ソシオネットワーク戦略ディスカッションペーパーシリーズ ISSN 2434-9445
第 116 号　2024 年 11 月
RISS Discussion Paper Series
No.116　November, 2024

On the Performance of Crémer–McLean Auction: An Experiment

Takeshi Nishimura, Nobuyuki Hanaki

文部科学大臣認定　共同利用・共同研究拠点

関西大学ソシオネットワーク戦略研究機構

Research Institute for Socionetwork Strategies,
Kansai University
Joint Usage / Research Center, MEXT, Japan
Suita, Osaka, 564-8680, Japan
URL: https://www.kansai-u.ac.jp/riss/index.html
e-mail: riss@ml.kandai.jp
tel. 06-6368-1228
fax. 06-6330-3304

On the Performance of Crémer–McLean Auction: An Experiment*

Takeshi Nishimura† Nobuyuki Hanaki‡

November 12, 2024

Abstract

The paradoxical full-surplus-extraction (FE) result, which can impair the mechanism design paradigm, is a longstanding concern in the literature. We tackle this problem by experimentally testing the performance of an FE auction, which is a second-price (2P) auction with lotteries. In FE treatment, overbid amounts given entry increased and entry rates decreased through rounds, and FE failed. By contrast, a majority of subjects learned value bidding in 2P treatment. To find the causes of failure in FE, we take an evolutionary-game approach. We show that the FE auction with risk-neutral bidders has exactly two symmetric equilibria, value bidding with full or partial entry, and only the partial-entry equilibrium is (evolutionarily or asymptotically) stable. Replicator dynamics with vanishing trends well explain observed dynamic bidding patterns. Our results suggest that the FE outcome is non-robust to bidders' trial-and-error learning.

Keywords: Full surplus extraction, Crémer–McLean auction, auction experiment, evolutionary stability, perturbed replicator dynamics.

*We gratefully acknowledge John Duffy, Yoshio Kamijo, Charles Noussair, Tiffany Tsz Kwan Tse, Nicolaas Vriend, and the participants at RISS Workshop, Waseda Workshop for Experimental Economics, SWET 2024, and 2024 Asia-Pacific ESA Conference for their helpful comments and suggestions. We are extremely grateful to Kazuhito Ogawa and Yuta Shimodaira for their support for running the experiment. The experiment reported in this study was approved by the Institutional Review Boards of the Institute of Social and Economic Research (ISER), Osaka University (No. 20221201) and the Research Institute for Socionetwork Strategies (RISS), Kansai University (No. 2022028). The experiment was pre-registered at AsPredicted.org (#119225, https://aspredicted.org/55gs-7f5r.pdf). We acknowledge the support of JSPS KAKENHI Grant Numbers 18K12747, 18K19954, 20H05631, 23H00051, and 23H00055 along with the Joint Usage/Research Center for Behavioral Economics at the ISER, Osaka University and the Joint Usage/Research Center at the RISS, Kansai University.

†Faculty of Economics, Hosei University.

‡Institute of Social and Economic Research, Osaka University; University of Limassol.

1 Introduction

Full-surplus-extraction (FE) auctions designed by Crémer and McLean (1988, CM) have been regarded as unrealistic. CM proved that these auctions allow a seller of an item to extract full surplus on average from risk-neutral bidders with private information. This FE result holds true if there is a (generically satisfied) statistical dependence among bidders' private values. Yet, Börgers (2015, p. 124) argue that "one should view the Crémer–McLean result as a paradox rather than a guidance to the construction of mechanisms that could work in practice."[1] Moreover, CM themselves concluded that "Economic intuition and informal evidence (we know of no way to test such a proposition) suggest that this result is counterfactual, and several explanations can be suggested." The purpose of this study is to conduct a formal experimental test whether FE auctions perform as predicted by CM and investigate underlying causes if full surplus extraction fails.

Perhaps, the most serious problem about the FE result is that an auction rule depends on the commonly known joint distribution of values. As CM pointed out, the seller should not only know the true distribution but also share this information with bidders in a credible way. An experimental approach, however, removes these hurdles because we can construct a common prior distribution and present it to subjects in a laboratory. Specifically, we consider a two-bidder environment with a positive correlation between their private values. This distribution produces a "surplus-extracting lottery" for each bidder, which is simple in the sense that the prize is monotonically decreasing in the opponent's bid.[2] In the second-price (2P) auction with this lottery, value bidding is a dominant strategy given entry, and risk-neutral bidders always "break even" if both bid their values. The seller can thus extract full surplus in the *full-entry* equilibrium of this FE auction. In our main treatment labelled FE, subjects played this auction game knowing the common prior.[3]

The second problem raised by CM is that the FE result relies on bidders' risk neutrality. For risk-averse bidders in the FE auction, value bidding is still dominant given entry, but the best response to value bidding is opt-out. The full-entry equilibrium is then lost. We address this problem by introducing another

[1]See Milgrom (2004, p. 165) and Carroll (2019, p. 142) for similar arguments.

[2]We used a compound distribution as in Carroll (2019). This specification allows us to characterize the lottery.

[3]Strictly speaking, we added a unit point to the lottery to endow the full-entry equilibrium with strictness.

treatment labelled PE (partial extraction), in which the lottery function is shifted upward compared to FE. In other words, the seller pays "risk premia" for entry.[4] The fixed amount of premia was selected so that the seller's expected revenue in this auction is nearly equal to that in the 2P (or ascending-clock) auction with the optimal reserve price, given value bidding in both auctions.[5] This revenue equivalence implies that if the actual average revenue in PE is lower than the expected revenue, then the PE auction is practically useless for the seller. We also elicited subjects' risk preferences by using multiple price lists (Laury and Holt, 2005) to investigate whether they have effects on entry decisions.

Our between-subjects experiment has three treatments: FE, PE, and 2P. Each treatment has 20 auction rounds. Applying group matching (Tan, 2020), we randomly divided subjects into groups of size 6, with random pairing within a group in each round. 2P is a control treatment in which a pair of subjects plays the 2P auction game with neither lottery nor reserve price. In experiments on 2P auctions, overbidding (i.e., bidding higher than values) has been frequently observed (Kagel and Levin, 2016). This behavior has been explained by, among others, spite motives (Bartling and Netzer, 2016; Cooper and Fang, 2008; Morgan et al., 2003; Nishimura et al., 2011; Tan, 2020) and cognitive limitations (Breitmoser and Schweighofer-Kodritsch, 2022; Li, 2017; Schneider and Porter, 2020). To control these confounding factors as much as possible in all of the three treatments, we applied several methods proposed in the literature on auction experiments.

First, in order to control social preferences, we used the amended random payment scheme developed by Lim and Xiong (2021) in their English-auction experiment on jump bidding. Under this scheme, one payment round is randomly chosen for each subject, who is uninformed about this round until the end of a session. The ingenuity of this scheme is that each subject is informed about whether each round is the current opponent's payment round. Subjects with social preferences should thus care only about their own material payoffs in opponents' *non*-payment rounds. Without any control, spite motives would cause overbidding as in previous experiments, and even worse, spite might be encouraged by the lotteries in FE and PE, where the opponent's prize is decreasing in own bid. Therefore, we adopted this scheme to induce selfish preferences in opponents'

[4]Alternatively, we could adopt a binary lottery procedure (Roth and Malouf, 1979) to induce risk-neutral preferences, but we did not. The reason is that evidence on its performance in auction experiments (Rietz, 1993; Walker et al., 1990) is mixed. Selten et al. (1999) also showed that this procedure was counterproductive in their lottery-choice experiment.

[5]There is sufficient evidence that subjects immediately follow value bidding in ascending-clock auctions with private values. See, e.g., Kagel et al. (1987), Li (2017), and Breitmoser and Schweighofer-Kodritsch (2022).

non-payment rounds. We focus on bidding data in these rounds in statistical analyses unless otherwise stated.

The second method is ascending-clock presentation developed by Breitmoser and Schweighofer-Kodritsch (2022) for alleviation of cognitive problems. Under this feedback method, a pair of subjects watches an "ascending price" after they submit sealed bids. Breitmoser and Schweighofer-Kodritsch (2022) provided experimental evidence that this method greatly reduces the deviations from value bidding in the 2P auction. Moreover, in FE and PE, we cannot resort to a dynamic ascending-clock format as in Li (2017), because the clock stops as soon as a bidder drops out, and hence, the loser's lottery prize is not well-defined. For these reasons, we adopted the static format with ascending-clock presentation, instead of a dynamic one, in all of the three treatments. We also elicited subjects' cognitive abilities by using a cognitive reflection test (Frederick, 2005; Toplak et al., 2014) to investigate their relationships with bidding behavior, following Schneider and Porter (2020).

Our experimental results show that the FE auction did not perform as predicted by the full-entry equilibrium. The seller's average revenues in the experiment are ranked as follows: Full surplus $>$ FE $\approx$ PE $>$ 2P. Full surplus extraction thus failed in FE. In 2P, a majority of subjects bid their values in most rounds, and many others who underbid in early rounds learned value bidding through rounds. This is in stark contrast to bidding behavior in FE and PE, where the average overbid amount conditional on entry increased and the entry rate decreased through rounds.

Where were subjects going in FE and PE? To answer this question, we take an evolutionary-game approach (Fudenberg and Levine, 1998; Weibull, 1997). Since subjects were randomly paired in each group, we analyze symmetric equilibria in the auction game. First, it turns out that the FE auction with risk-neutral bidders has the other symmetric equilibrium called the *partial-entry* equilibrium, in which bidders opt out if values are low, and opt in more often as values are higher, with value bidding given entry. Second, the partial-entry equilibrium is *evolutionarily stable*—whereas the full-entry equilibrium is *asymptotically unstable* in replicator dynamics. Hence, the former is a unique stable state to which interior paths can converge. These replicator dynamics are approximations of reinforcement-learning processes as shown by Börgers and Sarin (1997). The partial-entry equilibrium also changes continuously with bidders' risk preference, which allow participants to learn these equilibria.

Our clustering analysis suggests that, in the second half of FE, two clusters of subjects (42%) exhibited behavior in line with the partial-entry equilibria, while another cluster (30%) bid their values with full entry and two clusters (19%) overbid with full or partial entry. Remarkably, FE had some extreme overbidders, who often bid the maximum amount. In the replicator dynamics, some paths from interior (e.g., uniform) initial states explain the decrease in entry rates in FE. It is, however, difficult to explain the increase in overbids because overbidding is dominated for rational and selfish bidders, and hence, should approach "extinction." The fact that irrational overbidders coexisted with rational value-bidders in later rounds is not consistent with selection pressure.

We approach this problem by incorporating *vanishing trends* toward overbidding and value bidding into the dynamics. The perturbed dynamics well explain dynamic bidding patterns in FE as well as PE and 2P. Overbidding in FE and PE indeed leads to a further decrease in entry rates because this behavior, together with the CM lotteries, deters entry. In 2P, overbidding has no such effect. Given estimated parameters, a predicted long-run state for FE is close to a partial-entry equilibrium for slightly risk-averse bidders. In this equilibrium, the seller earns slightly less than in the 2P auction with the optimal reserve price. This underperformance result provides a clue for why CM auctions are not used in practice, within a conventional game-theory framework.

Related Literature By extending an insightful example of Myerson (1981), Crémer and McLean (1985, 1988) designed FE mechanisms in interdependent- and private-values environments with risk-neutral agents. CM characterized, for each of Bayesian and dominant incentive compatibility, the joint value distributions that guarantee full surplus extraction. In this study, we apply their result for dominant incentive compatibility. Their Bayesian implementation result was extended to a general environment with continuum type spaces by McAfee and Reny (1992),[6] who wrote the following famous passage:

> Although the paper develops tools for solving mechanism design problems with correlated information, the results (full rent extraction) cast doubt on the value of the current mechanism design paradigm as a model of institutional design.

Their pessimistic opinion is supported by the fact that arbitrarily small amounts of correlation result in full rent extraction. Heifetz and Neeman (2006) and Chen

[6]More recently, Mezzetti (2007) proposed two-stage FE mechanisms in environments with independent types but interdependent values. Noda (2019) constructed FE mechanisms in general dynamic environments by using intertemporal correlations of agents' types.

and Xiong (2013a), with different concepts of genericity, then proved nongenericity and genericity results on full surplus extraction in infinite-dimensional spaces of priors (e.g., the collection of all priors on the universal type space), respectively. Although these two studies reached opposite conclusions, they shared the same motivation—the evaluation of the validity of mechanism design paradigm. Chen and Xiong (2013a, p. 826) argue that the classical model as in CM remains subject to the "full-surplus-extraction critique." Our experiment, with only a single joint value distribution, is not designed to empirically study genericity problems. As CM suggested, it may be impossible to test their generic implementation results even experimentally.

To establish genericity, Chen and Xiong (2013a) showed that CM auctions are robust to common priors about types. Pham and Yamashita (2024), however, reveal that CM auctions are non-robust to heterogeneity in priors. As shown by Pham and Yamashita (2024), even slight perturbations of bidders' priors from the true prior about values can drastically change interim beliefs and equilibria in CM auctions, and the equilibrium revenues in CM (and other) auctions are no higher than the optimal dominant-strategy mechanism (e.g., the 2P auction with the optimal reserve price).[7] Our experimental design would ensure that the prior distribution is common knowledge among subjects. In addition, interim beliefs were explicitly displayed to subjects. Our evolutionary analyses with experimental evidence indicate that, even in this clean environment, CM auctions have another non-robustness problem—the instability with respect to bidders' trial-and-error learning.

Although CM recognized the problem of multiple equilibria, they regarded the full-entry value-bidding equilibrium as focal and focused on partial implementation. Yet, the instability of this equilibrium implies that their view might be optimistic. Full implementation of the FE outcome has been studied by Maskin and Riley (1980), Brusco (1998), Matsushima (2007), and Chen and Xiong (2013b). Our experiment has no treatment for their fully-implementing mechanisms, because they are more complicated than CM auctions (or experimentally infeasible). Experiments on these mechanisms are left for future research.

The partial-entry equilibrium revenue in FE treatment is decreasing in the degree of bidders' risk aversion. Relatedly, the effects of risk aversion on rent extraction have been theoretically studied. Robert (1991) proved that if bidders are

[7]See Chung and Ely (2007) for some foundations of dominant-strategy mechanisms. However, these foundations are weakened by Börgers (2017) who constructed a mechanism that "dominates" the optimal dominant-strategy mechanism.

risk averse (or have limited liability), then the seller's optimal revenue is continuous in value distributions, especially at independent distributions. The optimal auction with risk-averse bidders and correlated values was characterized by Eső (2005). The auction rule depends not only on the value distribution but also on bidders' risk preferences. Hence, to experimentally test its performance, one needs to elicit subjects' risk preferences in advance.

Finally, there are a few experimental studies of surplus extraction. Nishimura (2022) experimentally compared two CM auctions: the 2P and 1P (first-price) auctions with CM lotteries. In the 2P-CM auction, overbidding persisted throughout all rounds. However, the overbid amount slightly decreased through rounds in the 1P-CM auction. Although this tendency toward value bidding is of interest, this experiment is simpler than the current one. In particular, no bidder had the opt-out option. For public-goods problems, Krajbich et al. (2009) and Krajbich et al. (2017) experimentally tested the performances of their "neurometrically informed mechanisms." In this novel mechanism, each agent's tax payment depends on an ex-post signal correlated with their payoff type, as in Riordan and Sappington (1988). These non-manipulable signals are interpreted as noisy neural measures of subjects' preferences provided by neurometric technologies.[8] The neurometrically informed mechanisms, which are efficient, dominant incentive compatible, and interim individually rational, performed surprisingly well in their experiments. In contrast with their mechanisms, a bidder's lottery prize in the FE auction depends on the opponent's bid, not on the own neural signal. This additional strategic interaction between bidders might cause the underperformance of the FE auction.

The rest of the paper is organized as follows: Section 2 presents our theoretical framework with predictions. Section 3 demonstrates our experimental design. Section 4 shows the experimental results. Section 5 analyzes the dynamics of bidding strategies. Section 6 discusses some future directions. Several lemmas and proofs are presented in Appendix A. The Supplementary Material (SM) contains some additional information including English translations of instructions for the experiment.

[8]Krajbich et al. (2009) actually used functional magnetic resonance imaging (fMRI) to obtain signals of subjects' induced values, while Krajbich et al. (2017) used computers to simulate noisy signals as if they were generated by a neurometric technology.

2 Theoretical Framework

In this section, we construct an auction model following CM to derive theoretical predictions for our experiment. We provide a partial characterization of the set of symmetric equilibria and refine them applying stability concepts in evolutionary game theory. We also illustrate how stable equilibria in the experimental setting change with bidders' risk attitude.

2.1 Model

There are two bidders who compete for a single item auctioned by a seller. Each bidder $i \in \{1, 2\}$ has a value v_i for the item. The set of values is $V := \{v^1, v^2, ..., v^n\} \subset \mathbb{R}_{++}$ with $n \geq 2$. Bidders have the same von Neumann–Morgenstern utility function $u : \mathbb{R} \to \mathbb{R}$ over surpluses (or profits) with $u(0) = 0$. We assume that u is strictly increasing.

