the probability of winning the auction and the value of the good. Therefore, we did not estimate it either.

6.1 Theory

6.1.1 Cursed Equilibrium (CE)

In the CE model, each bidder systematically underestimates the correlation between the opponent's bid and private information. In a χ-cursed equilibrium, all bidders believe that with probability χ there is no correlation and with probability $(1-\chi)$ other bidders are also χ-cursed, with $\chi \in [0,1]$. The model is equivalent to NE when $\chi = 0$. Subjects are said to be "fully cursed" when $\chi = 1$. In our setting and following Eyster and Rabin (2005), the expected utility of a cursed bidder A can be computed analytically. When $r < N/2$, assuming bidder B bids according to an increasing bidding function $b^{CE,r}(X_B^r)$ and denoting A's bid by $b_A^{CE,r}$, the expected utility can be written as:

$$\begin{aligned} U_A^{CE,r} \quad = \quad & \Pr\left(b_A^{CE,r} \geq b^{CE,r}(X_B^r)\right)\Big[\chi\left(X_A^r + E[X_\emptyset^r] + E[X_B^r]\right) \\ & +(1-\chi)\left(X_A^r + E[X_\emptyset^r] + E\left[X_B^r \mid b_A^{CE,r} \geq b^{CE,r}(X_B^r)\right]\right) - b_A^{CE,r}\Big] \end{aligned}$$

The CE bid can then be determined using the same procedure as we did for the NE bid, and we get:

$$b^{CE,r}(X_A^r) = E[X_\emptyset^r] + \chi E[X_B^r] + (2-\chi)\left(X_A^r - \frac{\int_{\underline{X}^r}^{X_A^r} F^r(s)ds}{F^r(X_A^r)}\right)$$

When $r > N/2$, we just need to replace $E[X_\emptyset^r]$ by X_{Pub}^r and $F^r(\cdot)$ by $F^{N-r}(\cdot)$. When $r = N/2$ the term $E[X_\emptyset^r]$ disappears and $F^r(\cdot)$ becomes $F^{N/2}(\cdot)$. Comparing the CE and NE bids, $b^{CE,r}(X_j^r)$ and $b^r(X_j^r)$, we obtain the following result.

Proposition 2. (Cursed equilibrium) *CE predicts overbidding for all continuous c.d.f. $G(x)$ and for all r. Overbidding is entirely driven by private information since a CE bidder treats public information and common uncertainty just like a NE bidder.*

Cursed bidders always overbid compared to Nash players. This occurs because cursed bidders do not fully comprehend the link between the bid of the opponent and the signal he has about the value of the item. Higher levels of cursedness (higher χ) imply more overbidding. Also, a larger component of private information implies more aggregate overbidding. Therefore, for any given χ, cursed equilibrium predicts the highest level of overbidding in round 3, and the smallest in rounds 1 and 5. Finally but importantly, CE handles public information and common uncertainty as NE does. This is true because CE is a model of bounded rationality about the correlation between actions and private information, and does not predict any deviations relative to the two other types of information.

6.1.2 Cognitive Hierarchy (CH)

The CH model relaxes the assumption of accurate and homogeneous beliefs. It assumes there are different levels of strategic thinking in the population, and each level best responds to a mixture of lower levels. The model is built around an anchoring type, level-0, that behaves in a non-strategic way. In this paper, we follow the approach of Crawford and Iriberri (2007) and assume that each type believes everyone else's type corresponds to the level immediately below. Formally, level-1 best responds to level-0, level-2 best responds to level-1, and so on.[16]

[16] Instead, in the original theory (Camerer et al., 2004) each level believes that the population is distributed among all levels below.

The definition of level-0 is crucial since it anchors the beliefs and therefore the actions of all other types. Here again, we take a conservative route and follow Crawford and Iriberri (2007) who define two possible types of level-0: random and truthful. A random level-0 (henceforth RL_0), the most common in cognitive hierarchy models, chooses a bid uniformly random over all the possible bids. A truthful level-0 (henceforth TL_0) bids the amount of the boxes he observes plus the expected amount of the boxes he does not observe, that is, he does not take into account any strategic or informational effects.[17] We then define RL_k and TL_k as the level types k ($\in \{1,2\}$) anchored on RL_0 and TL_0, respectively. We do not consider levels 3 and above because they are unlikely to be observed empirically.

By construction, this model delivers heterogeneous behavior. Unlike the CE model, we cannot determine a priori whether a level-k subject will underbid or overbid, as it will depend on the parameters of the model, $G(x)$ and N. Figure 5 displays in our specific setting ($G(x) = x/50$ and $N = 6$) the predicted bids for each truthful and random level-k (with $k \in \{0,1,2\}$) in rounds 1, 2 and 3 as a function of their private information.[18] Notice that TL_0 overbid in all rounds for all levels of private information since they treat public information and common uncertainty correctly but do not shade their bids due to private information. TL_1 and TL_2 bid close to NE in round 1 for some values of private information whereas RL_1 and RL_2 have a tendency to underbid. Perhaps more surprising is the fact that the bidding functions of TL_1 and RL_2 have flat portions. It occurs because some types want to win the auction for sure for a range of values of their private information. This is ensured by matching the maximum possible bid of the rival. Conversely, other types want to lose the auction for another range of values of their private information. Again, this is ensured by matching the minimum possible bid of the rival. The overall bidding functions are therefore kinked and flat on some domains (Figure 5). These conclusions are summarized below. They are determined numerically for $G(x) = x/50$ and $N = 6$.

Proposition 3. (Cognitive Hierarchy) *CH predicts heterogeneous behavior. It can generate both underbidding and overbidding and predicts that the bidding functions of some types should be flat for some range of private information values.*

[17] Notice the importance of the definition of the anchoring type. For example, TL_0 is very sophisticated in some dimensions, as it treats public information and common uncertainty just like a NE player does. Also, even the most unsophisticated subjects behave differently in a first- and a second-price auction. Although here we only consider a first-price auction, one may wonder what would be a natural definition of RL_0 and TL_0 in a second-price auction then.

[18] It is not possible to obtain analytical solutions for all bidding functions of all levels in all rounds. Also, bidding behavior in rounds 4 and 5 depends on the realization of X^r_{Pub} and is omitted. Naturally, it coincides with behavior in rounds 2 and 1 respectively whenever $X^r_{\text{Pub}} = E[X^r_{\emptyset}]$.

6.2 Estimation

In this section we estimate these two behavioral models to check how well they each fit the data. For both models we perform two sets of estimations, one where we constrain the parameters to be the same for all rounds and another where we allow the parameters to differ across rounds. If the CE and CH models are robust, the parameters in the constrained and unconstrained estimations should be similar. It is important to stress the difficulty of a direct comparison between CE and CH, due to the fact that the CH model has six parameters and the CE only one. One way to overcome this obstacle would be to allow for different levels of cursedness in the population. However, this poses a technical problem: each individual should best respond to the population and solving for that system of equations is beyond the capacity of a normal computer.[19]

We use the following econometric specification:

$$b_{om} = b_w(X_{om}) + \varepsilon_{om}$$

where b_{om} is the bid observed in the data for observation o of subject m, $b_w(X_{om})$ is the bid predicted by model $w \in \{\text{CE}, \text{CH}\}$, and ε_{om} is an error term assumed to be independently distributed and following the normal distribution $N(0, \sigma)$. Therefore, $\Pr[b_{om} \mid X_{om}] = f(b_{om} - b_w(X_{om}))$, where $f(\cdot)$ is the density of a normal distribution with mean zero and standard deviation σ.[20] For each model we find the parameters that maximize the log-likelihood of our sample. Since the CH model assigns types to subjects, we construct the likelihood function per subject and then sum all subjects' likelihoods. Therefore, if we have O observations per subject, M subjects and L types with proportions π_l that sum to one, we get the following log-likelihood function:

$$LL(\pi, \sigma \mid b) = \sum_{m=1}^{M} \log \left(\sum_{l=1}^{L} \pi_l \prod_{o=1}^{O} \Pr[b_{om} \mid X_{om}, l] \right)$$

There are no different types in the CE model and the bidding function has the level of cursedness χ as the only parameter. Therefore, the log-likelihood function is:

$$LL(\chi, \sigma \mid b) = \sum_{m=1}^{M} \sum_{o=1}^{O} \log \left(\Pr[b_{om} \mid X_{om}, \chi] \right)$$

[19] Crawford and Iriberri (2007) allow for different levels of cursedness but assume that each subject believes that everyone in the population has the same level of cursedness as him/herself. This is an interesting compromise. However, we chose not to follow that route because, in our view, a systematic bias regarding other player's behavior goes somewhat against the fundamental premises of CE.