Bidders' values $(\tilde{v}_1, \tilde{v}_2)$ are jointly distributed on V^2 according to the following compound distribution:[9] With probability $\rho \in (0, 1)$, bidders have a common value $\tilde{v}$ distributed according to a full-support probability mass function p on V. With the remaining probability $1 - \rho$, they have values $(\tilde{w}_1, \tilde{w}_2)$ that are independently and identically distributed according to the mass function p on V. Hence, the parameter ρ is the correlation coefficient between $\tilde{v}_1$ and $\tilde{v}_2$. We also denote by $\tilde{w}$ a random variable with distribution p.

The auction rule is defined as follows: The set of sealed bids is $B := V \cup \{0\}$. Any positive bid $b_i > 0$ means opt-in (entry). The opt-out bid $b_i = 0$ always gives bidder i zero profit. We append the following *(full-extraction) lottery* $l_0 : B \to \mathbb{R}$ to the 2P auction:

$$l_0(b) := \frac{1-\rho}{\rho} \left((1-\rho) E \left[\max\{\tilde{w}_i - \tilde{w}_j, 0\} \right] - E \left[\max\{b - \tilde{w}, 0\} \right] \right) \tag{1}$$

for each $b \in V$, and $l_0(0) := 0$. The item is sold in the same way as in the 2P auction. In addition, any bidder i who opts in obtains a lottery prize $l_0(b)$ that depends on the opponent j's bid b. We assume $v^1 < l_0(v^1)$ for simplicity, which holds if v^1 is close to zero.

This is a Vickrey–Clarke–Groves mechanism, so that value bidding $b_i = v_i$ (weakly) dominates non-value bidding $b_i \in V \setminus \{v_i\}$. Moreover, the lottery allows the seller to extract full surplus on average from risk-neutral bidders who bid their

[9] Random variables are denoted with tildes.

values, and each bidder i then breaks even as follows:

$$E\left[v - \min\{v, \tilde{v}_j\} + l_0(\tilde{v}_j) \mid \tilde{v}_i = v\right] = 0 \tag{2}$$

for each $v \in V$. The break-even conditions (2) indeed pin down the lottery l_0 as the function (1) (Lemma 1). Note that the lottery l_0 is strictly decreasing on V. The prize $l_0(b)$ is positive (i.e., bonus) if the opponent's positive bid b is low, and negative (i.e., penalty) if b is high. Since entry can entail losses, risk-averse bidders may rationally choose to opt out. We therefore introduce a more general lottery $l : B \to \mathbb{R}$ defined as $l(b) := l_0(b) + \pi$ for $b \in V$ and $l(0) := 0$, where the fixed amount $\pi \geq 0$ is called a *premium*.

The *auction game* proceeds as follows: First, values $(\tilde{v}_1, \tilde{v}_2)$ are realized according to the compound distribution. Each bidder i privately knows the realized value v_i. Second, each bidder i submits a sealed bid $b_i \in B$. Third, a bidder i with the higher bid $b_i > b_j$ wins the item at price b_j. If the tie $b_1 = b_2 > 0$ occurs, then either bidder wins the item at price b_1 with equal probability. If $b_1 = b_2 = 0$, the item is unsold. Finally, each bidder i has a utility $u(x_i)$, where the ex-post surplus or profit x_i is determined as follows: $x_i = v_i - b_j + l(b_j)$ if i opts in and wins, $x_i = l(b_j)$ if i opts in but loses, and $x_i = 0$ if i opts out.

A (behavioral) strategy is denoted by $\sigma \in \Delta(B)^V$, where $\sigma(v, b)$ is the probability with which a bidder with value v submits bid b. Let $U^\sigma(v, b)$ denote the interim utility of a bidder with value v bidding b given the opponent's strategy σ. Specifically, $U^\sigma(v, 0) = 0$ and

$$U^\sigma(v, b) = E\left[\sum_{b' \leq b} u(v - b' + l(b'))\frac{\sigma(\tilde{v}_j, b')}{1 + I(b' = b)} + \sum_{b' \geq b} u(l(b'))\frac{\sigma(\tilde{v}_j, b')}{1 + I(b' = b)} \mid \tilde{v}_i = v\right]$$

for $b > 0$, where I is the indicator function. With some abuse of notation, we also define $U^\sigma(v, \sigma') := \sum_b U^\sigma(v, b)\sigma'(v, b)$ and $U(\sigma', \sigma) := E\left[U^\sigma(\tilde{v}_i, \sigma')\right]$ for i's own strategy σ'. Our main focus is on symmetric Bayesian Nash equilibrium, that is, a strategy σ such that $U(\sigma, \sigma) \geq U(\sigma', \sigma)$ for each strategy σ'. We simply call it an *equilibrium*.

2.2 Equilibrium and Stability

This subsection is devoted to equilibrium and stability analyses. As already mentioned, value bidding is dominant given entry. Moreover, every equilibrium σ requires bidders with any risk preference to do so (i.e., $\sigma(v, v) + \sigma(v, 0) = 1$). Using this fact, we often identify an equilibrium with the *opt-out strategy* $z \in [0,1]^V$ such that a bidder with value v submits $b = 0$ and $b = v$ with probabilities $z(v)$ and $1 - z(v)$ respectively

Given the opponent's opt-out strategy z, we denote by $U^z(v)$ a bidder's interim utility from value bidding:

$$U^z(v) = E\left[u(v)z(\tilde{v}_j) + u(v - \min\{v, \tilde{v}_j\} + l(\tilde{v}_j))(1 - z(\tilde{v}_j)) \mid \tilde{v}_i = v\right]. \qquad (3)$$

This utility is increasing in the opt-out probability $z(v_j)$ if and only if the externality $l(v_j) - \min\{v, v_j\}$ from the opponent's entry is nonpositive. Hence, bidders' entry decisions are strategic substitutes (as in the "hawk-dove" game) if each bidder i has a high value $v_i > l(v_i)$, and strategic complements (as in the "stag-hunt" game) if each i has a low value $v_i < l(v_i)$. This result is the key to understanding our theoretical and experimental results. If bidders are risk neutral with $u(x) \equiv x$, then the break-even condition (2) leads to the following equation:

$$U^z(v) = \rho(v - l(v))z(v) + (1 - \rho)E[(\min\{v, \tilde{w}\} - l(\tilde{w}))z(\tilde{w})] + \pi. \qquad (4)$$

We classify equilibria z into the two categories: the *full-entry equilibrium* if $z(v) = 0$ for each $v \in V$, and a *partial-entry equilibrium* if $z(v) > 0$ for some $v \in V$. Since the premium π is nonnegative, the full-entry equilibrium exists, provided bidders are risk neutral (or risk loving). The following theorem provides a partial characterization of the set of equilibria for risk-neutral bidders. In particular, the game with $\pi = 0$ has exactly two equilibria.

Theorem 1. *Suppose that bidders are risk neutral. Then, there exists a threshold* $\underline{\pi} > 0$ *of premium with the following properties: (i) If* $\pi < \underline{\pi}$*, then the auction game has a unique equilibrium* z^c *among those* z *that satisfy* $z(v) > 0$ *for some* $v > l(v)$*. If* $\pi = 0$*, then* z^c *is the unique partial-entry equilibrium. (ii) If* $\pi \geq \underline{\pi}$*, then every equilibrium* z *satisfies* $z(v) = 0$ *for each* $v > l(v)$*.*

The partial-entry equilibrium z^c itself is characterized in the proof. It has a cutoff value $v^c > l(v^c)$ such that $z^c(v^1) = \cdots = z^c(v^{c-1}) = 1 > z^c(v^c) > \cdots > z^c(v^n) > 0$. Hence, bidders always break even in this equilibrium. Indeed, the indifference between opt-in and opt-out for high values implies that bidders with low values $v < l(v)$ would suffer a loss $U^{z^c}(v) < 0$ from entry. Roughly speaking, mixed entry for bidders with high values $v \geq v^c$ follows from the fact that entry decisions are strategic substitutes for them. Recall that the unique symmetric equilibrium in the hawk-dove game is also mixed.

We then argue that the full-entry equilibrium for low premia is implausible from the perspective of evolutionary game theory (Fudenberg and Levine, 1998; Weibull, 1997). Following Selten (1983), we first define a *(direct) evolutionarily stable strategy (ESS)* as an equilibrium σ such that $U(\sigma,\sigma) = U(\sigma',\sigma)$ for $\sigma' \neq \sigma$ implies $U(\sigma,\sigma') > U(\sigma',\sigma')$. Equivalently, an ESS is a behavioral strategy σ such that, for each $\sigma' \neq \sigma$, there exists a number $\varepsilon_{\sigma'} \in (0,1]$ with $U(\sigma, \varepsilon\sigma' + (1-\varepsilon)\sigma) > U(\sigma', \varepsilon\sigma' + (1-\varepsilon)\sigma)$ for all $\varepsilon \in (0, \varepsilon_{\sigma'})$. Here, σ is called an *incumbent*, σ' a *mutant*, and $\varepsilon_{\sigma'}$ an *invasion barrier* of σ against σ'. We next define the *replicator dynamics* as the system of $n(n+1)$ differential equations:

$$\dot{\sigma}(v,b) = p(v)\left(U^{\sigma}(v,b) - U^{\sigma}(v,\sigma)\right)\sigma(v,b) \tag{5}$$

for each $v \in V$ and $b \in B$, where the dot denotes the time derivative, with the time variable being omitted.[10] The *state space* is given by $\Delta(B)^V$. A strategy σ is called an *asymptotically stable strategy (ASS)* if the population state σ is asymptotically stable in the replicator dynamics (5). We can verify that every ESS is an ASS, by applying standard methods with a so-called Lyapunov function (Weibull, 1997, Proposition 3.10).

Remark 1. In the 2P auction (i.e., $l(b) \equiv 0$), value bidding is the unique ESS for bidders with any risk preference. This is because value bidding is the unique best response to itself for bidders with $v > v^1$, and any mutant σ' with non-value bidding by $v = v^1$ earns less against itself than value bidding σ does against σ' (i.e., $U(\sigma',\sigma') < U(\sigma,\sigma')$).

We now obtain the following (in)stability results for the full-entry equilibrium:

Theorem 2. *Suppose that bidders are risk neutral. Let σ denote the full-entry equilibrium. (i) If $\pi = 0$, then σ is not an ASS. (ii) If $\pi > 0$, then σ is an ESS. The highest invasion barrier $\varepsilon_{\sigma'}(\pi)$ of the incumbent σ against each mutant $\sigma' \neq \sigma$ is increasing in premium π on $\mathbb{R}_{++}$, and $\varepsilon_{\sigma'}(\pi) \to 0$ as $\pi \to 0$ for some mutant σ'.*

The proof of Theorem 2 is simple and summarized as follows: If $\pi = 0$, then opt-out spreads in the subpopulation of each value $v < l_0(v)$ (i.e., the

[10] This is a model for the agent normal form where selection of bids takes place within the "subpopulation" of each value v. We can also consider standard replicator dynamics for the normal form. In the latter model for our auction game, the state space is the set $\Delta(B^V)$ of *mixed* strategies, where B^V is the set of *pure* strategies. Any result presented in this paper that a strategy is an ASS (or not an ASS) still holds in the dynamics for the normal form of our auction game. Moreover, the normal-form dynamics cause computational problems in estimation reported in Section 5 since $|B^V| = (n+1)^n$ is large. Therefore, we use the more tractable model (5).

strategic-complements region) given value bidding in the subpopulation of each value $v > l_0(v)$ (i.e., the strategic-substitutes region). If $\pi > 0$, then the full-entry equilibrium is strict, and hence, an ESS. However, the decrease in premium π lowers the invasion barrier against mutants because the full-entry incumbent then loses relative advantages over them. In particular, the invasion barrier against a mutant that opts in if $v < l_0(v)$ and opts out otherwise converges to zero as $\pi \to 0$. An important implication for our experiment is that, given low premia π, the full-entry equilibrium may collapse if a few subjects with low values opt out of the auction.

The instability of the full-entry equilibrium shifts our attention toward partial-entry equilibria. However, the auction game with positive premia π has multiple partial-entry equilibria in general, as in the stag-hunt game with three symmetric equilibria. As a benchmark, we introduce a *minimal-entry equilibrium* that has the lowest entry probability $1 - E[z(\tilde{v}_i)]$ among all equilibria z. In particular, this benchmark is useful for a treatment with high premia. To characterize this equilibrium for risk-neutral bidders, we construct an opt-out strategy z^d as follows: we first define z^d given any cutoff value v^d so that $z^d(v) := 1$ if $v < v^d$ and $z^d(v) := 0$ otherwise, and then choose the highest one among all cutoff values v^d that satisfy $U^{z^d}(v) < 0$ if $v < v^d$ and $U^{z^d}(v) \geq 0$ otherwise.[11] The next theorem characterizes the minimal-entry equilibrium:

Theorem 3. *Suppose that bidders are risk neutral. Let $\underline{\pi}$ denote the threshold of premium in Theorem 1. (i) If $\pi < \underline{\pi}$, then the partial-entry equilibrium z^c in Theorem 1 is the unique minimal-entry equilibrium. (ii) If $\pi \geq \underline{\pi}$, then the strategy z^d is the unique minimal-entry equilibrium. If $\pi > \underline{\pi}$, then it is an ESS.*

We can expect that the partial-entry equilibrium z^c in Theorem 1 is also an ESS, recalling that a mixed-strategy equilibrium in the hawk-dove game is evolutionarily stable. There are, however, multiple subpopulations in the hawk-dove region $v > l(v)$ in general. Although this prevents us from showing a general result, we have confirmed that the partial-entry equilibrium z^c is an ESS in our experimental setting. This is shown in the next subsection.

2.3 Stable Equilibria in the Experimental Setting

This subsection illustrates stable equilibria in our experimental setting. The environmental conditions are specified as follows: the set of values is given by $V =$

[11] At least one such value v^d (e.g., v^1) exists.

$\{100, 200, ..., 10000\}$, the correlation coefficient by $\rho = 1/3$, and the probability mass function by $p(v) \equiv 1/100$. The full surplus is $E[\max\{\tilde{v}_1, \tilde{v}_2\}] = 6161$. The full-extraction lottery (1) is given by $l_0(v) = 2222 - (v/100)((v/100) - 1)$ and the threshold in Theorem 1 by $\underline{\pi} \approx 159$. The premium is set at either $\pi = 1$ or $\pi = 991$. We now claim that if $\pi = 1$, the minimal-entry equilibrium z^c is the unique ESS (and ASS) among all partial-entry equilibria:[12]

Claim 1. *Suppose that bidders are risk-neutral. In our experimental setting with $\pi = 1$, (i) the minimal-entry equilibrium z^c in Theorem 3 is an ESS, and (ii) no other partial-entry equilibrium is an ASS.*

Finally, we show how the minimal-entry equilibria z^c and z^d in Theorem 3 change with bidders' risk attitude. We assume that bidders have the constant-absolute-risk-aversion (CARA) utility function u defined as follows:

$$u(x) := \frac{1 - \exp(-rx)}{r} \tag{6}$$

for each $r \neq 0$ and $u(x) := x$ for $r = 0$. Given the low premium $\pi = 1$ and some CARA parameters r, Figure 1 illustrates ESSs that are "connected" to the minimal-entry equilibrium z^c for risk-neutral bidders. Similarly, given the high premium $\pi = 991$ and some r, Figure 2 illustrates ESSs that are connected to the minimal-entry equilibrium z^d. Table 1 summarizes aggregate equilibrium outcomes.

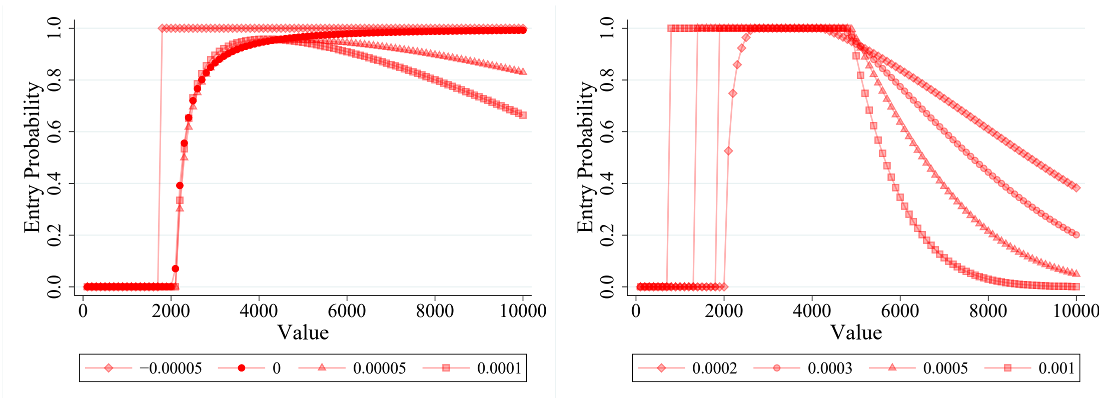


Figure 1: ESS z for Each CARA Parameter r. *Note*: The premium is $\pi = 1$. The entry probability is $1 - z(v)$. The ESS for $r = 0$ is the minimal-entry equilibrium z^c in Theorem 3 (i) with the cutoff $v^c = 2100$.

[12]For $\pi = 0$, we can confirm that z^c is an ESS by using the same way as Claim 1.

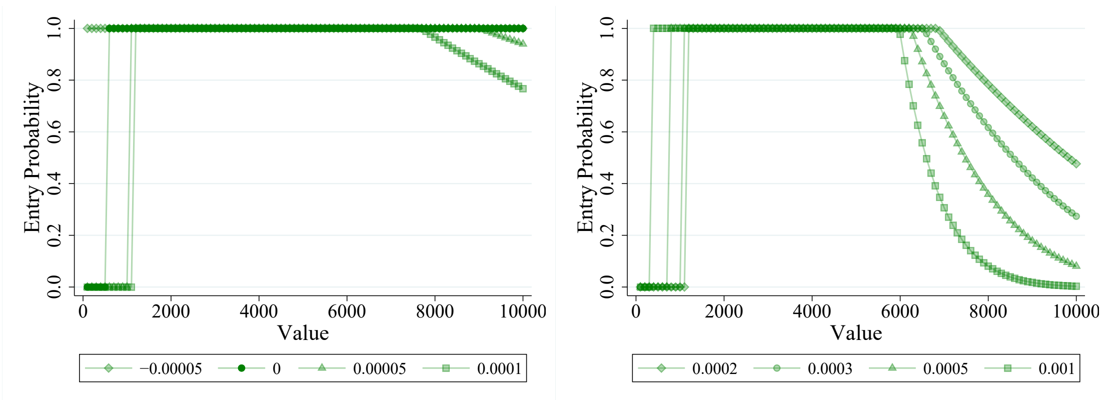


Figure 2: ESS z for Each CARA Parameter r. *Note*: The premium is $\pi = 991$. The entry probability is $1 - z(v)$. The ESS for $r = 0$ is the minimal-entry equilibrium z^d in Theorem 3 (ii) with the cutoff $v^d = 600$.

Table 1: Aggregate outcomes for each ESS z in Figures 1 and 2.

	$\pi = 1$				$\pi = 991$			
CARA Parameter r	Seller's Revenue	Social Surplus	Bidder's Profit	Entry Prob.	Seller's Revenue	Social Surplus	Bidder's Profit	Entry Prob.
−0.00005	6316	6087	−114	0.83	4179	6161	991	1.00
0	5968	5968	0	0.74	4456	6155	849	0.95
0.00005	5010	5728	359	0.70	4574	6116	771	0.89
0.0001	4143	5450	653	0.66	3979	5947	984	0.86
0.0002	2818	4858	1019	0.61	2741	5493	1375	0.79
0.0003	1941	4292	1175	0.58	1814	5013	1599	0.75
0.0005	1014	3460	1222	0.54	716	4231	1757	0.69
0.001	276	2555	1139	0.51	−202	3259	1748	0.64

Note: The aggregate outcomes are evaluated as ex-ante expected values. The seller's ex-ante revenue is $E[(1 - z(\tilde{v}_1))(1 - z(\tilde{v}_2))(\min\{\tilde{v}_1, \tilde{v}_2\} + l(\tilde{v}_1) + l(\tilde{v}_2))]$. The ex-ante social surplus from trade is $E[(1 - z(\tilde{v}_1))((1 - z(\tilde{v}_2)) \max\{\tilde{v}_1, \tilde{v}_2\} + 2(1 - z(\tilde{v}_1))z(\tilde{v}_2)\tilde{v}_1)]$. This surplus does not include risk premia for non-risk-neutral bidders. Each bidder's ex-ante profit is half of the difference between the social surplus and the seller's revenue. The ex-ante entry probability is $1 - E[z(\tilde{v}_i)]$.

Figures 1 and 2 show that the entry probability $1 - z$ for risk-averse bidders is nonmonotonic in their values. In particular, it is decreasing for high values. The intuition for this result is as follows: a risk-averse bidder i with a high value v with $l(v) < 0$ is reluctant to opt in because the opponent, who has the same value v

with high probability $\rho + (1-\rho)p(v) = 51/150$, bids the value v given entry, which in turn causes bidder i the ex-post loss $u(l(v)) < 0$ from entry. As bidders become extremely risk-averse (i.e., $r \to \infty$), this partial-entry equilibrium converges to the opt-out strategy z such that $z(v) = 0$ if $l(v) \geq 0$, and $z(v) = 1$ if $l(v) < 0$. As shown by Table 1, the seller's ex-ante revenue rapidly declines as bidder are more risk-averse. This is because high-value bidders opt out more often, and thus, the seller fails to collect high lottery penalties.

3 Experimental Design

This section presents the procedure used in our experiment. As explained in Section 1, we control confounding factors as much as possible by applying several methods developed in the auction-experiment literature.

3.1 Procedure

Our between-subjects experiment consists of three treatments: *full extraction* (FE), *partial extraction* (PE), and *second price* (2P). Given the environmental parameters V, ρ, p in Section 2.3, each treatment is defined as follows: First, each pair of subjects in FE and PE plays the auction game in Section 2.1 with the premium $\pi = 1$ and $\pi = 991$ respectively. The premium $\pi = 1$ is the lowest integer amount that endows the full-entry equilibrium with evolutionary stability, but this amount is so low that its invasion barrier against mutants is also low, as shown by Theorem 2. The little premium also provides risk-averse subjects only weak incentives to opt in. By contrast, the premium $\pi = 991$ is high enough. We chose this amount so that the seller's ex-ante revenue in the full-entry equilibrium is approximately equal to the dominant-equilibrium revenue (4179) in the 2P auction with the optimal reserve price (3800). Next, 2P is a benchmark treatment where each pair of subjects plays the 2P auction game with neither lottery nor reserve price.

Each session has 20 auction rounds. We adopted the following group matching (Tan, 2020) as a matching rule: subjects were randomly divided into groups of 6 at the beginning of a session, and 3 pairs were randomly formed within a group in each round. This rule allows us to regard each group as the unit of independent observation in statistical analyses. We chose this rule to obtain sufficient statistical power when testing our theoretical hypotheses, and at the same time to mitigate repeated-game effects as much as possible. Each group corresponds to a single population in the evolutionary-game model of Section 2.2 although the population size is not so large.

Each session proceeded as follows: First, subjects with instruction handouts watched a video explaining the rule of the experiment.[13] The instruction announced that the initial endowment for each subject is JPY 3,000 (USD 23 at the time of the experiment). Second, each subject chose one lottery for each problem in gain and loss domains. The multiple price lists (MPL) are shown by Table 1 of SM (Holt and Laury, 2002; Laury and Holt, 2005). Third, subjects played 2 practice rounds of the auction game, where their opponents' bids were predetermined to be 5000 and 0. Fourth, subjects had a comprehension quiz about the rule of the experiment. Each group could not proceed until all group members correctly answered all questions. Fifth, subjects played 20 rounds of the game. The unit of profit in the game was called point. The exchange rate was JPY 1 per 10 points in a payment round. Sixth, subjects took the cognitive reflection test (CRT), where they answered 7 questions in Table 3 of SM (Frederick, 2005; Toplak et al., 2014). Each subject obtained JPY 60 per correct answer. Finally, the session ended with a questionnaire that asked gender, age, and how they had made entry and bidding decisions.

The experiment was programmed in oTree (Chen et al., 2016). Figures 2–4 of SM illustrate the screenshots of bidding and result pages. The joint distribution of values $\tilde{v}_1$ and $\tilde{v}_2$ was shown to subjects in a table (Figure 1 of SM). The bidding page also displayed the conditional probability of the opponent's value v_j given bidder i's own value v_i. The probability is equal to $\rho + (1-\rho)p(v_j) = 51/150$ if $v_j = v_i$, and $(1-\rho)p(v_j) = 1/150$ if $v_j \neq v_i$. We gave this additional description because subjects may not update beliefs according to Bayes' rule, while it may help them correctly calculate interim profits (or utilities). Subjects in FE and PE could check the graph of the lottery l and use a calculator that displays the lottery prize for each bid (Figure 3 of SM).[14] The result page displayed the following items: winner, value v_i, bid b_i, the opponent's bid b_j, price $\min\{b_i, b_j\}$, profit $v_i - \min\{b_i, b_j\}$ (if i won), lottery prize $l(b_j)$, and total profit x_i.[15] On the bidding page, subjects could check these results in previous rounds.

To minimize the effect of social preferences on bids, we used the amended random payment scheme developed by Lim and Xiong (2021). Under this scheme, one payment round $t \in \{1, ..., 20\}$ is randomly chosen for each subject, who is

[13]Appendix D of SM provides the English translation of the instruction in FE.

[14]To avoid any framing effect, the lottery prize $l(b_j)$ for bidder i was called the profit adjustment (or change).

[15]The lottery prize and total profit were displayed only in FE and PE.

uninformed about t until the end of a session. Each subject is, however, informed about whether each round is the current opponent's payment round (Figure 2 of SM). Subjects with social preferences should then care only about their own material payoffs in opponents' *non*-payment rounds. We will focus on these rounds in data analyses unless otherwise stated.

Furthermore, to alleviate cognitive problems, we applied ascending-clock presentation developed by Breitmoser and Schweighofer-Kodritsch (2022). Under this feedback method, a pair of subjects watches an "ascending price" after they submit sealed bids. Specifically, on the result page (Figure 4 of SM), the black bar gets longer until it reaches the second-highest bid. We adopted a static auction format with ascending-clock presentation in all of the three treatments.

3.2 Administrative Details

The experiment was conducted at the laboratories of Osaka and Kansai Universities in Japan in 2023. In total, 367 subjects (students at Osaka and Kansai Universities) participated in our experiment. They were recruited from the subject pools of the ISER at Osaka University and the RISS at Kansai University through the ORSEE system (Greiner, 2015). Since our matching rule prescribes that the group size is 6, remainder subjects in each session, if any, joined a group with computerized bidders. The total number of subjects who joined human-only groups was 330. We will focus on these subjects in statistical analyses. The numbers of sessions, groups, and subjects in each treatment, together with summary of subjects' characteristics, are shown in Table 2.[16] See Table 4 of SM for the details of sessions.

[16]Following the pre-registration, we will exclude data as outliers for subjects (2 in FE, 2 in PE, and 3 in 2P) who cumulatively made 20 or more mistakes in the comprehension quiz. Due to machine troubles, subjects in 2 sessions (8 groups) of FE could not watch the instruction video. However, they could hear the sound of the video, looking at their instruction handouts. There seems to be no essential difference in bidding behavior between sessions with and without these troubles (Figures 5 and 6 and Table 5 of SM). Therefore, we will include data in these two sessions.

Table 2: Treatments.

		FE	PE	2P	FE vs. PE	FE vs. 2P	PE vs. 2P
# of Sessions		8	5	5			
# of Groups		27	14	14			
# of Subjects		162	84	84			
Gender	Male	45.6%	45.2%	57.1%			
	Female	47.5%	51.1%	38.0%			
	Other	2.4%	0%	3.5%			
	No Answer	4.3%	3.5%	1.1%			
Age		21.4	21.5	21.2	0.6104	0.7025	0.4574
# of Safe Choices (Gain)		5.7	5.9	5.7	0.5915	0.9028	0.5105
# of Safe Choices (Loss)		6.0	5.8	5.8	0.3721	0.6520	0.7359
CRT Score		5.3	5.2	5.5	0.4500	0.3558	0.1201
# of Mistakes in Quiz		4.0	4.2	3.5	0.9752	0.0114	0.0261

Note: This table summarizes information about the treatments. For each treatment, the table shows the gender composition, and the average age, numbers of safe choices in the MPL, CRT score, and number (cumulative total) of mistakes in the comprehension quiz. The last three columns show the exact p-values of Wilcoxon rank-sum tests for each treatment pair.

Each session lasted approximately 80 minutes on average, including 12 (FE and PE) or 10 (2P) minutes for the instruction. The average total payment per subject was JPY 3,499. The highest and lowest payments were JPY 4,720 and JPY 1,930 respectively. Our samples are well-balanced across treatments. For every treatment pair and each item, Table 2 presents a p-value from the Wilcoxon rank-sum test of the null hypothesis that subjects' characteristics in the two treatments come from the same distribution.[17] In the quiz, subjects made fewer mistakes in 2P than in FE and PE ($p = 0.0114$ for FE vs. 2P; $p = 0.0261$ for PE vs. 2P). This is a natural result since the rule of FE (or PE) is more difficult than that of 2P, and the former treatment has 10 problems in the quiz while the latter 8 problems.

The histograms for subjects' characteristics are shown by Figure 7 of SM. In every treatment, the mode categories of risk attitudes are risk aversion and risk neutrality in the gain and loss domains respectively.[18] Table 2 of SM also shows the range of CARA parameters given the number of safe choices in the MPL. The

[17] All statistical tests in this paper are two-sided.

[18] In the gain domain, 93% (FE), 95% (PE), and 92% (2P) of subjects correctly switched lottery choices at most once. In the loss domain, 95% (FE), 97% (PE), and 95% (2P) did so. There was no irrational subject who chose the safe (risky) lottery in the last problem in the gain (loss) domain.

CARA parameters in Table B.II are comparable with those in Figures 1 and 2 since the exchange rate was JPY 1 per 10 points.

4 Experimental Results

In this section, we report experimental results and compare them with the theoretical predictions derived in Section 2. According to the stability results in Theorems 2 and 3 with Claim 1, the main predictions for FE and PE are the minimal-entry equilibria z^c and z^d in Figures 1 and 2 respectively—and the full-entry equilibrium. The full-entry equilibrium in FE, however, is implausible because of the low invasion barrier against mutants as shown by Theorem 2 (ii). From Remark 1, the unique prediction for 2P is the full-entry equilibrium. In Section 4.1, we report results on the seller's revenue, which provide suggestive evidence against the full- and minimal-entry equilibria in FE. We then investigate subjects' bidding behavior in Section 4.2 and classify them according to their behavior in Section 4.3.

4.1 Seller's Revenue

We begin with auction performances for the seller.[19] Figure 3 shows the seller's average realized revenues in the experiment and the equilibrium ones. The realized revenues were 5370, 4980, and 3461 points in FE, PE, and 2P respectively. Given the realized values in the experiment, the full-entry equilibrium revenues are 6308, 4311, and 3982 points in FE, PE, and 2P respectively. Similarly, the minimal-entry equilibrium revenues are 6113 and 4580 points in FE and PE respectively.[20] The realized revenue in FE is thus lower than the full- and minimal-entry equilibrium ones by around 15% and 12% respectively.

We then test the equilibrium hypotheses about the seller's revenue. The first hypothesis is the full-entry equilibrium. For each treatment, we conduct a Wilcoxon matched-pairs signed-rank test of the null hypothesis that the realized and equilibrium revenues come from the same distribution. The exact p-values are given by $p = 0.0121$ for FE, $p = 0.1040$ for PE, and $p = 0.0040$ for 2P respectively. Second, let us consider the minimal-entry equilibria in FE and PE as predictions.

[19] We have revised the pre-registered plan for statistical analyses. See Appendix C of SM for details.

[20] Note that, in PE, the full-entry revenue is less than the minimal-entry revenue because the latter equilibrium prevents low-value bidders' entry with high prizes for opponents (Figure 2).

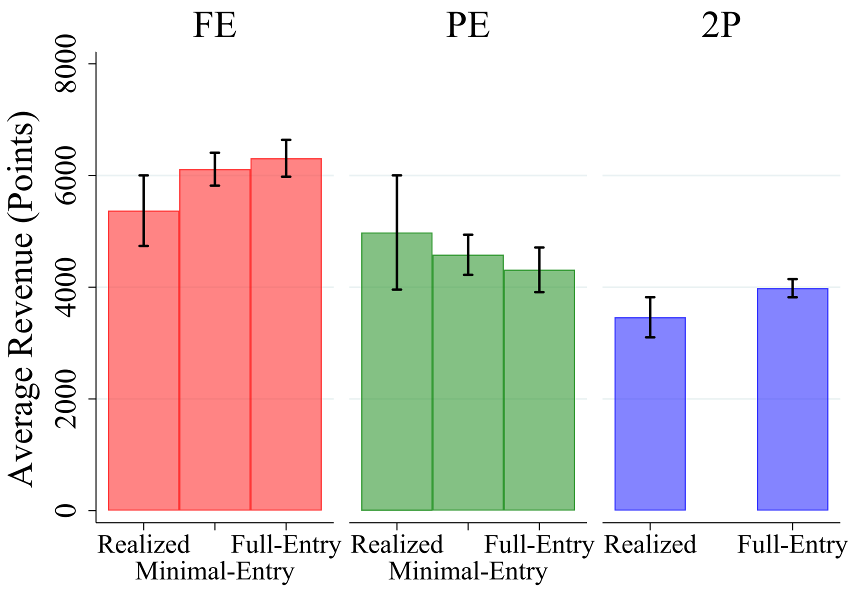


Figure 3: Realized and Equilibrium Revenues for the Seller. *Note*: The minimal-entry equilibria are given by Theorem 3. The unit of observation is a group. The capped spikes represent 95% confidence intervals.

The exact p-values from the signed-rank tests are p = 0.0340 for FE and p = 0.3910 for PE respectively.