[20] An alternative estimation strategy was to use a logit model. However, given the large number of actions, many of which were not observed in the data, we ran into an empty cell problem.

Table 6 displays the estimation results of the two behavioral models. The column labeled *all rounds* reports the findings when we constrain the parameters to be the same for all rounds. For each of the two sets of estimations, we compute two information criteria: the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The lower the value of these criteria, the better the model fits the data. According to both criteria the CH model fits the data better than the CE model. However, we do not want to stress too much this result because AIC and BIC are not the most appropriate tests for non-nested models.

The results are presented in a more intuitive way in Figure 6. It plots the densities of the empirical deviations from NE as well as the deviations from NE estimated by each model. On the x-axis are the deviations from NE and on the y-axis is the empirical and estimated probability density functions. As we can see from the graphs, each model captures different aspects of the data. CE captures accurately the average of the data (the peak of the CE density matches well the peak of the data) but not the extent of the asymmetry in the dispersion, due in part to the restricted number of parameters. Conversely, CH does not capture the average of the data in some rounds but the heterogeneity of the different levels captures quite well the dispersion.

Notice from Table 6 that the estimated level of cursedness increases with the amount of private information (it is highest in round 3 and lowest in rounds 1 and 5). The differences in cursedness between rounds r and $r+1$ are all significant at the 95% level. Also, when we perform the likelihood ratio test comparing the model where we restrict χ to be the same for all rounds with the model where we allow χ to be different across rounds, we find that the differences are statistically significant at the 98% level. Remember that the CE model predicts overbidding compared to NE (the estimation of $\chi = 0$ in round 1 reflects the observed underbidding in that round). It also predicts that, for a given level of cursedness, overbidding will be stronger the greater the total amount of private information (see Proposition 2). Our result shows that subjects increase their bids from rounds 1 to 3 and decrease their binds from rounds 3 to 5 more than what the model predicts. We summarize the results of the CE estimations as follows.

Result 5. *The CE model captures: (i) the peak of the empirical distribution of bids, (ii) the strongly monotonic bidding functions, and (iii) the departure from NE regarding private information.*

The CE model does not capture: (i) the dispersion of the empirical distribution of bids, (ii) the heterogeneity in the sample, and (iii) the departure from NE regarding common uncertainty and public information.

The results we obtain for the CH model are more puzzling. As mentioned above, the model captures well the dispersion of the data. On the other hand, the proportions of the different types change drastically across rounds, and we have not been able to find any systematic pattern. Also, the differences between the model where we restrict the proportions of the types to be the same for all rounds with the model where we allow the proportions of the types to differ across rounds are significant at the 98% confidence level according to the likelihood ratio test.

In order to gain a better understanding of the performance of the CH model, we performed two further exercises. First, we classified each subject into a level and compared the theoretical bidding functions obtained for each level with the data and the best cubic fit.[21] Figure 7 shows the graphs of the most interesting types, RL_2 and TL_1, where the CH theory predicts bidding functions that are not strictly monotonic (graphs of the other types are available upon request). Even though the data roughly follows the patterns of the CH predictions, none of the actual bidding functions exhibit the flat portions. Also and except for TL_1 in round 1, the bidding functions are more reactive to private information than what the CH model predicts.

Second, it is also instructive to compare the results obtained for the CH model with the cluster analysis conducted in section 5. That analysis showed that our sample can be divided in several groups with typical bidding attitudes. Given this heterogeneity, the CH model offers an adequate structure to explain the data. Note in particular that the classification of subjects according to levels is nothing but a specific clustering exercise in which levels are exogenously fixed by the CH model. If the CH model predicts bidding behavior accurately, levels should coincide to a large degree with our (endogenously determined) clusters. Table 7 shows how subjects in our six clusters are classified across levels. Notice that subjects in a cluster are mapped almost exclusively into one of the CH types, except for cluster 1 which is mapped into two types. This is an encouraging result for the CH model. However, the converse is not true: subjects of a given type (most notably TL_0 and TL_1) are found in more than one cluster. In particular, subjects in clusters 1 and 2 classified as TL_1 are pooled with subjects in cluster 5. This indicates the CH model does not make a distinction between the two drastically different patterns of behavior evidenced in those clusters. It captures the 'closeness' to NE but it cannot explain the attitude vis-à-vis each type of information across rounds. As for the rest, subjects in cluster 3 are classified as RL_2. This type tends to underbid and captures to some extent the systematic underbidding of that cluster. It is also not surprising that subjects in cluster 4 are classified as TL_0. These subjects offer the paradigmatic example of the tendency to overbid in common value

[21] We classified each subject to the CH type that minimizes the quadratic distance between the subject's bids and the CH type's predictions. We considered the six functions corresponding to each type of the CH model (RL_k and TL_k with $k \in \{0, 1, 2\}$).

auctions. Subjects in cluster 6 are classified as RL_1 because they exhibit extreme underbidding behavior. Finally, note that we do not find subjects classified as RL_0 in our sample.

A closer inspection of the classification of clusters into levels suggests the following. Subjects in clusters 4 and 6, the most naive bidders in our sample, belong to the lowest levels of the CH model with positive representation: TL_0 and RL_1. An intuitive improvement over the CH model considered previously would be to assume that clusters 4 and 6 are the actual lowest levels of the CH model. The case is easy to make for cluster 4 which is already classified as TL_0. As for subjects in cluster 6, their extreme underbidding strategy is consistent with bidding randomly on the interval $[0, 150]$. Let us denote this type by random* level-0 (henceforth, R^*L_0). Because of the properties of the uniform distribution, this new definition of RL_0 does not change the optimal bidding function of the next levels. The results of this exercise are presented in Table 8. With this alternative hierarchy of random players, the proportion of R^*L_0 increases, the proportions of RL_1 and RL_2 decrease and the proportions of the truthful types do not change. In essence, the subjects in cluster 6 who were "wrongly" classified as either RL_1 or RL_2 are now classified as R^*L_0. The fit of the model is improved both under the AIC and the BIC criteria. However, according to the likelihood ratio test, the differences in the estimated parameters between the constrained and unconstrained models are still significant at the 98% level.

Summing up, the CH model explains to some extent how different levels treat different types of information in a different way, and generates a behavior that is roughly consistent with the observed bidding. However, the model cannot explain the different treatment of the different types of information across rounds. For instance, the behavior of cluster 5 cannot be accounted for by the CH model. Moreover, the model predicts that some bidding functions will not be strictly monotonic, a feature that is not present in our data. The result is summarized as follows.

Result 6. *The CH model captures: (i) the dispersion of the empirical distribution of bids, (ii) the heterogeneity of behavior and the mapping between types and clusters, and (iii) to some extent, the departure from NE regarding the three types of information.*

The CH model does not capture: (i) the peak of the empirical distribution of bids, and (ii) the strongly monotonic bidding functions.