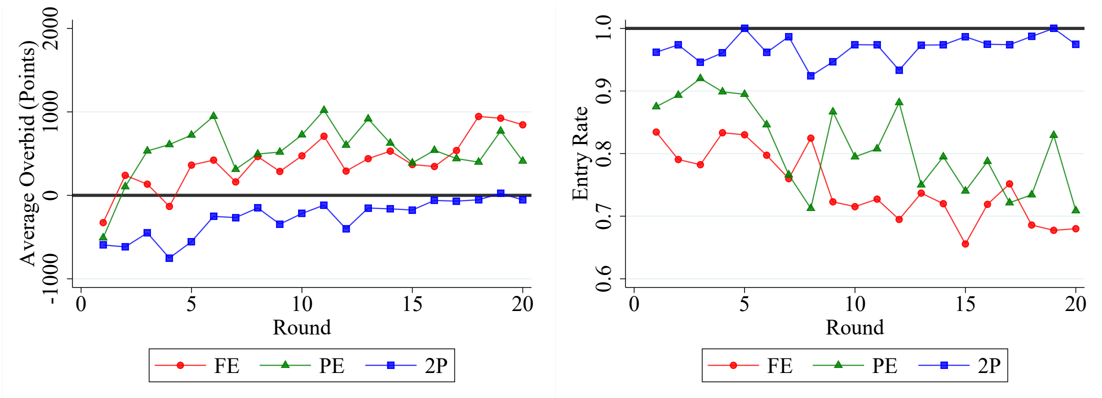
We also evaluate average treatment effects on revenues. For every treatment pair, we conduct a Wilcoxon rank-sum test of the null hypothesis that the realized revenues in these two treatments come from the same distribution. The exact p-values are p = 0.5408 for FE vs. PE, p = 0.0002 for FE vs. 2P, and p = 0.0088 for PE vs. 2P respectively.

4.2 Bidding Behavior

The results in Section 4.1 suggest that the performance in FE is worse than in the full- and minimal-entry equilibria. To find the causes of these underperformances, we investigate how subjects made bidding decisions in FE as well as in the other two treatments.

In panel data analyses, let $b_{it} \in B$ and $v_{it} \in V$ denote subject i's bid and value in round $t \in \{1, 2, ..., 20\}$ respectively. Subject i's overbid amount and deviation (from value bidding) given entry are measured as $b_{it} - v_{it}$ and $|b_{it} - v_{it}|$ with $b_{it} > 0$ respectively.

Figures 4 illustrates the time series of average overbid amounts given entry and entry rates. Figure 4a shows that the average overbid amount increases through



(a) Average Overbid Amount Given Entry (b) Entry Rate

Figure 4: Time Series of Average Overbid Amounts Given Entry and Entry Rates.

rounds in every treatment. A crucial difference between the treatments is that the amount deviates from zero in FE and PE, while it approaches zero from below in 2P. This trend away from underbidding toward value bidding is similar to those in the 2P ascending-clock-presentation treatments of Breitmoser and Schweighofer-Kodritsch (2022, Fig. 3). Figure 4b shows that the entry rate decreases through rounds in FE and PE, while it approaches one in 2P. The entry rates in FE and PE are around 0.7 in the final round. These rates are compatible with some equilibria for risk-averse bidders (Table 1), but the departure from value bidding given entry is not.

We conduct more rigorous analyses to examine the determinants of bidding behavior. Specifically, we report the regression results of multilevel mixed-effects models for each treatment. Each model in Tables 3–5 has group- and subject-level random effects. In these regression analyses, we include bidding data in opponents' payment rounds to examine whether and how subjects changed behavior in these rounds.[21]

[21]When reporting any effect in these analyses, we present a p-value from the z-test of the null hypothesis that the associated coefficient is zero. The z-statistic is defined as the estimate divided by the robust standard error clustered at the group level.

Table 3: Tobit regression models for bids.

Dependent Variable	Bid (in Entry Round)		
Treatment	FE	PE	2P
Value	0.74 (0.03)	0.70 (0.03)	0.90 (0.15)
Round	53.67 (17.74)	18.83 (21.83)	39.27 (9.21)
Last Opponent's Bid	0.06 (0.01)	0.08 (0.02)	−0.003 (0.01)
Opponent's Payment Round	−587.64 (405.63)	−840.39 (458.15)	−738.78 (475.19)
# of Mistakes in Quiz	−73.10 (53.03)	−49.01 (52.28)	−27.31 (29.06)
CRT Score	74.09 (95.20)	24.63 (117.08)	−144.75 (71.82)
Male	115.62 (320.99)	−633.48 (358.23)	−199.55 (215.07)
Constant	1221.51 (595.44)	2240.16 (580.80)	863.26 (393.76)
Var. of Group Random Effect	356992	117437	179579
Var. of Subject Random Effect	3739427	2084651	769620
Var. of Residual	5478177	5433625	1853798
Log Pseudolikelihood	−17453	−9970	−12571
# of Observations	2267	1264	1489

Note: Bid = $b_{it} > 0$. Value = v_{it}. Round = $t > 1$. Last opponent's bid = b_{jt-1}, where subject j is i's opponent in round $t-1$. Opponent's payment round = 1 if t is a payment round for the current opponent, and 0 otherwise. Male = 1 if subject i is male, and = 0 otherwise. Robust standard errors clustered at the group level are shown in parentheses. Var. denotes a variance.

Table 4: Linear regression models for deviations from value bidding.

Dependent Variable	Deviation (in Entry Round)		
Treatment	FE	PE	2P
Round	12.53 (10.93)	−35.36 (9.26)	−42.81 (9.75)
Last Opponent's Bid	0.01 (0.01)	0.03 (0.02)	−0.005 (0.01)
Opponent's Payment Round	1152.42 (326.77)	740.68 (230.27)	1129.32 (356.40)
# of Mistakes in Quiz	23.39 (29.66)	35.68 (33.13)	33.35 (30.51)
CRT Score	86.22 (49.99)	−88.42 (61.75)	−39.41 (63.68)
Male	31.48 (208.61)	−41.13 (279.51)	50.48 (191.32)
Constant	793.17 (303.74)	2235.32 (361.91)	1298.53 (394.39)
Var. of Group Random Effect	64148	76888	16585
Var. of Subject Random Effect	1318473	656056	833219
Var. of Residual	3243918	3261763	1447716
Log Pseudolikelihood	−20362	−11332	−12773
# of Observations	2267	1264	1489

Note: Deviation = $|b_{it} - v_{it}|$ with $b_{it} > 0$. The note of Table 3 is also applied.

Table 5: Linear regression models for entry decisions.

Dependent Variable	Entry Dummy × 100		
Treatment	FE	PE	2P
Value / 10^2	1.17 (0.10)	1.29 (0.22)	0.41 (0.14)
Value × Value / 10^4	−0.006 (0.0009)	−0.008 (0.001)	−0.002 (0.001)
Round	−0.81 (0.13)	−0.84 (0.15)	0.01 (0.09)
Last Opponent's Bid / 10^4	−8.15 (1.73)	−5.61 (2.03)	1.93 (1.55)
Opponent's Payment Round	7.15 (3.16)	7.02 (2.91)	−4.38 (3.80)
# of Safe Choices (Gain)	−2.44 (0.97)	−1.06 (1.31)	0.15 (0.31)
# of Safe Choices (Loss)	−1.14 (1.58)	1.97 (1.09)	−0.76 (0.38)
# of Mistakes in Quiz	0.58 (0.55)	−0.23 (0.43)	0.13 (0.20)
CRT Score	−0.55 (1.18)	−2.19 (0.51)	0.83 (0.60)
Male	−3.29 (3.56)	−5.37 (4.21)	0.94 (1.84)
Constant	72.44 (14.30)	66.51 (11.54)	82.27 (8.23)
Var. of Group Random Effect	5.06 / 10^{13}	34	1.62 / 10^{10}
Var. of Subject Random Effect	416	136	24
Var. of Residual	1177	1032	262
Log Pseudolikelihood	−13797	−7298	−5949
# of Observations	2755	1482	1406

Note: Entry Dummy $= I(b_{it} > 0)$. The note of Table 3 is also applied.

Table 3 presents tobit models in which the bids b_{it} in entry rounds, censored at the upper limit of 10000, are regressed on several explanatory variables. A robust result is that subjects with higher values bid higher ($p < 0.0001$ in all treatments). Subjects also raise their bids through rounds in FE and 2P ($p = 0.0024$ in FE; $p = 0.3884$ in PE; $p < 0.0001$ in 2P), and become more aggressive as their last opponents bid higher in FE and PE ($p < 0.0001$ in FE; $p = 0.0010$ in PE) although the impacts are small in size. The coefficients on the opponent-payment-round dummy are all negative ($p = 0.1474$ in FE; $p = 0.0666$ in PE; $p = 0.1200$ in 2P), contrary to spite hypotheses, but the standard errors of the coefficients are relatively high. It seems that subjects have heterogeneous social preferences since some subjects bid higher in opponents' payment rounds while others lower in these rounds. Table 4, which presents linear regression models for the deviations $|b_{it} - v_{it}|$ in entry rounds, indeed shows that the deviations are larger in opponents' payment rounds ($p = 0.0004$ in FE; $p = 0.0013$ in PE; $p = 0.0015$ in 2P). The deviations increase by around 700–1100 points in these rounds. This table also shows that the deviations decrease through rounds in PE and 2P ($p = 0.0001$ in PE; $p < 0.0001$ in 2P). There is no such declining trend in FE.

Table 5 shows linear models in which the entry dummy $I(b_{it} > 0)$ is regressed on several variables. We include the squared term of values since some equilibrium entry probabilities shown in Figures 1 and 2 are nonmonotonic in values. The estimated coefficients on v_{it} and v_{it}^2 imply that the entry rates reach the peaks at around $v = 9750$ (FE), $v = 8062$ (PE), and $v = 10250$ (2P). In FE and PE, subjects are less likely to opt in through rounds (p $<$ 0.0001 in both) and as the last opponents bid higher (p $<$ 0.0001 in FE; p $=$ 0.0057 in PE). In FE, as expected, more risk-averse subjects (in the gain domain) are less likely to opt in (p $=$ 0.0117). In PE, subjects with higher CRT scores are less likely to opt in (p $<$ 0.0001).

4.3 Subject Clusters

Subjects' bidding behavior in FE and PE is highly heterogeneous. To explore what types of strategies they used in each treatment, we classify subjects into clusters by applying the method of k-means clustering (Hastie et al., 2009; Macqueen, 1967).[22] Here, we are interested in bidding behavior of subjects who have learned group members' behavior to some extent. For this reason, we focus on bidding data in the second half of the experiment (i.e., rounds 11–20). As we will see below, even in these later rounds, no single equilibrium can explain subjects' behavior in FE (and PE).

Our clustering procedure is described as follows: For each subject i, we denote by $T_i' \subset \{11, 12, ..., 20\}$ the set of opponents' *non*-payment rounds and by $T_i := \{t \in T_i' \mid b_{it} > 0\}$ the subset of entry rounds. If subject i chose entry in at most one round (i.e., $|T_i| \leq 1$), then this subject is classified into cluster 0 ("infrequent-entry" cluster). Let us then suppose that subject i chose entry more than once (i.e., $|T_i| > 1$). To define a *feature vector* of this subject, we combine multiple estimates since the pure or behavioral strategy in the auction game is high-dimensional. For subject i, we obtain the OLS estimates $(\hat{\beta}_0, \hat{\beta}_1)$ and $(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2)$ from the two linear regression models:

$$b_{it} = \beta_0 + \beta_1 v_{it} + \varepsilon_{it}, \quad t \in T_i,$$
$$I(b_{it} > 0) = \gamma_0 + \gamma_1 v_{it} + \gamma_2 v_{it}^2 + \eta_{it}, \quad t \in T_i',$$

[22] Clustering approaches based on machine learning algorithms have been successfully introduced into the experimental economics literature. To organize high-dimensional behavioral data, Fréchette et al. (2022) used the k-means algorithm and Romero and Rosokha (2023) a clustering algorithm called affinity propagation.

where the errors ε_{it} and η_{it} are independently and identically distributed. These formulations follow the regression models in Tables 3 and 5.[23] We also calculate subject i's average overbid amount $\hat{o} := \frac{1}{|T_i|}\sum_{t \in T_i}(b_{it} - v_{it})$ in entry rounds and entry rate $\hat{e} := |T_i|/|T_i'|$. Subject i thus has a 7-dimensional feature vector $(\hat{\beta}_0, \hat{\beta}_1, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{o}, \hat{e})$.[24] Finally, for each treatment, the k-means algorithm classifies subjects into k clusters to (locally) minimize within-cluster variances of feature vectors, where the number k of clusters is specified by us. We set $k = 8$ for FE and $k = 5$ for PE and 2P, including cluster 0.

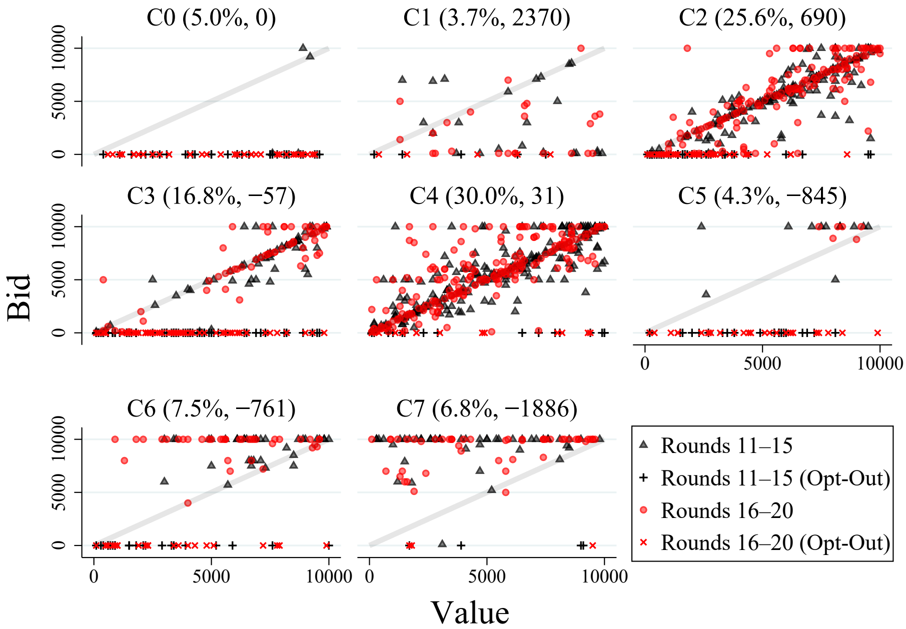


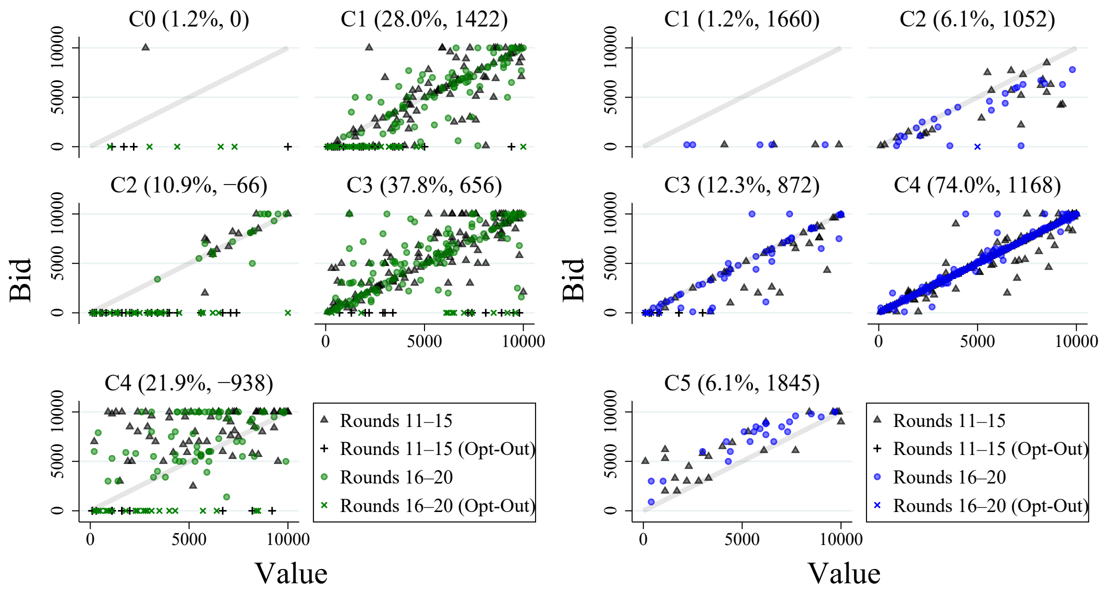
Figure 5: Subject Clusters in FE. *Note*: The capital C stands for Cluster. These clusters (excluding cluster 0) are consecutively numbered according to the average overbid amount in entry rounds within each cluster. For each cluster, the percentage and the number in parentheses show the proportion of subjects and the average profit in rounds 16–20, respectively. Triangle and plus symbols denote bids in rounds 11–15, and circle and cross symbols bids in rounds 16–20. The 45-degree line shows value bidding.

Figure 5 presents the clustering result for FE. Cluster 4 includes value-bidders, who almost always chose bids close to their values, with some noisy bidders. This cluster accounts for 30% of subjects. The bidding patterns in clusters 2 and 3

[23]We have chosen these simple models since the number of observations for each subject is at most ten. The use of the probit or the Heckman selection model requires much more variation in entry decisions.

[24]Subjects' feature vectors are standardized so that each coordinate has zero mean and unit variance within each treatment.

are qualitatively consistent with some partial-entry equilibria in Figure 1. The subject-level data (Figures 8 and 9 of SM) also suggest that subjects in these clusters had their own “threshold values” v at which they changed entry decisions. Clusters 2 and 3 together account for 42%. Clusters 5–7, which together account for 19%, have many overbidders with full or partial entry. In particular, cluster 7 includes 4 extreme overbidders who always bid the maximum amount. As shown in Figure 5, the average profits of clusters 5–7 in rounds 16–20 are negative. Cluster 1 has underbidders.



(a) PE (b) 2P

Figure 6: Subject Clusters in PE and 2P. *Note*: The note of Figure 5 is applied.

Figure 6a shows the clusters in PE. Cluster 3 includes value-bidders, with some subjects who opted out when their values were relatively *high*. This behavior is in accordance with partial-entry equilibria for risk-averse bidders (Figure 2). This cluster accounts for 38% of subjects. Clusters 1 and 2, which together account for 39%, are similar to clusters 2 and 3 in FE respectively. It again seems that subjects in these two clusters in PE had their own threshold values v for entry decisions (Figures 10 and 11 of SM). Their thresholds are, however, so high that the partial-entry equilibria in Figure 2 cannot explain their entry decisions. Cluster 4, which accounts for 22%, has some noisy overbidders.

Finally, Figure 6b shows the clusters in 2P. The clustering result is fairly simple: 74% of subjects are value-bidders in cluster 4, 20% are underbidders in clusters 1–3, and 6% are overbidders in cluster 5. Almost all subjects in 2P always chose entry in rounds 11–20.

5 Dynamics of Bidding Strategies

Finally, we analyze the dynamics of bidding strategies. Since some clusters in FE and PE earned negative profits as shown by Figures 5 and 6a, subjects might be still learning strategies even in later rounds. We explain their bidding behavior by using a model of perturbed replicator dynamics. This model then provides a long-run prediction about their behavior.

5.1 Perturbed Replicator Dynamics

Our clustering analysis for FE reveals that some subjects exhibited behavior in line with the full- or partial-entry equilibria, while others overbidding behavior. Altogether, Figure 5 shows that a majority of bids are on the "letter Z." Indeed, extreme overbidding $b = 10000$, value bidding $b = v$ and opt-out $b = 0$ account for 15%, 14% and 29% of all bids in rounds 16–20 of FE, respectively. The standard replicator dynamics (5) have difficulty in explaining this mixed bidding pattern for several reasons.

Most importantly, overbidding is dominated for selfish bidders with any risk preference, and hence, should approach "extinction." Specifically, $U^\sigma(v, v) \geq U^\sigma(v, b)$ holds for each value v, non-value bid $b \in V \setminus \{v\}$ and strategy σ, and the inequality is strict if $\sigma(v, b) = 1$. Hence, this non-value bid $b \neq v$ vanishes from the subpopulation of v in the limit $t \to \infty$ along every interior solution path to the replicator dynamics (5).[25] Of course, "irrational" overbidding can prevail during early periods in the dynamics, but this may not coexist with the "rational" strategy of value bidding given entry. Recall also that we adopted the amended random payment scheme to induce selfish preferences.

Secondly, the ESSs in Figure 1 require bidders with low values to opt out with probability one, while subjects with low values chose entry too often. In particular, this is the case with clusters 4 (value-bidders) and 7 (overbidders) in Figure 5.

We approach this problem by incorporating trends toward (extreme) overbidding and (full-entry) value bidding into the replicator dynamics (5). To implement this idea, we perturb the dynamics as follows:

$$\dot{\sigma}(v, b) = p(v) \left[(1 - \delta(t)) \left(U^\sigma(v, b) - U^\sigma(v, \sigma) \right) \sigma(v, b) + \delta(t) \left(\theta(v, b) - \sigma(v, b) \right) \right], \tag{7}$$

[25]We can prove this extinction result by applying Proposition 3.2 of Weibull (1997) to our dynamics.

where $\delta(t) \in [0, 1)$ denotes an error rate at time t and $\theta \in \Delta(B)^V$ is a strategy defined as $\theta(v, 10000) = \lambda$ and $\theta(v, v) = 1 - \lambda$ given a weight parameter λ. In words, the fractions $\delta(t)\lambda$ and $\delta(t)(1 - \lambda)$ of the subpopulation of v exogenously change bids from current ones to $b = 10000$ and $b = v$ respectively, while the original selection takes place in the other fraction $1 - \delta(t)$. Our formulation follows Gale et al. (1995), but has a crucial difference from their perturbed replicator dynamics in that they mainly focused on random drift (i.e., uniform noise), which here corresponds to $\theta \equiv 1/|B|$.[26] This is why we adopt the term "trend" rather than "drift," which reminds us of random genetic drift in biology. We do *not* insist that the trends toward overbidding and value bidding are really perpetual. We regard the exogenous changes of bids as naïve and temporary behavior. In Section 5.2, we will discuss possible reasons for these trends in our experiment. Here, let us assume that the error rate is strictly decreasing in time and will vanish as follows:

$$\delta(t) := \varepsilon e^{-\eta t}$$

for some initial error rate $\varepsilon \in [0, 1)$ and decay rate $\eta > 0$.

For estimation, we next introduce a discrete-time version of the perturbed replicator dynamics. In particular, we apply the reinforcement learning model of Börgers and Sarin (1997). This model is suitable for our experiment, in which subjects got less feedback about opponents' strategies. As explained in Appendix E of SM, this learning process is stochastic, and the expected movement of the state is determined by the following system of $n(n + 1)$ difference equations given an initial state $\sigma(0) \in \Delta(B)^V$:

$$\frac{\sigma_b^v(t+\tau) - \sigma_b^v(t)}{\tau p(v)} = (1 - \delta(t)) \left(U^{\sigma(t)}(v, b) - U^{\sigma(t)}(v, \sigma(t)) \right) \sigma_b^v(t) + \delta(t) \left(\theta_b^v - \sigma_b^v(t) \right) \tag{8}$$

at each time $t \in \{0, \tau, 2\tau, ...\}$, where σ_b^v and θ_b^v denote $\sigma(v, b)$ and $\theta(v, b)$ respectively. As the step size τ goes to zero, the model (8) converges to the continuous-time model (7).

We are interested in time t at which the state $\sigma(t)$ fits well the data in rounds 16–20. To avoid overfitting, we basically assume that the initial state $\sigma(0)$ is

[26] Their purpose is to show that small noises can rather stabilize imperfect Nash equilibria in ultimatum bargaining. See also Binmore and Samuelson (1999) for the role of evolutionary drift in equilibrium selection.

uniform.[27] Yet, the opt-out bid can be distinguished from positive bids. We therefore parameterize the initial state using a weight $\alpha \in [0,1]$ such that $\sigma_b^v(0) := \alpha/2 + (1-a)/101$ if $b = 0$ and $\sigma_b^v(0) := \alpha/200 + (1-a)/101$ if $b > 0$. The value $\alpha = 1$ implies that random decisions are made at "two stages," while there is no such distinction between $b = 0$ and $b > 0$ if $\alpha = 0$. Our design of the bidding page (Figure 2 of SM) enhances the former possibility. We also assume that bidders have the CARA utility function (6) with a parameter r. Since our learning model follows Börgers and Sarin (1997), a positive affine transformation is applied to the utility function u so that $u(x) \in (0,1)$ for all profit levels x in the experiment. See Appendix E of SM for details.

Our model thus contains the five parameters $(\alpha, \varepsilon, \lambda, r, t)$ to be estimated through a maximum-likelihood method. We set $\eta = 0.001$ since it turns out that lower decay rates provide better fits.[28] With the bidding data in rounds 16–20 of each treatment, we conducted a grid search over parameter values in the note of Table 6 to find a combination that maximizes the log-likelihood (LL):

$$\sum_{i \in N} \sum_{s \in T_i} \ln \sigma(v_{is}, b_{is})(t), \tag{9}$$

where σ is the solution path to the discrete-time dynamics (8) given the initial state $\sigma(0)$, N the set of subjects in a treatment, and $T_i \subset \{16, 17, ..., 20\}$ the set of non-payment rounds for subject i's opponents. The estimates are presented in Table 6, which also shows the estimates for three restricted models.[29] The full model provides the best fit (and the unperturbed model the worst fit) for every treatment according to the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

[27]This approach has a flavor similar to the logit solution (Goeree et al., 2016), which traces back to completely random behavior.

[28]We also set $\tau = 10$. As τ is larger, learning is faster. We chose this large step size to speed up computation.

[29]The CARA estimates $\hat{r}$ in 2P are relatively high. This may be because 2P had many underbidders (Figure 6b). Note that, for a bidder with a concave (convex) utility function, underbidding is a less (more) costly mistake than overbidding due to loss aversion (loving).

Table 6: Maximum likelihood estimates for the replicator dynamics.

Treatment	Model	$\hat{\alpha}$	$\hat{\varepsilon}$	$\hat{\lambda}$	$\hat{r}$	$\hat{t}$	LL	AIC	BIC
FE	Full	0.2	0.007	0.7	0.0001	8440	−2295	4601	4624
	Restricted ($\lambda = 1$)	0.1	0.004	1	0.0001	10530	−2485	4979	4998
	Restricted ($\lambda = 0$)	1	0.006	0	0	1920	−2631	5270	5289
	Unperturbed ($\varepsilon = 0$)	0.8	0	—	0	2330	−2836	5679	5693
PE	Full	0.4	0.005	0.8	0.0001	11520	−1298	2606	2626
	Restricted ($\lambda = 1$)	0.4	0.003	1	0.0001	13130	−1383	2774	2790
	Restricted ($\lambda = 0$)	1	0.003	0	0.00005	3600	−1426	2860	2876
	Unperturbed ($\varepsilon = 0$)	1	0	—	0.00005	4630	−1518	3042	3054
2P	Full	0.2	0.003	0.1	0.0003	28470	−851	1713	1733
	Restricted ($\lambda = 1$)	0.1	0.001	1	0.0002	45570	−1371	2750	2766
	Restricted ($\lambda = 0$)	0.1	0.006	0	0.0002	11900	−874	1757	1773
	Unperturbed ($\varepsilon = 0$)	0.1	0	—	0.00025	47990	−1398	2802	2814

Note: The estimated parameter values for the full model maximize the LL (9) over the following sets: $\alpha \in \{0.1m \mid m = 0, 1, ..., 10\}$, $\varepsilon \in \{0.001m \mid m = 0, 1, ..., 10\}$, $\lambda \in \{0.1m \mid m = 0, 1, ..., 10\}$, $r \in \{0.00005m \mid m = -6, -5, ..., 6\}$ and $t \in \{0, \tau, 2\tau, ..., 5000\tau\}$ given the step size $\tau = 10$ and the decay rate $\eta = 0.001$. The estimates for each restricted model maximize the LL over the associated subsets. The table also displays LL, AIC and BIC values.

5.2 Predicted Dynamics

Using the estimated parameters of the full models, we now analyze predicted dynamics. For each treatment, let us call the estimated time $\hat{t}$ in Table 6 the *medium run* and its fiftyfold $50\hat{t}$ the *long run*. The long run is so long that the computed solution path to the dynamics (8) given the other estimates $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda}, \hat{r})$ in Table 6 achieves enough convergence. This distinction between time spans follows Binmore and Samuelson (1999, p. 367).

Figures 7 and 8 demonstrate observed and estimated bidding behavior in the medium run and predicted behavior in the long run. We see from these figures that the medium-run state $\sigma(\hat{t})$ fits the data well. Moreover, the predicted time series shown by Figure 9 are qualitatively similar to the observed series in Figure 4. It should be emphasized that we did not use the bidding data in early rounds to obtain the estimates in Table 6. As an exercise, we then estimated time t at which the LL for bidding data in rounds 1–15 is maximized given the estimates $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda}, \hat{r})$ in Table 6. As explained in the note of Figure 9, the estimated time for each treatment is prior to the medium run $\hat{t}$. We see from this figure that the predicted time series (lines) also explain the data in rounds 1–15 (hollow symbols)

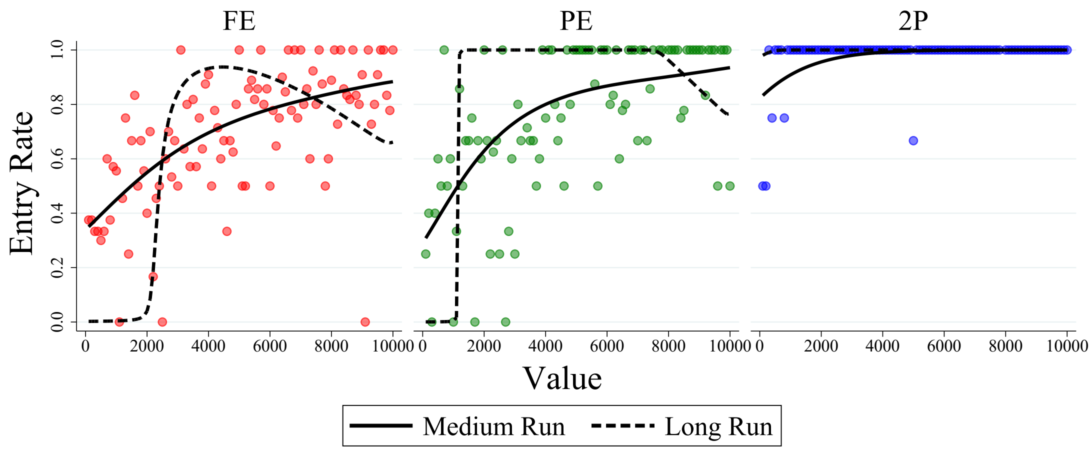


Figure 7: Entry Rates by Value. *Note*: Circles represent observed entry rates in rounds 16–20. The solid and dashed curves for each treatment show estimated medium-run rates (at time $\hat{t}$ in Table 6) and predicted long-run rates (at time $50\hat{t}$) respectively.

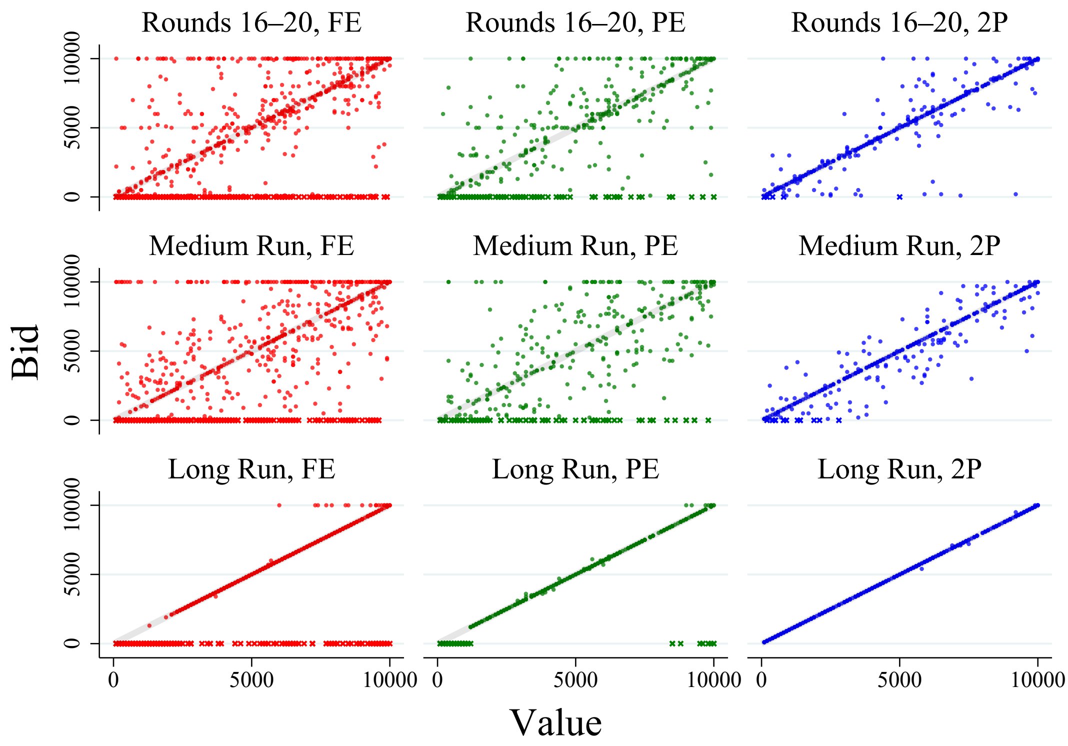
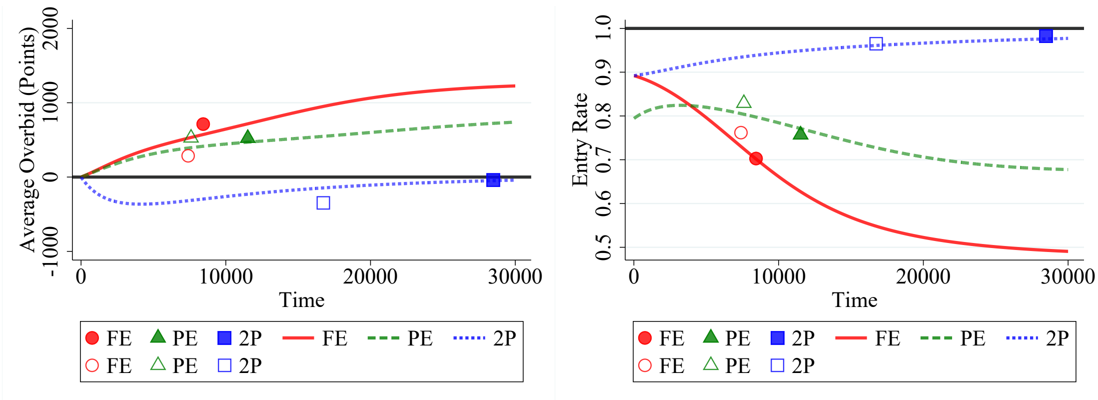


Figure 8: Bids by Value. *Note*: For each treatment, the scatter plot at the top shows observed bids in rounds 16–20. The plot at the middle shows simulated bids drawn from the medium-run distributions $\sigma_v(\hat{t}) \in \Delta(B)$ at the estimated time $\hat{t}$ in Table 6 (given the realized values v in rounds 16–20). The plot at the bottom shows simulated bids drawn from the long-run distributions $\sigma_v(50\hat{t})$ (given the realized values v in rounds 16–20).