Taken together, Results 5 and 6 suggest that each model captures some interesting aspects of the behavior of subjects. However, they both exhibit limitations. First and foremost, both models show inconsistencies across rounds. If we believe the role of a behavioral model is to deliver better predictions than NE theory, it must be consistent over a range of environments. More modestly, if the same subjects play in all rounds, we would like to obtain similar estimations of the parameters across rounds. None of the models has this property in our experiment. Second, each model 'fixes' a specific observed departures from NE theory. The CE model generates a bias with respect to private information only. As such, it cannot explain a biased attitude towards public information or common uncertainty. The CH model generates heterogeneous behavior. However, it is very sensitive to the specification of level-0 behavior. A single hierarchy of levels (either random or truthful) would not recover much of the observed behavior in our sample. Combining both hierarchies helps capture important aspects of the bidding patterns. Interestingly, the cluster analysis allows us to refine the behavior of the lowest levels. It indicates that all subjects think about the game, and that the most naive players are of two types: subjects who disregard informational effects and consequently overbid, and subjects who choose randomly over 'prudent' bids. Using revised hierarchies based on these premises helps capturing better the data. We summarize the main limitation of the models in the following result.

Result 7. *The estimated parameters in the behavioral models (cursedness in CE and proportion of hierarchy types in CH) are significantly different across rounds. The cluster analysis allows us to refine the behavior of lowest level random players in CH.*

7 Conclusions

This paper incorporates two novel features that facilitate the study of bidding behavior in common value auctions. First, it divides the goods into three (additively separable) elements: those known by both bidders, those known by one bidder, and those known by no bidder. Second, it varies the relative importance of each element holding everything else constant.

The paper replicates the overbidding tendency highlighted in previous research (Kagel and Levin (1986, 2008)) as well as the ability of the Cognitive Hierarchy model to capture reasonably well the behavioral heterogeneity of individuals (Crawford and Iriberri, 2007). More importantly, our two variants generate several new results. First, the regression analysis suggests that subjects depart from NE not only with respect to their treatment of private information but also with respect to their treatment of public information and common uncertainty. In particular, bidders on average react to public information less than predicted by theory. Second, the cluster analysis suggests that part of the heterogeneity across individuals is due to how different subjects treat the different information components. So, for example, clusters 1 and 4 react as Nash players with respect to private information and overreact with respect to common uncertainty. However, the former cluster underreacts with respect to public information whereas the latter overreacts. Third, although CE and CH can account for some interesting features of the data, these two models exhibit an important limitation. Indeed, the estimated parameters in both models change significantly (and rather unpredictably for CH) over rounds even though the same set of individuals plays in all of them.[22] Finally, interesting insights can be gained from the combination of cluster analysis and CH estimation. Quite remarkably, in five of the six clusters, most subjects within a cluster are mapped into the same level of the CH model, despite the fact that clustering is made on an orthogonal dimension. Cluster behavior can also be used to refine the definition of the anchoring type.

We conclude with two general comments. First, the auction studied in this paper has a signal extraction problem which is similar to other games with common values and private information: informational cascades, information aggregation through voting, and jury verdicts just to name a few. Our paper suggests that for these type of settings, one should pay close attention to the way in which information is presented. Indeed, different models may have qualitatively the same theoretical predictions. In practice, however, the behavior of subjects may be affected by the presence and relative importance of other elements such as public information and common uncertainty. Second, performing *both* a cluster analysis *and* a structural estimation of a theory based on heterogeneous types can improve our understanding of the behavior of subjects in complex games. The paper provides a first example of what this joint methodology has to offer but the results suggest that more work in this direction is still needed.

References

[1] Andreoni, J., Che, Y. and Kim, J. (2007). "Asymmetric Information About Rival's Types in Standard Auctions: An Experiment," *Games and Economic Behavior*, 59, 240-259.

[2] Avery, C. and Kagel, J. H. (1997). "Second-Price Auctions With Asymmetric Payoffs: An Experimental Investigation," *Journal of Economics & Management Strategy*, 6 (3), 573-603.

[3] Brocas, I., Carrillo, J., Wang, S. and Camerer, C. (2009). "Measuring Attention and Strategic Behavior in Games with Private Information," *Working Paper*.

[4] Camerer, C. and Ho, T. (1999). "Experience-weighted Attraction Learning in Normal-Form Games," *Econometrica*, 67 (4), 827-874.

[22] A different limitation has been provided in two elegant experiments by Charness and Levin (2009) and Ivanov et al. (2010). The papers show that subjects in the takeover game and second price common value auctions fall for the winner's curse even when the design explicitly rules out incorrect beliefs (the basis of CH and CE) as a potential explanation.

[5] Camerer, C., Ho, T. and Chong, J. (2004). "A Cognitive Hierarchy Model of Behavior in Games," *Quarterly Journal of Economics*, 119, 861-98.

[6] Carrillo, J.D. and Palfrey, T. (2009). "The Compromise Game: Two-sided Adverse Selection in the Laboratory," *American Economic Journal: Microeconomics*, 1(1), 151-181.

[7] Charness, G. and Levin, D. (2009). "The Origin of the Winner's Curse: A Laboratory Study," *American Economic Journal: Microeconomics*, 1(1), 20736.

[8] Costa-Gomes, M., Crawford, V. P. and Broseta, B. (2001). "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 69 (5), 1193-1235.

[9] Crawford, V. P. and Iriberri, N. (2007). "Level-k Auctions: Can a Non-equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," *Econometrica*, 75 (6), 1721-1770.

[10] Esponda, I. (2008). "Information Feedback in First Price Auctions," *RAND Journal of Economics*, 39 (2), 491-508.

[11] Eyster, E. and Rabin, M. (2005). "Cursed Equilibrium," *Econometrica*, 73 (5), 1623-1672.

[12] Fraley, C. and Raftery, A. E. (2002). "Model-based clustering, discriminant analysis, and density estimation," *Journal of the American Statistical Association*, 97, 611-631.

[13] Ivanov, A., Levin, D. and M. Niederle (2010) "Can Relaxation of Beliefs Rationalize the Winner's Curse?: An Experimental Study" forthcoming in *Econometrica*.

[14] Kagel, J. H. and Levin, D. (1986). "The Winner's Curse and Public Information in Common Value Auctions," *The American Economic Review*, 76 (5), 894-920.

[15] Kagel, J. H. and Levin, D. (2008). "Auctions: A Survey of Experimental Research, 1995-2008," *The Handbook of Experimental Economics*, Volume II, Kagel, J. H. and Roth, A. E. (eds). Princeton University Press.

[16] Klemperer, P. (1998). "Auctions with Almost Common Values," *European Economic Review*, 42, 757-769.

[17] Mares, V. and Shor, M. (2008). "Industry Concentration in Common Value Auctions: Theory and Evidence," *Economic Theory*, 35, 37-56.

[18] McKelvey, R. and Palfrey, T. (1995). "Quantal Response Equilibria in Normal Form Games," *Games and Economic Behavior*, 10, 6-38.

[19] Milgrom, P. R. and Weber, R. J. (1982). "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50 (5), 1089-1122.

A Appendix

A.1 Proofs

Proof of Proposition 1.

We restrict the attention to monotonic bidding strategies that are differentiable. Assume that bidder B bids according to such a function and denote it by $b^r(X_B)$.