(a) Average Overbid Amount Given Entry

(b) Entry Rate

Figure 9: Predicted Time Series of Average Overbid Amounts Given Entry and Entry Rates. *Note*: The lines for each treatment show time series predicted by the solution path σ to the perturbed replicator dynamics (8) given the full-model estimates $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda}, \hat{r})$ in Table 6. The circle, triangle and square, located at the estimated time 8440 (FE), 11520 (PE) and 28470 (2P) respectively, represent observed averaged data in rounds 16–20. The hollow circle, triangle and square, located at the estimated time 7400 (FE), 7600 (PE) and 16740 (2P) respectively, represent observed averaged data in rounds 1–15.

fairly well. The result of this exercise validates to some extent our model of the perturbed replicator dynamics with vanishing trends.

The medium-run states $\sigma(\hat{t})$ for FE and PE are not close to the ESSs given the CARA estimate $\hat{r} = 0.0001$ (Figures 1 and 2). First, the entry-rate curves are more "flattened" as shown by Figure 7. Second, there are many extreme overbids as shown at the middle of Figure 8. These two results are caused by the exogenous trends—with the CM lotteries. To illustrate, let us consider the subpopulation of $v = 6000$. The trend of overbidding enhances "entry deterrence," imposing the severe lottery penalty $l(10000)$ on the entire population. The maximum bid thus keeps spreading. Overbidding however hurts the subpopulation itself, and hence, selection fosters opt-out. In the subpopulation of a low value, say, $v = 100$, the high entry rates (0.3–0.4) are mainly due to the trend of value bidding. Extreme overbidding is too costly for low-value bidders.

What are the causes of these trends toward overbidding and value bidding in the first place? The trend of value bidding might be created by ascending-clock presentation used in our experiment. In FE and PE, this presentation might induce some subjects to naïvely follow value bidding with *full entry*. We now find several possible reasons for the trend of overbidding. The first one is strategic. As already mentioned, overbidding can enhance entry deterrence in FE and PE. The negative effects of the last opponent's bid in Table 5 provide evidence for this argument. It is likely that subjects who counted overbidding as effective in entry deterrence were inclined to do so. However, overbidding might not contribute their profits as suggested by the negative average profits of overbidding clusters (Figures 5 and 6a). The second reason is behavioral. The positive effects of the last opponent's bid in Table 3 suggest (indirect) reciprocal behavior within each group in FE and PE. Although we used the amended random payment scheme to control spite, some subjects might have motives for causing their opponents anxiety by overbidding. Moreover, the regression results in Tables 3 and 5 suggest that subjects' bidding and entry decisions in 2P are not affected by the last opponents' bid amounts. This clear distinction between the treatments implies that the CM lotteries played a key role for the observed trend of overbidding, irrespective of whether the underlying reason is strategic or behavioral.

These trends virtually vanish in the long run. As shown by Figures 7 and 8, the predicted long-run states in FE and PE are close to the partial-entry equilibria given the CARA estimate $\hat{r} = 0.0001$ (Figures 1 and 2). This suggests that the middle-run state $\sigma(\hat{t})$ for each treatment lies in the basin of attraction of this partial-entry equilibrium in the original dynamics (5), albeit the basin of attraction of the full-entry equilibrium in PE is relatively large. The seller's ex-ante revenues in those partial-entry equilibria are 4143 in FE and 3979 in PE (Table 1). These revenues are less than 4179, which is the revenue in the 2P auction with the optimal reserve price (Section 3.1). These underperformances give some clues for why CM auctions are not used in practice.

6 Discussion

We conducted a controlled laboratory experiment to investigate whether CM auctions perform as predicted by the theory, and found evidence against it for the full surplus extraction. The perturbed replicator dynamics with vanishing trends predict that, in the long run for FE treatment, the seller will earn slightly less than in the 2P auction with the optimal reserve price. While we considered a single joint distribution of values in our experiment, examining how the degree (or sign) of value correlation changes subjects' behavior in FE auctions would be a fruitful future research. On one hand, if the correlation coefficient ρ is close to one, then bidders have virtually common values. In the limit, the lottery (1) is always zero, and thus, the FE auction is equivalent to the standard 2P auction which may perform well with the aid of ascending-clock presentation. On the other hand, if the correlation coefficient ρ is close to zero, then bidders have virtually independent values. In this case, the lottery (1) imposes very high penalties if the opponent submits high bids, and hence, bidders seldom choose entry in the partial-entry equilibrium characterized by Theorem 1 (i). Therefore, we expect that the seller's revenue is low in this case.

We also argued that the underperformance of the FE auction may be caused by the (evolutionary and asymptotical) instability of the full-entry value-bidding equilibrium. While we proved the instability in the specific environment, we conjecture that this result holds true in more general environments. Although we leave more formal analyses to future research, for expositional purposes, let us suppose that there are two ex-ante symmetric (risk-neutral) bidders, and consider the 2P auction with CM lotteries. Because the full-surplus-extraction lotteries must sometimes give positive prizes (i.e., bonuses) to provide full-entry incentives, a mutant strategy that opts out (opts in) when values are lower (higher) than the lottery prizes for the opponent earns more than the incumbent strategy, as shown by Theorem 2 (i). Hence, it spreads in the single population of bidders.

A Lemmas and Proofs

The proofs of lemmas are provided in Appendix F of SM.

Lemma 1. *A function $l_0 : V \to \mathbb{R}$ satisfies equation (2) for each $v \in V$ if and only if l_0 satisfies equation (1) for each $b \in V$.*

Lemma 2. *Fix any opt-out strategy z for risk-neutral bidders. (i) If $z(v') > 0$ for some $v' > l(v')$, then $U^z(v') > U^z(v)$ for each $v \le l(v)$. (ii) If $z(v) > 0$ for some $v < l(v)$, then $U^z(v) < U^z(v')$ for each $v' \ge l(v')$.*

Lemma 3. *Fix any symmetric equilibrium z for risk-neutral bidders. Then: (i) z is decreasing on $\{v \in V \mid v \ge l(v)\}$. (ii) $z(v^n) < 1$. (iii) If $z(v^m) = 0$ for some $v^m \ge l(v^m)$, then $z(v) = 0$ and $U^z(v) = U^z(v^m) \ge 0$ for each $v > l(v)$. (iv) $U^z(v) \le \pi$ for each $v \in V$.*

Proof of Theorem 1. As a preliminary, we define the threshold $\underline{\pi} > 0$ as the unique solution to the equation

$$\pi = (1-\rho) \sum_{w \le l_0(w)+\pi} p(w)(l_0(w) + \pi - w). \qquad (10)$$

The uniqueness is shown as follows: The right side of (10) is strictly increasing in π. It is continuous in π, and positive at $\pi = 0$. Its left derivative with respect to π at $\pi > 0$ is given by $(1-\rho)\sum_{w-l_0(w)<\pi} p(w) < 1$. Hence, the right and left sides of (10) have a single-crossing point $\underline{\pi} > 0$. We see from equation (4) that equation (10) is the break-even condition $U^z(v) = 0$ for any $v \geq l(v)$ when high values $w > l(w)$ opt in and low values $w \leq l(w)$ opt out.

(i) Suppose $\pi < \underline{\pi}$. We characterize the equilibrium in the statement by using several necessary conditions. Let z be an equilibrium such that $z(v) > 0$ for some $v > l(v)$. Lemma 3 (ii) and (iii) then imply $z(v^n) \in (0,1)$. Lemma 2 (i) also implies that if $v \leq l(v)$, then $U^z(v) < U^z(v^n) = 0$ and $z(v) = 1$. Together with these facts, it follows from Lemma 3 (i) that there exists a cutoff $c \leq n$ with $v^c > l(v^c)$ such that $z(v) = 1$ if $v < v^c$, and $z(v) \in (0,1)$ if $v \geq v^c$. Hence, z is a solution to the following system of n linear equations:

$$z(v') = 1, \tag{11}$$

$$\rho\,(v - l(v))\,z(v) + (1-\rho)\sum_{w\in V} p(w)\,(\min\{v,w\} - l(w))\,z(w) = -\pi \tag{12}$$

for each $v' \in \{v^1, ..., v^{c-1}\}$ and $v \in \{v^c, ..., v^n\}$. Note that equation (12) is equivalent to $U^z(v) = 0$ from equation (4). Also, $U^z(v^{c-1}) \leq 0$ is necessary for preventing the entry of v^{c-1}. In the following 4 steps, we construct a unique equilibrium with such a cutoff c.

Step 1: Fix any integer $c \leq n$ with $v^c > l(v^c)$. We find a unique solution $z^c \in \mathbb{R}^n$ to the system of n linear equations (11) and (12). Let P denote the cumulative distribution function of p. By using the equation

$$\rho(l(v^m) - l(v^{m+1})) = (1-\rho)P(v^m)(v^{m+1} - v^m)$$

for each $m < n$ and subtracting equation (12) for $v = v^{n-1}$ from that for $v = v^n$, we obtain

$$z^c(v^{n-1}) = \frac{\rho(v^{n-1} - l(v^{n-1})) + (v^n - v^{n-1})}{\rho(v^{n-1} - l(v^{n-1}))} z^c(v^n).$$

More generally, a similar calculation yields the following one for each m with

$c \leq m < n$:

$$z^c(v^m) = \frac{\rho(v^m - l(v^m)) + (\rho + (1-\rho)P(v^{m+1}))(v^{m+1} - v^m)}{\rho(v^m - l(v^m))} z^c(v^{m+1}) + \frac{(1-\rho)(v^{m+1} - v^m)}{\rho(v^m - l(v^m))} \sum_{k \geq m+2} p(v^k) z^c(v^k). \tag{13}$$

Note that the coefficient on $z^c(v^{m+1})$ is greater than one and that on $z^c(v^k)$ is positive since $v^m > l(v^m)$. Hence, z^c is strictly decreasing on $\{v \mid v \geq v^c\}$ if $z^c(v^n) > 0$, and increasing on $\{v \mid v \geq v^c\}$ if $z^c(v^n) \leq 0$.

Proceeding recursively, we can represent $z^c(v^m)$ for each $m \geq c$ as a linear function of $z^c(v^n)$ with a positive slope. We then substitute these values $(z^c(v^c), z^c(v^{c+1}), ..., z^c(v^{n-1}))$ (with $z^c(v) = 1$ for each $v < v^c$) into equation (12) for $v = v^n$. In the resulting equation, the coefficient on $z^c(v^n)$ is positive. Solving this equation with respect to $z^c(v^n)$ yields a unique solution $z^c \in \mathbb{R}^n$ given the cutoff c.

For any pair of two vectors $z^c, z^{c-1} \in \mathbb{R}^n$ characterized in this way, we denote $\zeta^c(v) := z^c(v) - z^{c-1}(v)$ for each $v \in V$ and use the next equation:

$$\begin{aligned} &U^{z^c}(v) - U^{z^{c-1}}(v) \\ &= \rho\,(v - l(v))\,\zeta^c(v) + (1-\rho) \sum_{w \geq v^{c-1}} p(w)\,(\min\{v, w\} - l(w))\,\zeta^c(w), \end{aligned} \tag{14}$$

where the equality follows from $z^c(w) = z^{c-1}(w) = 1$ for each $w < v^{c-1}$.

Step 2: We find an integer $c \leq n$ that satisfies $v^c > l(v^c)$, $z^c(v^c) \in (0,1)$ and $U^{z^c}(v^{c-1}) \leq 0$. First, we claim that $v^c > l(v^c)$ implies $z^c(v^c) > 0$. Indeed, if $z^c(v^c) \leq 0$, then $z^c(v^c) \leq \cdots \leq z^c(v^n) \leq 0$ from step 1, and $U^{z^c}(v^n) < 0$ from equations (4) and (10) with $z^c(v^1) = \cdots = z^c(v^{c-1}) = 1$ and $\pi < \underline{\pi}$. This contradicts equation (12) for $v = v^n$ (i.e., $U^{z^c}(v^n) = 0$).

Second, we claim that the lowest integer c with $v^c > l(v^c)$ satisfies $U^{z^c}(v^{c-1}) < 0$. Indeed, our first claim with step 1 then implies that $z^c(v)$ is positive for each $v \geq v^c$. Hence, equation (4) with $(v^{c-1} - l(v^{c-1}))z^c(v^{c-1}) \leq 0$ implies $U^{z^c}(v^{c-1}) < U^{z^c}(v^c) = 0$.

Third, we claim that the highest integer c with $U^{z^c}(v^{c-1}) \leq 0$ is a required integer. Since this integer c satisfies $v^c > l(v^c)$ and $z^c(v^c) > 0$ from our second and first claims, it remains to show $z^c(v^c) < 1$. To derive a contradiction, we suppose $z^c(v^c) \geq 1$. This occurs only if $c < n$ from equations (3) and (12). Then, $\zeta^{c+1}(v^c) = 1 - z^c(v^c) \leq 0$. Hence, $\zeta^{c+1}(w) \geq 0$ for each $w \geq v^{c+1}$ because if not, then $\zeta^{c+1}(w) < 0$ for each $w \geq v^{c+1}$ from step 1, and $0 = U^{z^{c+1}}(v^n) < U^{z^c}(v^n) = 0$ from equations (12) and (14), which is a contradiction. Using these inequalities for ζ^{c+1} with (14), we obtain

$$U^{z^{c+1}}(v^c) = U^{z^{c+1}}(v^c) - U^{z^c}(v^c) \leq U^{z^{c+1}}(v^{c+1}) - U^{z^c}(v^{c+1}) = 0.$$

This contradicts the hypothesis that c is the highest integer with $U^{z^c}(v^{c-1}) \leq 0$.

Step 3: We show that the vector z^c in step 1 given the cutoff c in step 2 is a partial-entry equilibrium. The vector z^c is indeed an opt-out strategy because it follows from step 1 with $z^c(v^c) \in (0,1)$ that $z^c(v) \in (0,1)$ if $v \geq v^c$, and $z^c(v) = 1$ if $v < v^c$. Moreover, any bidder has no incentive to deviate from z^c, because $U^{z^c}(v) = 0$ for each $v \geq v^c$ from equation (12), $U^{z^c}(v^{c-1}) \leq 0$, and $U^{z^c}(v) < 0$ for each $v < v^{c-1}$ from equation (4) with $z^{c-1}(v) = 1$.

Step 4: We show that the auction game has no other equilibrium z that satisfies $z(v) > 0$ for some $v > l(v)$. To show this, fix any equilibrium with this property. As shown in step 1, this equilibrium is characterized as the vector z^m given some cutoff $m \leq n$ such that $v^m > l(v^m)$, $z^m(v^m) \in (0,1)$ and $U^{z^m}(v^{m-1}) \leq 0$. Since the cutoff c in step 2 is defined as the highest integer with $U^{z^c}(v^{c-1}) \leq 0$, we have $m \leq c$.

To derive a contradiction, suppose $m < c$. Since the latter catoff is higher, $z^{c-1}(v^{c-1}) \leq z^m(v^{c-1}) < 1$ from equation (13). Hence, $\zeta^c(v^{c-1}) = 1 - z^{c-1}(v^{c-1}) > 0$ and $\zeta^c(w) < 0$ for each $w \geq v^c$. Together with these inequalities, equation (14) implies

$$U^{z^c}(v^{c-1}) = U^{z^c}(v^{c-1}) - U^{z^{c-1}}(v^{c-1}) > U^{z^c}(v^c) - U^{z^{c-1}}(v^c) = 0,$$

which contradicts $U^{z^c}(v^{c-1}) \leq 0$. Therefore, $m = c$ and $z^m \equiv z^c$.

Finally, we show that if $\pi = 0$ (and therefore $l = l_0$), then this equilibrium z^c is the unique partial-entry equilibrium. To show this, fix any equilibrium z. If $z(v^n) > 0$, then $z \equiv z^c$ from the former result. Next, we consider the case of $z(v^n) = 0$. Lemma 3 (iii) then implies that $z(v) = 0$ if $v > l_0(v)$. We thus obtain

$$U^z(v^n) = (1-\rho) \sum_{w < l_0(w)} p(w)(w - l_0(w))z(w) \geq 0.$$

Hence, $z(w) = 0$ if $w < l_0(w)$. If $z(w) > 0$ for $w = l_0(w)$, then $U^z(v^1) = (1-\rho)p(w)(v^1 - l_0(w))z(w) < 0$, which contradicts $U^z(v^1) \geq 0$. We thus conclude that $z \equiv 0$ (i.e., the full-entry equilibrium) in this case.

(ii) Suppose $\pi \geq \underline{\pi}$. Fix any equilibrium z. To derive a contradiction, suppose $z(v) > 0$ for some $v > l(v)$. Then, $U^z(v) \leq 0$ since z is an equilibrium. Now, since $\pi \geq \underline{\pi}$, the premium π is weakly higher than the right side of equation (10). Hence, equation (4) with $l = l_0 + \pi$ and $(v - l(v))z(v) > 0$ implies $U^z(v) > 0$, which contradicts $U^z(v) \leq 0$. □

Proof of Theorem 2. (i) Suppose $\pi = 0$. Let us consider a pure opt-out strategy z such that $z(v) = 1$ for some $v < l_0(v)$ and $z(v) = 0$ for each $v \geq l_0(v)$. Equation (4) implies that a bidder with any value v incurs a loss $U^z(v) < 0$ from value bidding when matching with this pure strategy z. The replicator dynamics (5) then imply that, on the subset $\{\sigma \in \Delta(B)^V \mid \sigma(v,v) + \sigma(v,0) = 1 \text{ and } \sigma(v,v) \in (0,1) \text{ if } v < l_0(v), \sigma(v,v) = 1 \text{ if } v \geq l_0(v)\}$ of states, the proportion of value-bidders in the subpopulation of any value $v < l_0(v)$ is strictly decreasing because $\dot{\sigma}(v,v) = p(v)U^\sigma(v,v)(1-\sigma(v,v))\sigma(v,v) < 0$. Thus, the full-entry equilibrium is asymptotically unstable in the dynamics (5).

(ii) Suppose $\pi > 0$. Let z denote the full-entry equilibrium. Since a risk-neutral bidder with any value v obtains the positive interim profit $U^z(v) = \pi > 0$, value bidding is the unique best response to itself. The full-entry equilibrium is thus evolutionarily stable.

Let σ denote the full-entry equilibrium, and fix any mutant $\sigma' \neq \sigma$. Let $\varepsilon_{\sigma'}(\pi) \in (0,1]$ be the highest invasion barrier of σ against σ'. For each $\varepsilon \in (0, \varepsilon_{\sigma'}(\pi))$, the terms of π in the utilities $U(\sigma, \varepsilon\sigma' + (1-\varepsilon)\sigma)$ for the incumbent σ and $U(\sigma', \varepsilon\sigma' + (1-\varepsilon)\sigma)$ for the mutant σ' are given by

$$
\begin{aligned}
&((1-\varepsilon) + \varepsilon E[1 - \sigma'(\tilde{v}_j, 0)])\,\pi, \\
&((1-\varepsilon)E[1 - \sigma'(\tilde{v}_i, 0)] + \varepsilon E[(1 - \sigma'(\tilde{v}_i, 0))(1 - \sigma'(\tilde{v}_j, 0))])\,\pi
\end{aligned}
$$

respectively. Since the former is not lower than the latter, the barrier $\varepsilon_{\sigma'}$ is increasing in π.

We now consider the mutant σ' such that $\sigma'(v,0) = 1$ for each $v < l_0(v)$ and $\sigma'(v,v) = 0$ for each $v \geq l_0(v)$. With equation (4), some algebra shows that

$$
\varepsilon_{\sigma'}(\pi) = \frac{\Pr[\tilde{v}_i < l_0(\tilde{v}_i)]\pi}{\sum_{v_i < l_0(v_i)} \sum_{v_j < l_0(v_j)} \Pr[\tilde{v}_i = v_i, \tilde{v}_j = v_j]\,(l_0(v_j) + \pi - \min\{v_i, v_j\})}
$$

if the right side is lower than one, and $\varepsilon_{\sigma'}(\pi) = 1$ otherwise. Thus, the highest invasion barrier $\varepsilon_{\sigma'}(\pi)$ of σ against the mutant σ' converges to zero as $\pi \to 0$. □

Proof of Theorem 3. (i) Fix any equilibrium z. If $z(v) > 0$ for some $v > l(v)$, then z is the partial-entry equilibrium z^c from Theorem 1 (i). If $z(v) = 0$ for each $v > l(v)$, then $z(v) = 0 < z^c(v) \in (0,1)$ for each $v \geq v^c > l(v^c) \geq l(v)$ and $z(v) \leq 1 = z^c(v)$ for each $v < v^c$. Thus, z^c is the unique minimal-entry equilibrium.

(ii) By construction, the opt-out strategy z^d is an equilibrium. Fix another equilibrium z. We show that $z(v) \leq z^d(v)$ for each $v \in V$, and hence, z^d is the unique minimal-entry equilibrium. It follows from Theorem 1 (ii) that $z(v) = z^d(v) = 0$ for each $v > l(v)$. Let m be the lowest integer such that $z(v) = 0$ for each $v \geq v^m$. To derive a contradiction, suppose $v^d < v^m$. We now consider the opt-out strategy z' such that $z'(v) = 1$ if $v < v^m$ and $z'(v) = 0$ if $v \geq v^m$. Equation (4) then implies that $U^{z'}(v^1) < \cdots < U^{z'}(v^{m-1}) \leq U^z(v^{m-1}) \leq 0$, where the last inequality follows from $z(v^{m-1}) > 0$. Equation (4) also implies

$$U^{z'}(v^m) = U^{z'}(v^{m+1}) = \cdots = U^{z'}(v^n) = (1-\rho) \sum_{w \leq v^{m-1}} p(w)\,(w - l(w)) + \pi \geq 0,$$

where the last inequality follows from equation (10) with $w \leq v^{m-1} \leq l(v^{m-1}) \leq l(w)$ and $\pi \geq \underline{\pi}$. This contradicts the construction of the strategy z^d with the cutoff v^d.

Suppose now $\pi > \underline{\pi}$. Equation (10) then implies that $U^{z^d}(v) > 0$ for each $v \geq v^d$. Since $U^{z^d}(v) < 0$ for each $v < v^d$ by construction, the minimal-entry equilibrium z^d is strict, and hence, an ESS. □

Lemma 4. *Every ASS z with $z(v) = 0$ for each $v \geq l(v)$ satisfies $z(v) \in \{0,1\}$ for each $v < l(v)$.*

Proof of Claim 1. (i) The minimal-entry equilibrium z^c satisfies $U^{z^c}(v) < 0$ for each $v < v^{c-1}$. This strict inequality also holds for $v = v^{c-1}$ in the experimental setting with $\pi = 1$. This implies that any mutant σ' is a best response to the incumbent σ equivalent to z^c only if $\sigma'(v,0) = 1$ for each $v < v^c$. We now define the quadratic function

$$f(z) := E\left[(1 - z^c(\tilde{v}_i))U^z(\tilde{v}_i)\right] - E\left[(1 - z(\tilde{v}_i))U^z(\tilde{v}_i)\right] = E\left[(z(\tilde{v}_i) - z^c(\tilde{v}_i))U^z(\tilde{v}_i)\right]$$

on $Z := \{z \in [0,1]^V \mid z(v) = 1\ \forall v < v^c\}$. The minimal-entry equilibrium z^c is an ESS if and only if $f(z) > 0$ for each mutant opt-out strategy $z \in Z$ with $z \neq z^c$.

Since z^c is an equilibrium with $z^c(v) \in (0,1)$ for each $v \geq v^c$, the first-order condition $\frac{\partial f}{\partial z_v}(z^c) = 0$ is satisfied for each $v \geq v^c$, where $z_v := z(v)$. Let H denote the Hessian matrix of f with respect to $(v^c, ..., v^n)$. The diagonal and off-diagonal elements of H are given by

$$\frac{\partial^2 f}{\partial z_v^2} = p(v)p(v)(1-\rho)(2v - 2l(v)) + p(v)\rho(2v - 2l(v))$$
$$\frac{\partial^2 f}{\partial z_v \partial z_w} = p(v)p(w)(1-\rho)(2\min\{v, w\} - l(v) - l(w))$$

for each $v, w \geq v^c$ with $v \neq w$, respectively. If z^c minimizes f, then H should be positive semi-definite. We can indeed show that H is positive definite, and hence, f is strictly convex on Z. From Sylvester's criterion (Gilbert, 1991), the symmetric matrix H is positive definite if and only if all principal minors of H are positive. We can show by direct computations that H satisfies the latter condition in the experimental setting with $\pi = 1$. Thus, $f(z) > 0$ for each $z \in Z$ with $z \neq z^c$, so that z^c is an ESS.

(ii) Fix another partial-entry equilibrium $z \neq z^c$. Theorem 1 (i) with $\pi < \underline{\pi}$ implies that $z(v) = 0$ for each $v > l(v)$. Since FE treatment has no value v with $v = l(v)$, we can find a value $v < l(v)$ with $z(v) > 0$. To derive a contradiction, suppose that z is an ASS. Lemma 4 then implies $z(v) = 1$. A bidder with value v^1 should then opt out because

$$U^z(v^1) \leq (1-\rho)p(v)(v^1 - l(v)) + \pi = \frac{1}{150}(100 - l_0(v)) + \frac{149}{150} < 0, \tag{15}$$

where the first inequality follows from equation (4), and the last inequality from $l_0(v) \geq l_0(1800) > 1800$. A bidder with value v^n should then opt out because

$$U^z(v^n) \leq (1-\rho)p(v^1)(v^1 - l(v^1)) + \pi < 0,$$

where the last inequality follows from (15) with $p(v) = p(v^1)$ and $l(v) \leq l(v^1)$. This contradicts $z(v^n) = 0$. □

References

Bartling, Björn and Nick Netzer (2016): "An Externality-Robust Auction: Theory and Experimental Evidence," *Games and Economic Behavior*, 97, 186–204.

Binmore, Ken and Larry Samuelson (1999): "Evolutionary Drift and Equilibrium Selection," *The Review of Economic Studies*, 66 (2), 363–393.

Börgers, Tilman (2015): *An Introduction to the Theory of Mechanism Design*, Oxford University Press, USA.

——— (2017): "(No) Foundations of Dominant-Strategy Mechanisms: A Comment on Chung and Ely (2007)," *Review of Economic Design*, 21 (2), 73–82.

Börgers, Tilman and Rajiv Sarin (1997): "Learning through Reinforcement and Replicator Dynamics," *Journal of Economic Theory*, 77 (1), 1–14.

Breitmoser, Yves and Sebastian Schweighofer-Kodritsch (2022): "Obviousness around the Clock," *Experimental Economics*, 25 (2), 483–513.

Brusco, Sandro (1998): "Unique Implementation of the Full Surplus Extraction Outcome in Auctions with Correlated Types," *Journal of Economic Theory*, 80 (2), 185–200.

Carroll, Gabriel (2019): "Robustness in Mechanism Design and Contracting," *Annual Review of Economics*, 11 (1), 139–166.

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016): "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Chen, Yi-Chun and Siyang Xiong (2013a): "Genericity and Robustness of Full Surplus Extraction," *Econometrica*, 81 (2), 825–847.

——— (2013b): "Supplement to "Genericity and Robustness of Full Surplus Extraction"," *Econometrica*, https://www.econometricsociety.org/ecta/Supmat/10123_proofs.pdf.

Chung, Kim-Sau and Jeffrey C. Ely (2007): "Foundations of Dominant-Strategy Mechanisms," *The Review of Economic Studies*, 74 (2), 447–476.

Cooper, David J. and Hanming Fang (2008): "Understanding Overbidding in Second Price Auctions: An Experimental Study," *The Economic Journal*, 118 (532), 1572–1595.

Crémer, Jacques and Richard P. McLean (1985): "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent," *Econometrica*, 53 (2), 345–361.

——— (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56 (6), 1247–1257.

Eső, Péter (2005): "An Optimal Auction with Correlated Values and Risk Aversion," *Journal of Economic Theory*, 125 (1), 78–89.

Fréchette, Guillaume R., Alessandro Lizzeri, and Jacopo Perego (2022): "Rules and Commitment in Communication: An Experimental Analysis," *Econometrica*, 90 (5), 2283–2318.

Frederick, Shane (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19 (4), 25–42.

Fudenberg, Drew and David K. Levine (1998): *The Theory of Learning in Games*, vol. 2, MIT press.

Gale, John, Kenneth G. Binmore, and Larry Samuelson (1995): "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior*, 8 (1), 56–90.

Gilbert, George T. (1991): "Positive Definite Matrices and Sylvester's Criterion," *The American Mathematical Monthly*, 98 (1), 44–46.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey (2016): *Quantal Response Equilibrium: A Stochastic Theory of Games*, Princeton University Press.

Greiner, Ben (2015): "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE," *Journal of the Economic Science Association*, 1 (1), 114–125.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009): *The Elements of Statistical Learning: Prediction, Inference and Data Mining.*

Heifetz, Aviad and Zvika Neeman (2006): "On the Generic (Im)possibility of Full Surplus Extraction in Mechanism Design," *Econometrica*, 74 (1), 213–233.

Holt, Charles A. and Susan K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92 (5), 1644–1655.

Kagel, John H., Ronald M. Harstad, and Dan Levin (1987): "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study," *Econometrica*, 55 (6), 1275–1304.

Kagel, John H. and Dan Levin (2016): "Auctions: A Survey of Experimental Research," in *The Handbook of Experimental Economics, Volume 2*, ed. by John H. Kagel and Alvin E. Roth, Princeton University Press, 563–637.

Krajbich, Ian, Colin Camerer, John Ledyard, and Antonio Rangel (2009): "Using Neural Measures of Economic Value to Solve the Public Goods Free-Rider Problem," *Science*, 326 (5952), 596–599.

Krajbich, Ian, Colin Camerer, and Antonio Rangel (2017): "Exploring the Scope of Neurometrically Informed Mechanism Design," *Games and Economic Behavior*, 101, 49–62.

Laury, Susan K. and Charles A. Holt (2005): "Further Reflections on Prospect Theory," *Andrew Young School of Policy Studies Research Paper Series*, (06-11).

Li, Shengwu (2017): "Obviously Strategy-Proof Mechanisms," *American Economic Review*, 107 (11), 3257–3287.

Lim, Wooyoung and Siyang Xiong (2021): "Does Jump Bidding Increase Sellers' Revenue? Theory and Experiment," *Journal of Economic Behavior & Organization*, 189, 84–110.

Macqueen, James (1967): *Some Methods for Classification and Analysis of Multivariate Observations*, University of California Press.

Maskin, Eric and John G. Riley (1980): "Auctioning an Indivisible Object," Tech. rep.

Matsushima, Hitoshi (2007): "Mechanism Design with Side Payments: Individual Rationality and Iterative Dominance," *Journal of Economic Theory*, 133 (1), 1–30.

McAfee, R. Preston and Philip J. Reny (1992): "Correlated Information and Mechanism Design," *Econometrica*, 395–421.

Mezzetti, Claudio (2007): "Mechanism Design with Interdependent Valuations: Surplus Extraction," *Economic Theory*, 31 (3), 473–488.

Milgrom, Paul Robert (2004): *Putting Auction Theory to Work*, Cambridge University Press.

Morgan, John, Ken Steiglitz, and George Reis (2003): "The Spite Motive and Equilibrium Behavior in Auctions," *Contributions in Economic Analysis & Policy*, 2 (1), 1–25.

Myerson, Roger B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6 (1), 58–73.

Nishimura, Naoko, Timothy N. Cason, Tatsuyoshi Saijo, and Yoshikazu Ikeda (2011): "Spite and Reciprocity in Auctions," *Games*, 2 (3), 365–411.

Nishimura, Takeshi (2022): "An Experimental Study of Crémer–McLean Auctions," *Available at SSRN 4111200*.

Noda, Shunya (2019): "Full Surplus Extraction and Within-Period Ex Post Implementation in Dynamic Environments," *Theoretical Economics*, 14 (1), 39–69.

Pham, Hien and Takuro Yamashita (2024): "Auction Design with Heterogeneous Priors," *Games and Economic Behavior*, 145, 413–425.

Rietz, Thomas A. (1993): "Implementing and Testing Risk-Preference-Induction Mechanisms in Experimental Sealed-Bid Auctions," *Journal of Risk and Uncertainty*, 7, 199–213.

Riordan, Michael H. and David E.M. Sappington (1988): "Optimal Contracts with Public Ex Post Information," *Journal of Economic Theory*, 45 (1), 189–199.