Let $r < N/2$. The expected utility of bidder A when he bids b^r_A is $U^r_A = \Pr(b^r_A \geq b^r(X^r_B))\left(X^r_A + E[X^r_\emptyset] + E\left[X^r_B | b^r_A \geq b^r(X^r_B)\right] - b^r_A\right)$. Using the distribution of X^r_A, it can be rewritten as:

$$U^r_A = \left(X^r_A + E[X^r_\emptyset] - b^r_A\right) F^r(b^{r^{-1}}(b^r_A)) + \int_{\underline{X}^r}^{b^{r^{-1}}(b^r_A)} X^r_B f^r(X^r_B) dX^r_B \tag{1}$$

Maximizing U_A with respect to b^r_A and imposing the symmetry condition $b^r_A = b^r$, we get the following first-order condition:

$$\left(2X^r_A + E[X^r_\emptyset]\right) f^r(X^r_A) = F^r(X^r_A) b^{r'}(X^r_A) + b^r(X^r_A) f^r(X^r_A)$$

Integrating both sides yields the result. The ex-ante expected bid when $r < N/2$ is

$$E[b^r] = \int_{\underline{X}^r}^{\overline{X}^r} \left(E[X^r_\emptyset] + 2\left(X^r_A - \frac{\int_{\underline{X}^r}^{X^r_A} F^r(s)ds}{F^r(X^r_A)} \right) \right) f^r(X^r_A) dX^r_A$$

Integrating the last term by parts, we get:

$$\begin{aligned} E[b^r] &= E[X^r_\emptyset] + 2E[X^r_A] + 2\int_{\underline{X}^r}^{\overline{X}^r} \log(F^r(X^r_A)) F^r(X^r_A) dX^r_A \\ \iff E[b^r] &= E[V] + 2\int_{\underline{X}^r}^{\overline{X}^r} \log(F^r(X^r_A)) F^r(X^r_A) dX^r_A \end{aligned}$$

The ex-ante expected utility of bidder A is $E[U^r_A] = \int_{\underline{X}^r}^{\overline{X}^r} U^r_A(X^r_A) f^r(X^r_A) dX^r_A$ and at equilibrium $U^r_A(X^r_A) = \int_{\underline{X}^r}^{X^r_A} F^r(s)ds$. Integrating by parts, we get

$$E[U^r_A] = \int_{\underline{X}^r}^{\overline{X}^r} F^r(X^r_A)\left(1 - F^r(X^r_A)\right) dX^r_A$$

When $r = N/2$ there is no common uncertainty so we just need to remove the term $E[X^r_\emptyset]$. When $r > N/2$, We have a family of functions parameterized by X^r_{Pub}. Therefore, we can substitute $E[X^r_\emptyset]$ by X^r_{Pub} and follow the exact same procedure to get the result. □

Proof of Proposition 2.

From the definitions of $b^r(X^r_A)$ and $b^{CE,r}(X^r_A)$, we have:

$$b^{CE,r}(X^r_A) - b^r(X^r_A) = \chi \left[E[X^r_B] - X^r_A + \frac{\int_{\underline{X}^r}^{X^r_A} F^r(s)ds}{F^r(X^r_A)} \right]$$

When $X^r_A = \underline{X}^r$:

$$b^{CE,r}(\underline{X}^r) - b^r(\underline{X}^r) = \chi \left[E[X_B] - \underline{X}^r \right] > 0$$

When $X^r_A = \overline{X}^r$:

$$\begin{aligned} b^{CE,r}(\overline{X}^r) - b^r(\overline{X}^r) &= \chi \left[\int_{\underline{X}^r}^{\overline{X}^r} X^r_B f^r(X^r_B) dX^r_B - \overline{X}^r + \int_{\underline{X}^r}^{\overline{X}^r} F^r(s)ds \right] \\ &= \chi \left[[X^r_B F^r(X^r_B)]_{\underline{X}^r}^{\overline{X}^r} - \int_{\underline{X}^r}^{\overline{X}^r} F^r(X^r_B) dX^r_B - \overline{X}^r + \int_{\underline{X}^r}^{\overline{X}^r} F^r(s)ds \right] = 0 \end{aligned}$$

Finally notice that:

$$\frac{\partial (b^{CE,r}(X^r_A) - b^r(X^r_A))}{\partial X^r_A} = \chi \frac{f^r(X^r_A) \int_{\underline{X}^r}^{X^r_A} F^r(s)ds}{F^r(X^r_A)^2} \geq 0 \quad \forall X^r_A \in [\underline{X}^r, \overline{X}^r],$$

and the proposition follows. □

A.2 Quantal response equilibrium

QRE is an equilibrium model with noisy best responses. The model is characterized by a parameter λ that measures the precision of the best responses. When $\lambda = 0$, agents behave randomly, choosing each strategy with equal probability. When $\lambda \to \infty$, agents behave as in NE. In our model the strategy space is the set of bids, $b \in [N\underline{x}, N\overline{x}]$. Since agents have private information about the value of the item, we have to compute the probability of choosing each bid conditional on the level of private information. Formally, assuming only integer bids can be chosen:

$$\Pr(b^r_A \mid X^r_A) = \frac{\exp(\lambda U(b^r_A \mid X^r_A))}{\sum_{b=N\underline{x}}^{N\overline{x}} \exp(\lambda U(b \mid X^r_A))} \quad \forall b^r_A \in \{N\underline{x}, ..., N\overline{x}\}, X^r_A \in [\underline{X}^r, \overline{X}^r]$$

When $r \leq N/2$, we have:

$$\begin{aligned} U(b^r_A \mid X^r_A) \quad = \quad & \int_{\underline{X}_B}^{\overline{X}^r_B} \left(\frac{1}{2} \Pr(b^r_A \mid X^r_B) + \sum_{i=N\underline{x}}^{b^r_A - 1} \Pr(b = i \mid X^r_B) \right) \\ & \times \left(X^r_A + E[X^r_\emptyset] + X^r_B - b^r_A \right) f^r(X^r_B) dX^r_B \end{aligned}$$

When $r = N/2$ we simply need to remove $E[X^r_\emptyset]$. When $r > N/2$, we replace $E[X^r_\emptyset]$ by X^r_{Pub} and $F^r(\cdot)$ by $F^{N-r}(\cdot)$. The term $P(b^r_A \mid X^r_B)/2$ appears because it is assumed that each bidder wins with probability 1/2 in case of a tie.

To understand bidding QRE, we need to consider three factors. First, for low bids the payoff is bounded below at zero (losing the item), while for high bids the payoff may be negative (paying more than the value of the item). Therefore, lower bids are more likely than higher bids. This effect makes QRE predict underbidding relative to NE. Second, if other agents are underbidding, it is possible to win the item at an even lower price, and there are incentives to reduce bids even further. Third, in a common value auction the signals of all bidders affect the value of the item and therefore, directly affect the utility of all bidders. If an agent anticipates that his rival is underbidding, he has incentives to increase his own bid. This is true because the anticipated low bid of his rival is still associated with a high signal. Therefore, the value of the item is higher, and it is worth to bid more for it. Which effect dominates depends on the distribution of the private signals. In our specification of the model, the combined effect is that QRE predicts underbidding relative to NE.

A.3 Figures

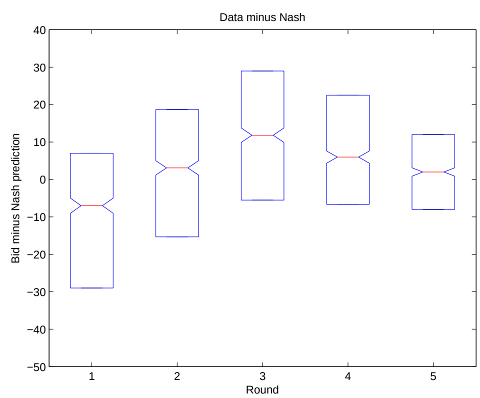

Figure 1: Deviations from NE

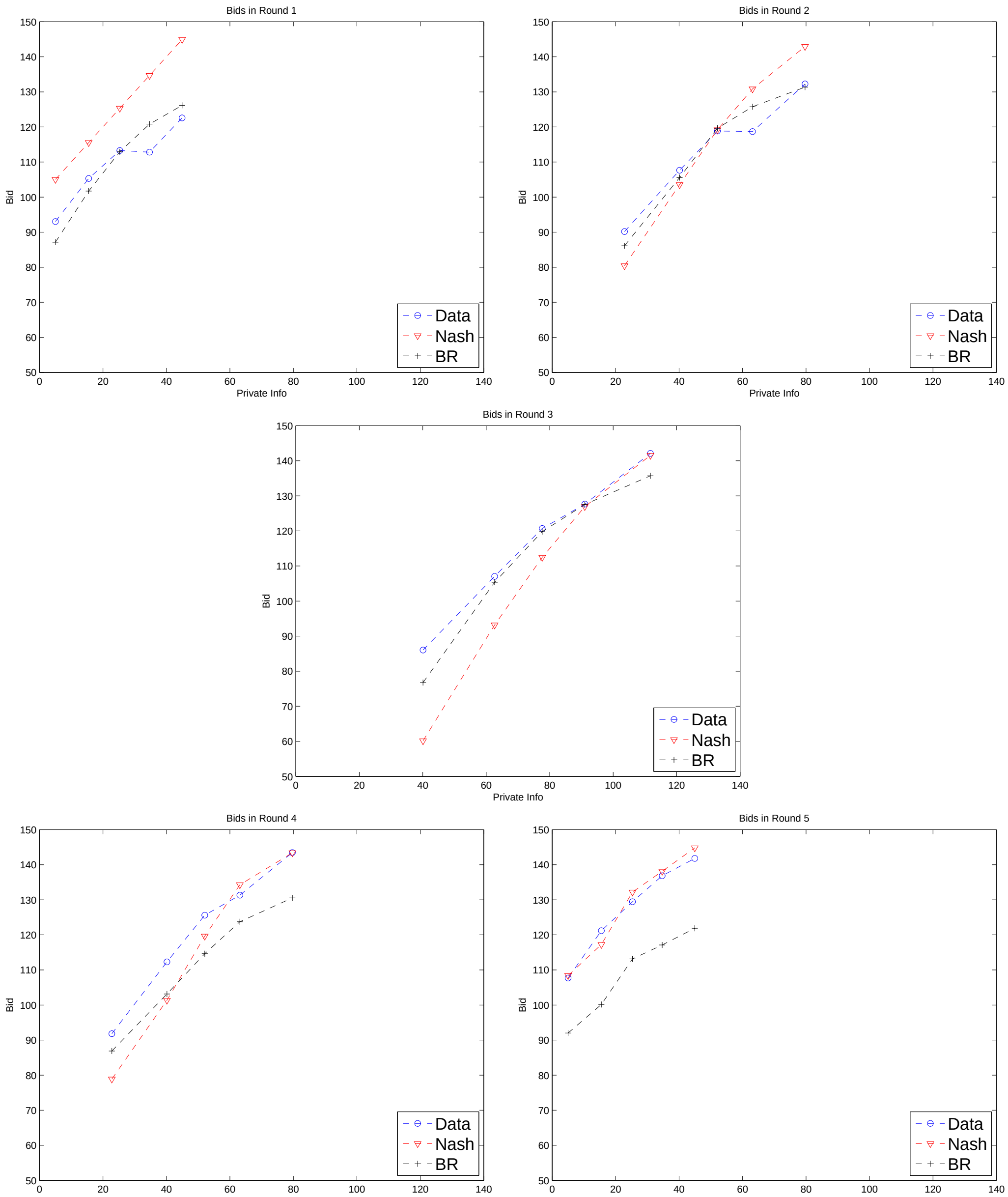

Figure 2: Average bids as a function of private information

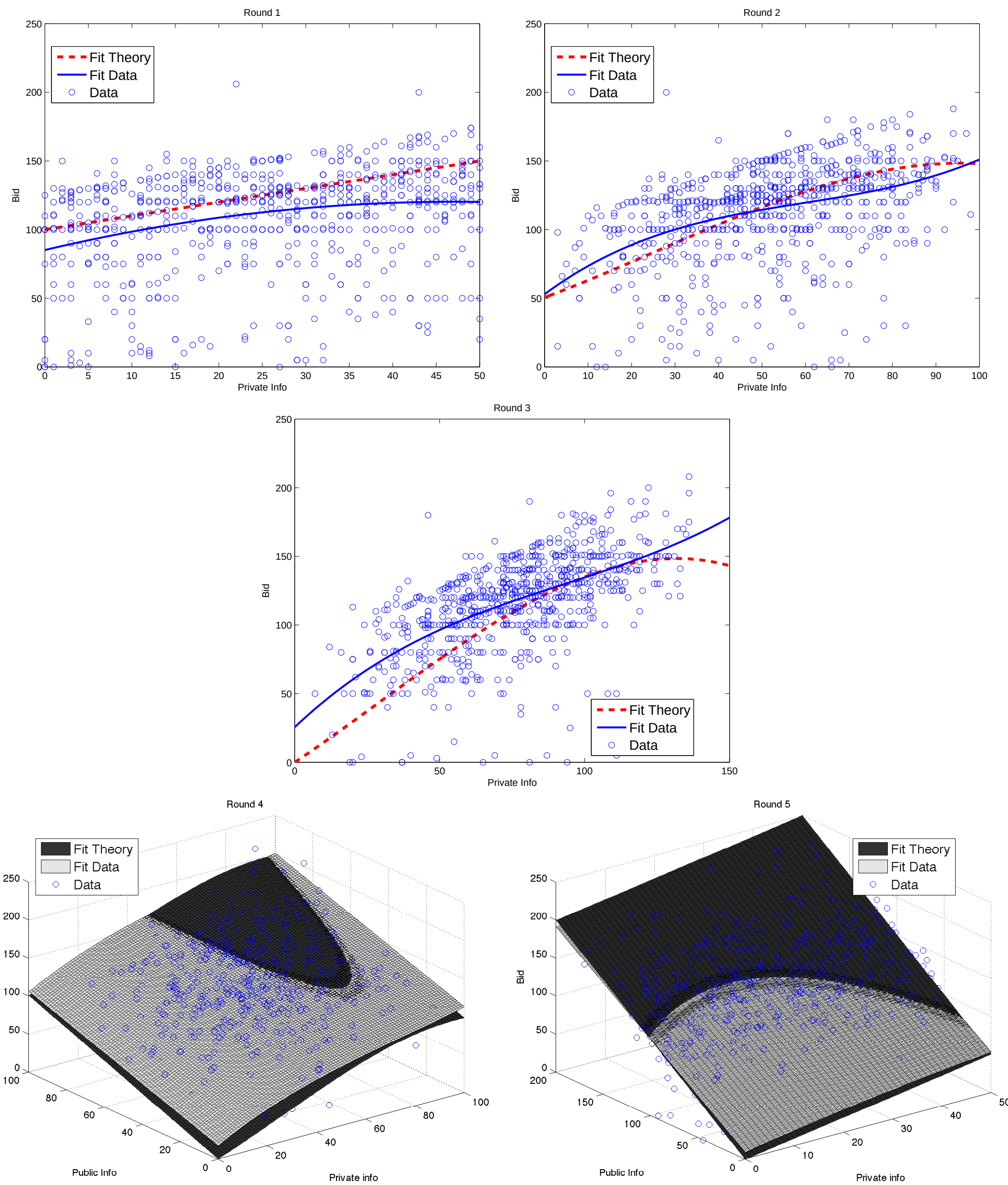

Figure 3: NE and FGLS regression

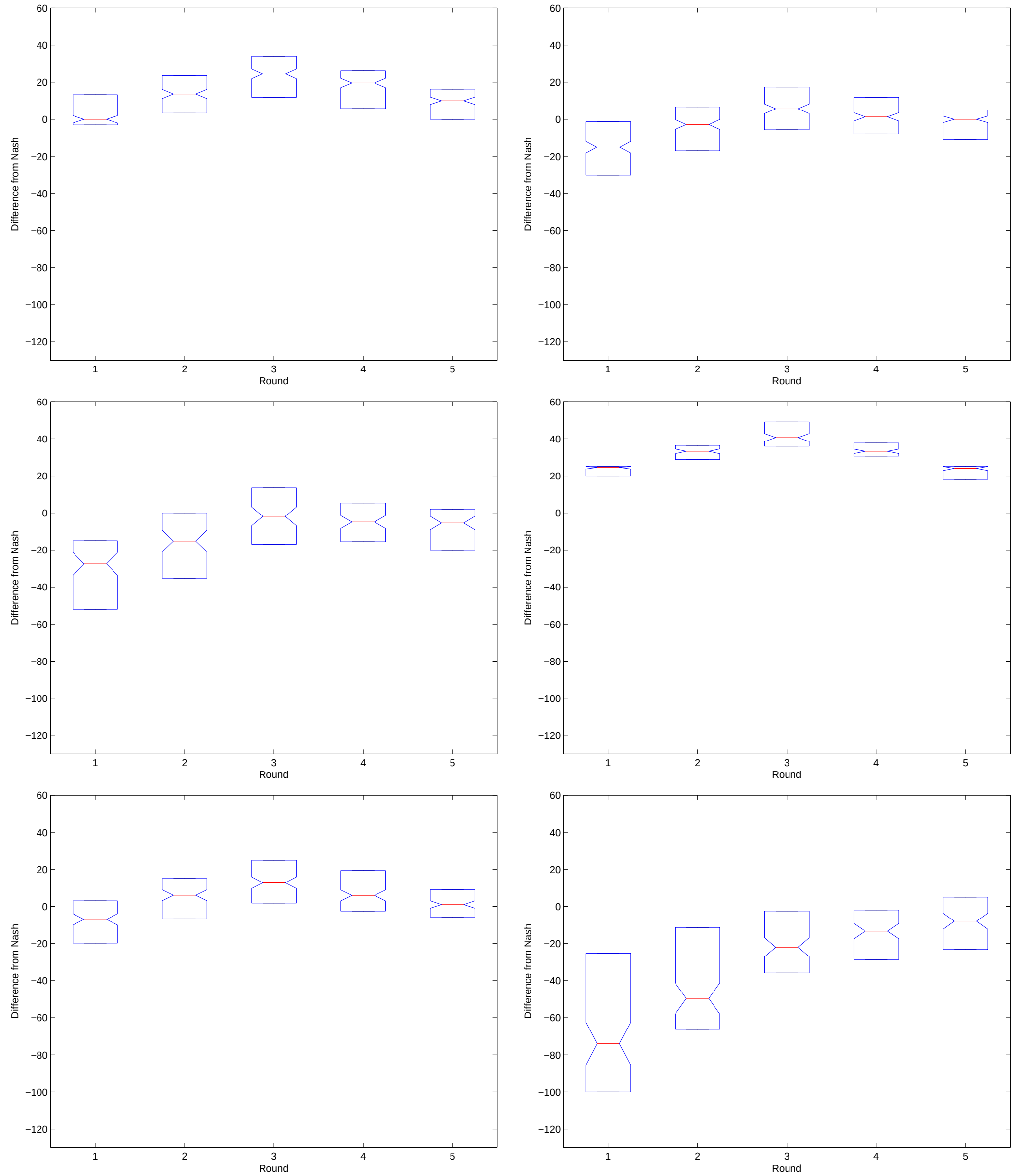

Figure 4: Deviations from NE per cluster

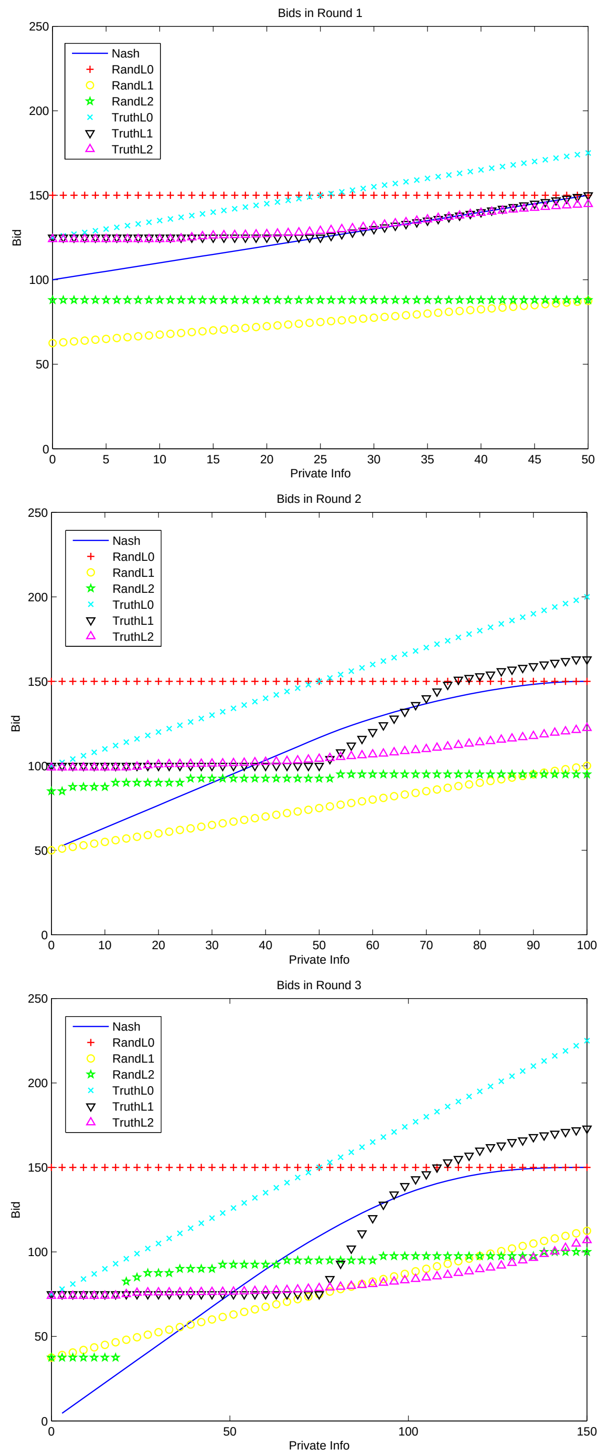

Figure 5: Level-k bids per round

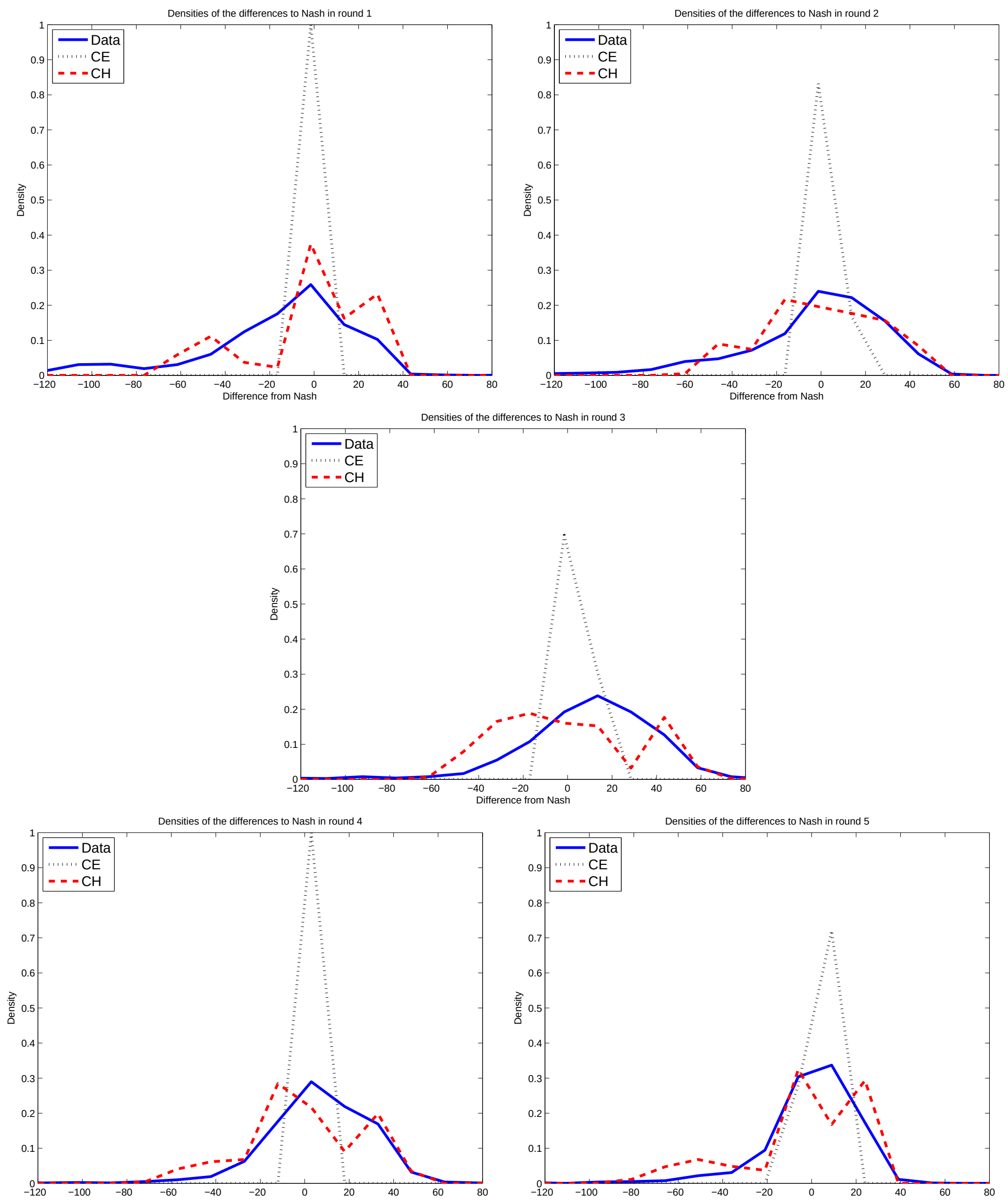

Figure 6: Densities of deviations from NE per round

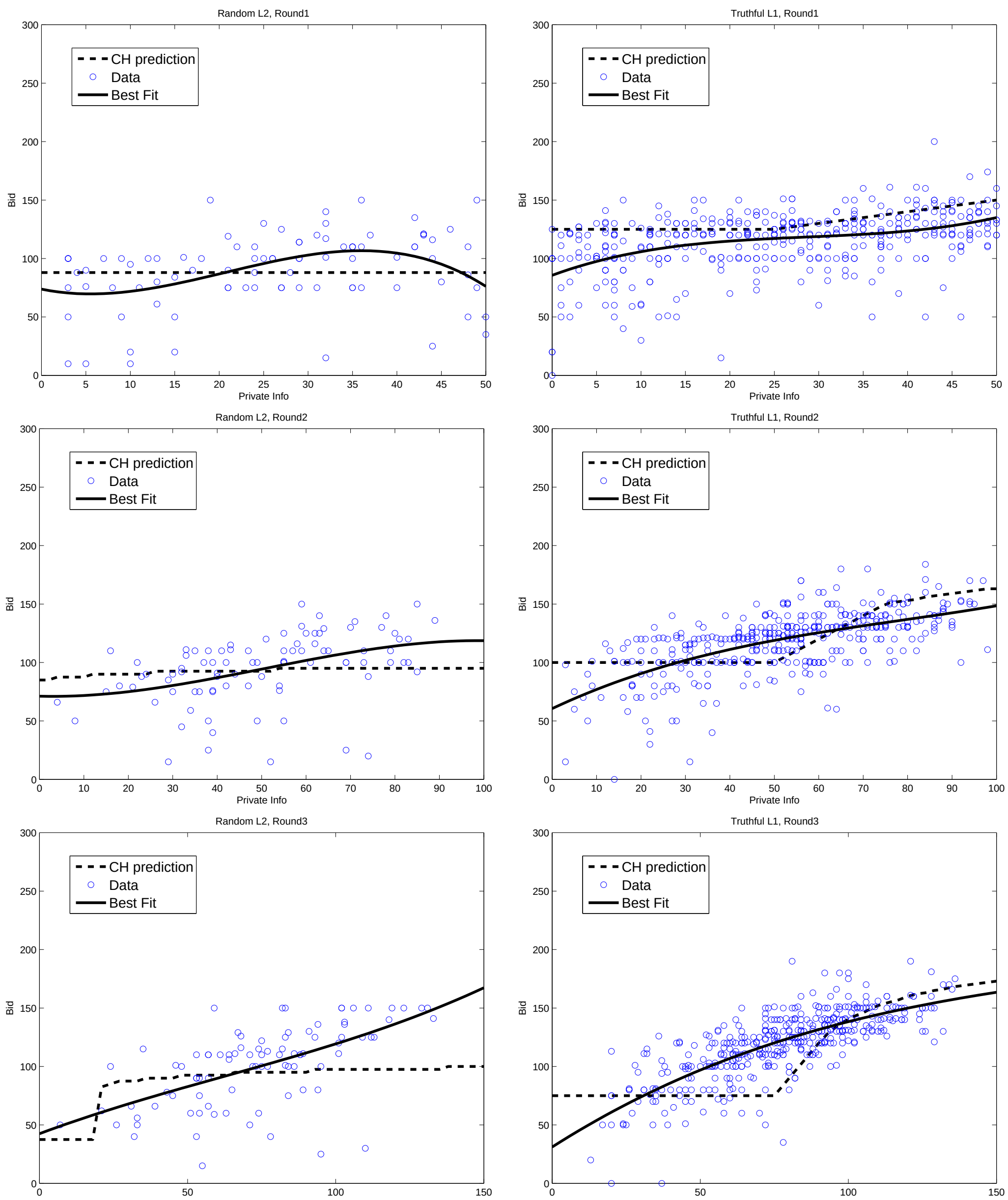

Figure 7: Bidding behavior of RL2 and TL1: theory and data

A.4 Tables

Table 1: Average bids

Round	1	2	3	4	5
Mean Data	109.41	113.52	116.70	120.90	127.40
	(1.28)	(1.22)	(1.22)	(1.25)	(1.40)
Mean NE	125.10***	115.39	106.85***	115.51***	128.16
	(0.51)	(0.81)	(1.06)	(1.12)	(1.20)
Mean BR	109.75	113.69	113.03***	111.79***	108.86***
	(0.51)	(0.61)	(0.78)	(0.80)	(0.98)

Standard errors in parenthesis

*, **, ***: Significantly different from data at 90%, 95% and 98% confidence level

Table 2: Average gains

Round	1	2	3	4	5
Mean Data	12.33	10.69	9.19	8.38	6.27
	(1.09)	(0.96)	(0.84)	(0.67)	(0.50)
Mean NE	10.03**	12.52*	14.95***	12.46***	8.50***
	(0.90)	(0.82)	(0.77)	(0.61)	(0.43)
Mean BR	17.40***	15.87***	12.76***	10.82***	6.04
	(1.10)	(0.96)	(0.81)	(0.69)	(0.51)

Standard errors in parenthesis

*, **, ***: Significantly different from data at 90%, 95% and 98% confidence level

Table 3: NE and FGLS regression per round

	Round 1		Round 2		Round 3		Round 4		Round 5	
	NE	Data	NE	Data	NE	Data	NE	Data	NE	Data
Priv	1	1.49***	1.18	2.34***	1.42	1.99***	1.18	1.80***	1	1.17***
		(0.0097)		(0.0188)		(0.0167)		(0.0127)		(0.0066)
$(\text{Priv})^2$	0	−0.016***	0.007	−0.031***	0.004	−0.014***	0.007	−0.016***	0	−0.006***
		(0.0002)		(0.0004)		(0.0002)		(0.0003)		(0.0001)
$(\text{Priv})^3$	NA	NA	-0.0001	0.0002***	-0.0001	0.0001***	-0.0001	0.0001***	NA	NA
				(0)		(0)		(0)		
Pub	NA	NA	NA	NA	NA	NA	1	0.8814***	1	0.9102***
								(0.0013)		(0.0009)
Constant	100	85.17***	53.00†	56.04***	−0.05†	25.59***	0.59†	17.60***	0	9.17***
		(0.1843)		(0.2922)		(0.3951)		(0.2090)		(0.1196)
F-test		59.6465***		15.3015***		50.9397***		18.8071***		2.1902
Adjusted R^2		0.075		0.181		0.349		0.519		0.634

Standard errors in parenthesis

*, **, ***: Significantly different from NE at 90%, 95% and 98% confidence level

†: These values are not 50, 0 and 0 respectively because we are using a polynomial approximation of NE

Table 4: NE and FGLS regression per round and cluster

	Nash	Cluster					
		1	2	3	4	5	6
Round 1							
Constant	100	110.32^{***}	91.06^{***}	84.37^{***}	127.29^{***}	89.36^{***}	43.56^{***}
		(0.31)	(0.36)	(0.55)	(0.20)	(0.76)	(0.72)
Priv	1	0.64^{***}	0.65^{***}	0.37^{***}	0.73^{***}	1.05^{**}	0.70^{***}
		(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
F-test		8.53^{***}	60.92^{***}	54.57^{***}	417.02^{***}	12.82^{***}	122.84^{***}
Round 2							
Constant	53	80.12^{***}	56.70^{***}	54.82	96.27^{***}	34.62^{***}	17.17^{***}
		(0.57)	(0.82)	(1.18)	(0.33)	(1.60)	(2.16)
Priv	1.18	1.55^{***}	1.41^{***}	2.11^{***}	1.04^{***}	3.47^{***}	2.91^{***}
		(0.04)	(0.06)	(0.15)	(0.02)	(0.07)	(0.18)
$(\text{Priv})^2$	0.007	-0.015^{***}	-0.006^{***}	-0.039^{***}	0.000^{***}	-0.045^{***}	-0.057^{***}
		(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.004)
$(\text{Priv})^3$	-0.0001	0.0000^{***}	0.0000^{***}	0.0003^{***}	0.0000^{***}	0.0002^{***}	0.0004^{***}
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
F-test		62.30^{***}	8.07^{***}	14.27^{***}	497.74^{***}	4.65^{***}	36.68^{***}
Round 3							
Constant	-0.05	61.45^{***}	46.87^{***}	34.70^{***}	73.37^{***}	-12.12^{***}	-67.12^{***}
		(1.92)	(1.22)	(1.61)	(0.49)	(1.09)	(2.96)
Priv	1.42	1.20^{***}	0.74^{***}	1.87^{***}	1.14^{***}	3.11^{***}	4.54^{***}
		(0.08)	(0.06)	(0.11)	(0.03)	(0.06)	(0.17)
$(\text{Priv})^2$	0.004	-0.006^{***}	0.004	-0.020^{***}	-0.005^{***}	-0.021^{***}	-0.047^{***}
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)
$(\text{Priv})^3$	-0.0001	0.0000^{***}	0.0000^{***}	0.0000^{***}	0.0000^{***}	0.0000^{***}	0.0002^{***}
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
F-test		743.91^{***}	302.07^{***}	42.26^{***}	1249.15^{***}	235.80^{***}	10.36^{***}
Round 4							
Constant	0.59	28.35^{***}	11.05^{***}	21.96^{***}	36.66^{***}	19.56^{***}	-29.08^{***}
		(0.71)	(0.75)	(2.16)	(1.78)	(0.65)	(2.64)
Priv	1.18	2.32^{***}	1.87^{***}	0.85^{*}	1.70^{***}	0.97^{***}	3.72^{***}
		(0.05)	(0.06)	(0.19)	(0.12)	(0.05)	(0.16)
$(\text{Priv})^2$	0.007	-0.030^{***}	-0.017^{***}	0.003	-0.015^{***}	0.006	-0.054^{***}
		(0.001)	(0.001)	(0.005)	(0.002)	(0.001)	(0.003)
$(\text{Priv})^3$	-0.0001	0.0002^{***}	0.0000^{***}	0.0000	0.0001^{***}	0.0000^{***}	0.0003^{***}
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Pub	1	0.83^{***}	0.92^{***}	0.83^{***}	0.99^{***}	0.99	0.76^{***}
		(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
F-test		94.24^{***}	5.05^{***}	3.11^{*}	272.42^{***}	9.58^{***}	6.77^{***}
Round 5							
Constant	0	20.93^{***}	11.79^{***}	0.09	15.10^{***}	0.61	-1.93
		(0.29)	(0.32)	(0.54)	(0.44)	(0.50)	(1.67)
Priv	1	0.81^{***}	0.84^{***}	0.72^{***}	1.05^{***}	0.91^{***}	1.17^{***}
		(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
Pub	1	0.92^{***}	0.89^{***}	0.93^{***}	1.05^{***}	1.02^{***}	0.80^{***}
		(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)
F-test		37.94^{***}	7.25^{***}	8.53^{***}	237.38^{***}	0.43	8.65^{***}
N		11	13	6	6	9	7

Standard errors in parenthesis

*, **, ***: Significantly different from NE at 90%, 95% and 98% confidence level

Table 5: Average gains by cluster

Cluster	1	2	3	4	5	6
Mean	9.56	10.12	10.46	4.73	12.31	6.97
Standard Deviation	5.39	6.38	5.55	4.57	6.35	5.77

Table 6: Normal Estimation of CE and CH models

		All rounds	Round 1	Round 2	Round 3	Round 4	Round 5
Cursed Equilibrium	χ	0.21	0.00^{**}	0.14^{***}	0.52^{***}	0.37^{***}	0.01
	AIC	37633	37498				
	BIC	37639	37529				
	LL	$18816^{\dagger\dagger\dagger}$	18744				
Cognitive Hierarchy	RL_0	0.00	0.00	0.00	0.00	0.00	0.00
	RL_1	0.08	0.14	0.10	0.04	0.00	0.00
	RL_2	0.12	0.24	0.00	0.12	0.06	0.09
	TL_0	0.20	0.00	0.13	0.40	0.20	0.05
	TL_1	0.55	0.00	0.45	0.43	0.61	0.00
	TL_2	0.05	0.63	0.32	0.01	0.13	0.86
	AIC	31345	31237				
	BIC	31387	31450				
	LL	$15665^{\dagger\dagger\dagger}$	15583				

*,**,***: Significantly different from the next round at 90%, 95% and 98% confidence level.

†††: Significantly different from the unconstrained model at 98% confidence level.

Table 7: Classification of subjects by hierarchy type and cluster

		Cluster						
		1	2	3	4	5	6	Total
CH	TL_0	4			6			10
	TL_1	7	12			9		28
	TL_2		1	1				2
	RL_0							0
	RL_1						6	6
	RL_2			5			1	6
	Total	11	13	6	6	9	7	52

Table 8: Normal Estimation of CH with R^*L_0

	All rounds	Round 1	Round 2	Round 3	Round 4	Round 5
R^*L_0	0.05	0.09	0.05	0.04	0.04	0.04
RL_1	0.06	0.04	0.06	0.03	0.00	0.00
RL_2	0.10	0.24	0.00	0.10	0.01	0.06
TL_0	0.20	0.00	0.13	0.40	0.20	0.05
TL_1	0.55	0.00	0.45	0.43	0.62	0.00
TL_2	0.05	0.62	0.32	0.00	0.13	0.85
AIC	31239	31132				
BIC	31282	31346				
LL	$15613^{\dagger\dagger\dagger}$	15531				

†††: Significantly different from the unconstrained model at 98% confidence level.