Robert, Jacques (1991): "Continuity in Auction Design," *Journal of Economic Theory*, 55 (1), 169–179.

Romero, Julian and Yaroslav Rosokha (2023): "Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma," *Econometrica*, 91 (6), 2295–2331.

Roth, Alvin E. and Michael W. Malouf (1979): "Game-Theoretic Models and the Role of Information in Bargaining," *Psychological review*, 86 (6), 574.

Schneider, Mark and David Porter (2020): "Effects of Experience, Choice Architecture, and Cognitive Reflection in Strategyproof Mechanisms," *Journal of Economic Behavior & Organization*, 171, 361–377.

Selten, Reinhard (1983): "Evolutionary Stability in Extensive Two-Person Games," *Mathematical Social Sciences*, 5 (3), 269–363.

Selten, Reinhard, Abdolkarim Sadrieh, and Klaus Abbink (1999): "Money Does not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse," *Theory and Decision*, 46 (3), 213–252.

Tan, Charmaine H.Y. (2020): "Overbidding and Matching Rules in Second-Price Auctions: An Experimental Study," *Journal of Behavioral and Experimental Economics*, 84, 101507.

Toplak, Maggie E., Richard F. West, and Keith E. Stanovich (2014): "Assessing Miserly Information Processing: An Expansion of the Cognitive Reflection Test," *Thinking & Reasoning*, 20 (2), 147–168.

Walker, James M., Vernon L. Smith, and James C. Cox (1990): "Inducing Risk-Neutral Preferences: An Examination in a Controlled Market Environment," *Journal of Risk and Uncertainty*, 3, 5–24.

Weibull, Jörgen W. (1997): *Evolutionary Game Theory*, MIT Press.

Supplementary Material for "On the Performance of Crémer–McLean Auction: An Experiment"

Takeshi Nishimura* Nobuyuki Hanaki†

November 12, 2024

B Figures and Tables

Bidder A's value \ Bidder B's value	100	200	· · ·	9900	10000	Total
100	$\frac{51}{15000}$	$\frac{1}{15000}$	· · ·	$\frac{1}{15000}$	$\frac{1}{15000}$	$\frac{1}{100}$
200	$\frac{1}{15000}$	$\frac{51}{15000}$	· · ·	$\frac{1}{15000}$	$\frac{1}{15000}$	$\frac{1}{100}$
· · ·	· · ·	· · ·	· · ·	· · ·	· · ·	· · ·
9900	$\frac{1}{15000}$	$\frac{1}{15000}$	· · ·	$\frac{51}{15000}$	$\frac{1}{15000}$	$\frac{1}{100}$
10000	$\frac{1}{15000}$	$\frac{1}{15000}$	· · ·	$\frac{1}{15000}$	$\frac{51}{15000}$	$\frac{1}{100}$
Total	$\frac{1}{100}$	$\frac{1}{100}$	· · ·	$\frac{1}{100}$	$\frac{1}{100}$	1

Figure 1: Joint Probability Distribution of Values. *Note*: The instruction told subjects that the probability of each cell along the diagonal is 51 out of 15,000, while the probability of any other cell is 1 out of 15,000.

*Faculty of Economics, Hosei University.
†Institute of Social and Economic Research, Osaka University; University of Limassol.

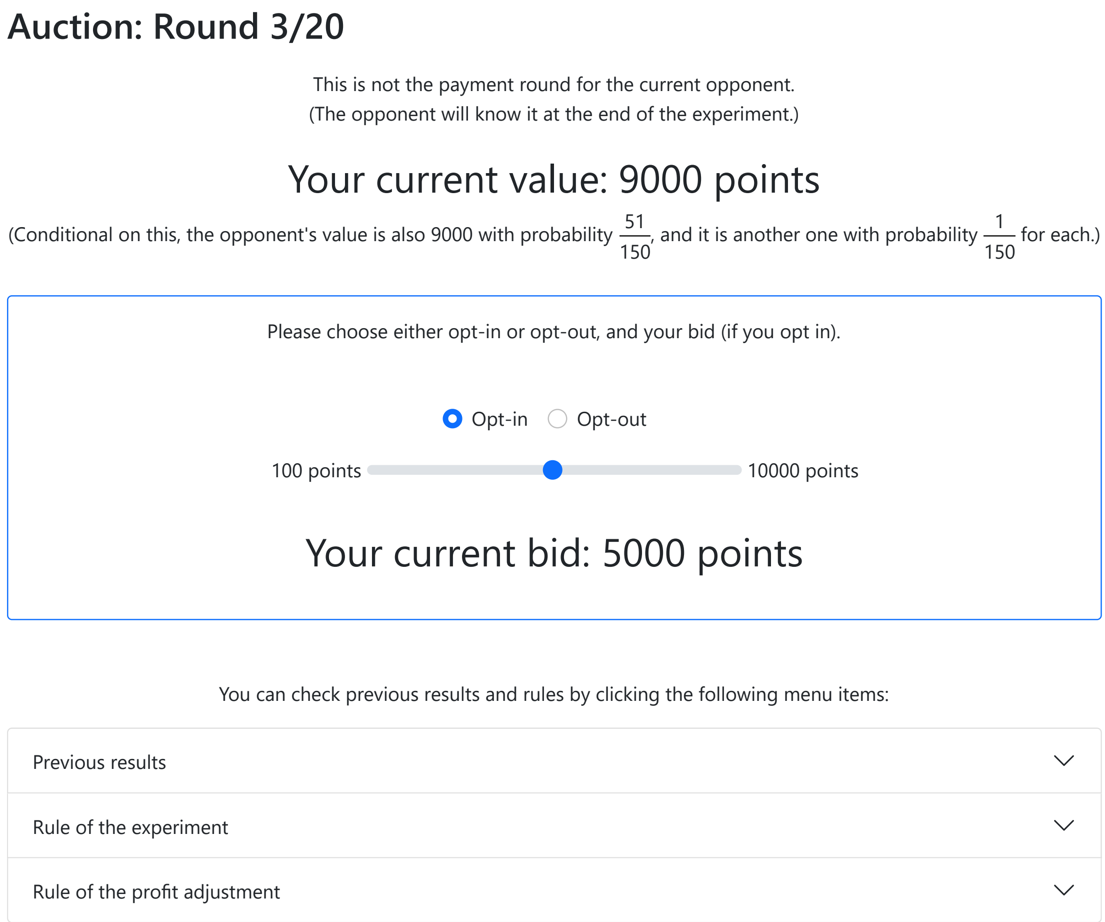


Figure 2: Screenshot of the Bidding Page. *Note*: In 2P, there is no menu item for the rule of the profit adjustment (i.e., CM lottery). The default bid is 5000. The current bid is automatically submitted after 45 seconds pass. In the experiment, the captions on this page were displayed in Japanese.

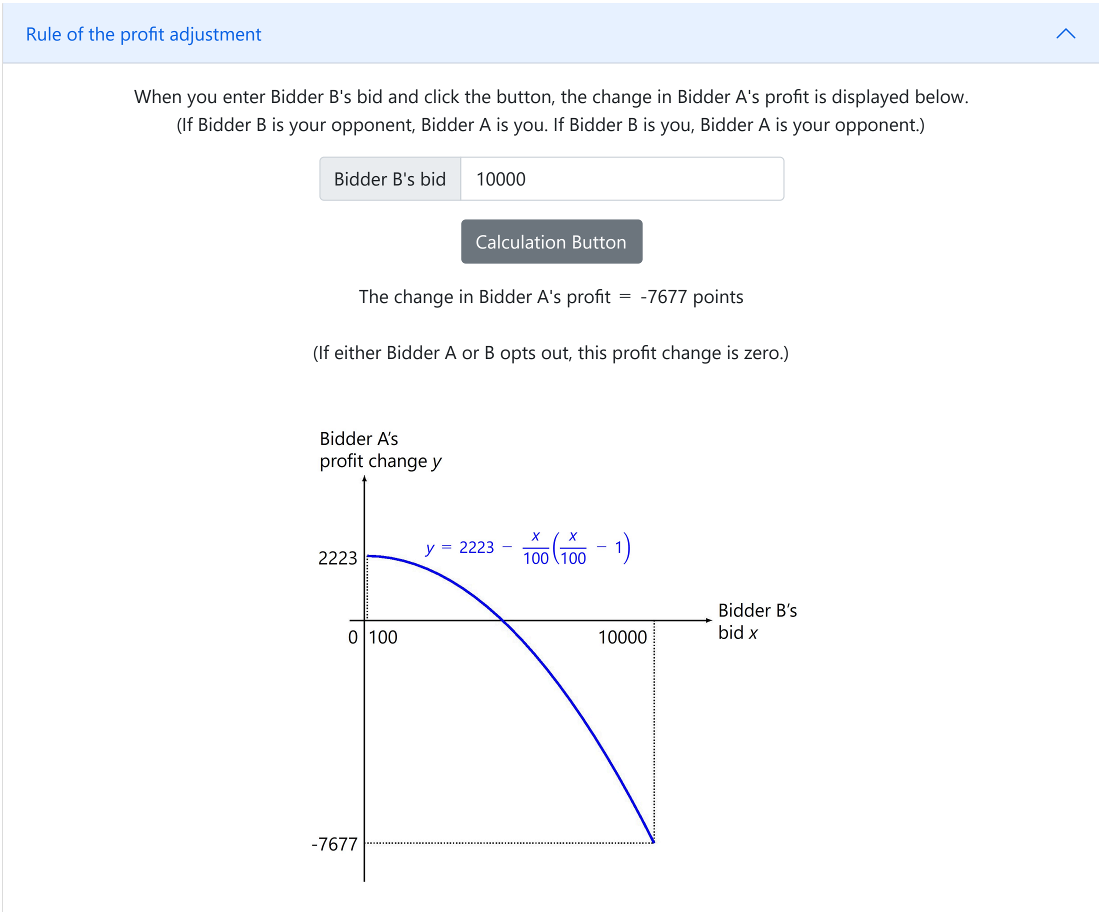


Figure 3: Rule of the Profit Adjustment (CM Lottery). *Note*: In the experiment, the captions on this page were displayed in Japanese.

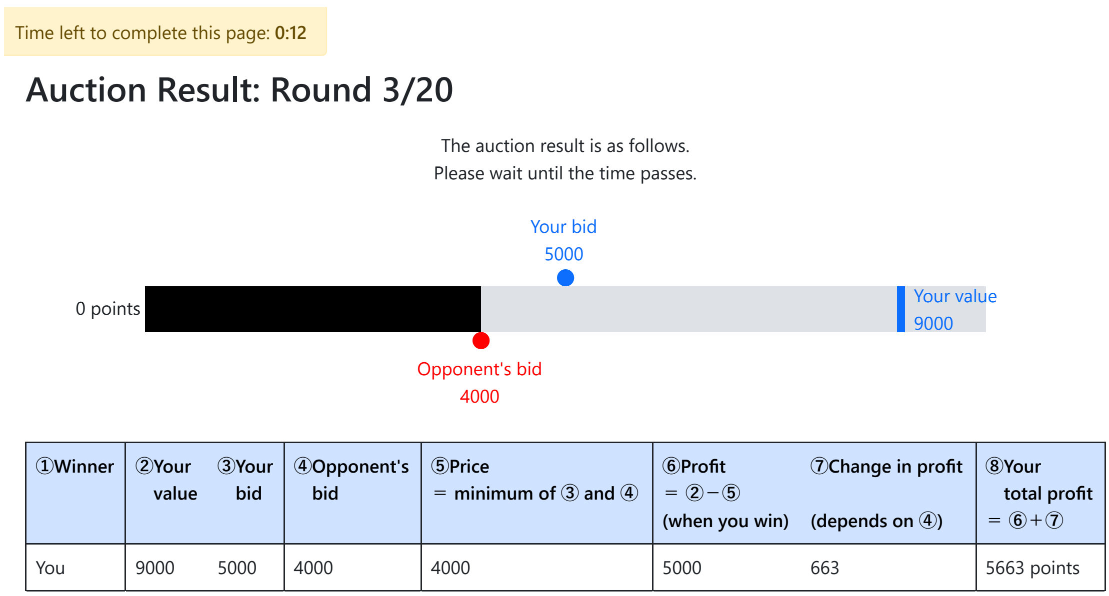


Figure 4: Screenshot of the Result Page. *Note*: In 2P, the last two columns (the change in profit and the total profit) are not displayed. Subjects must wait until 15 seconds pass. In the experiment, the captions on this page were displayed in Japanese.

Table 1: MPL for the elicitation of risk attitudes in the gain and loss domains.

Choice	Lottery A (Safe)				Lottery B (Risky)			
Problem	Prob.	Payoff	Prob.	Payoff	Prob.	Payoff	Prob.	Payoff
1	10%	200	90%	160	10%	385	90%	10
2	20%	200	80%	160	20%	385	80%	10
3	30%	200	70%	160	30%	385	70%	10
4	40%	200	60%	160	40%	385	60%	10
5	50%	200	50%	160	50%	385	50%	10
6	60%	200	40%	160	60%	385	40%	10
7	70%	200	30%	160	70%	385	30%	10
8	80%	200	20%	160	80%	385	20%	10
9	90%	200	10%	160	90%	385	10%	10
10	100%	200	0%	160	100%	385	0%	10
Choice	Lottery A (Safe)				Lottery B (Risky)			
Problem	Prob.	Payoff	Prob.	Payoff	Prob.	Payoff	Prob.	Payoff
1	10%	−200	90%	−160	10%	−385	90%	−10
2	20%	−200	80%	−160	20%	−385	80%	−10
3	30%	−200	70%	−160	30%	−385	70%	−10
4	40%	−200	60%	−160	40%	−385	60%	−10
5	50%	−200	50%	−160	50%	−385	50%	−10
6	60%	−200	40%	−160	60%	−385	40%	−10
7	70%	−200	30%	−160	70%	−385	30%	−10
8	80%	−200	20%	−160	80%	−385	20%	−10
9	90%	−200	10%	−160	90%	−385	10%	−10
10	100%	−200	0%	−160	100%	−385	0%	−10

Note: This table shows the multiple price lists (MPL). Each subject chooses either lottery A or B for all choice problems in 3×2 minutes. At the end of a session, one problem in each domain is randomly selected and the lottery chosen in this problem is drawn by a computer. The unit of payoff is JPY.

Table 2: CARA parameters given the number of safe choices in the MPL.

# of Safe Choices	Range of CARA Parameters	
	Gain Domain	Loss Domain
0	$r < -0.0093$	NA (Irrational)
1	$-0.0093 < r < -0.0055$	$r < -0.0119$
2	$-0.0055 < r < -0.0030$	$-0.0119 < r < -0.0078$
3	$-0.0030 < r < -0.0009$	$-0.0078 < r < -0.0051$
4	$-0.0009 < r < 0.0010$	$-0.0051 < r < -0.0029$
5	$0.0010 < r < 0.0029$	$-0.0029 < r < -0.0010$
6	$0.0029 < r < 0.0051$	$-0.0010 < r < 0.0009$
7	$0.0051 < r < 0.0078$	$0.0009 < r < 0.0030$
8	$0.0078 < r < 0.0119$	$0.0030 < r < 0.0055$
9	$0.0119 < r$	$0.0055 < r < 0.0093$
10	NA (Irrational)	$0.0093 < r$

Note: This table shows the range of constant-absolute-risk-aversion (CARA) parameters given the number of safe lotteries chosen by a decision maker who correctly switches lottery choices at most once in the MPL. The CARA utility function is defined as $u(y) = (1 - \exp(-ry))/r$ (if $r \neq 0$) and $u(y) = y$ (if $r = 0$) for monetary payoffs $y \in \mathbb{R}$.

Table 3: Questions in the CRT.

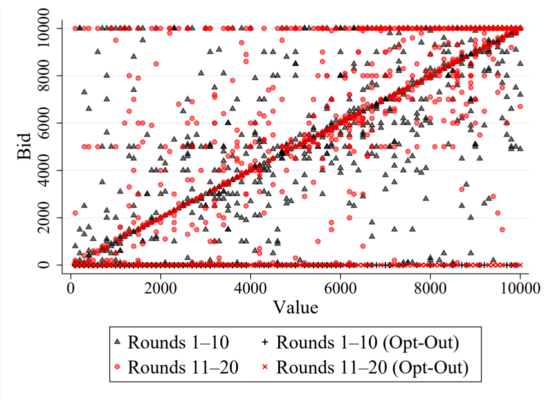
	Question	Answer
1	A book and a pen cost 1,100 yen in total. The book costs 1,000 yen more than the pen. How much does the pen cost?	50 yen
2	If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?	5 minutes
3	In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?	47 days
4	If Taro can drink one barrel of water in 6 days, and Hanako can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together?	4 days
5	Jiro received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?	29 students
6	A man buys a pig for 6,000 yen, sells it for 7,000 yen, buys it back for 8,000 yen, and sells it finally for 9,000 yen. How much has he made?	2,000 yen
7	Saburo decided to invest 800,000 yen in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Saburo, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Saburo has: a. broken even in the stock market, b. is ahead of where he began, c. has lost money.	c

Note: This table shows the cognitive reflection test (CRT). Each subject answers all questions in 6 minutes, obtaining JPY 60 per correct answer. In the experiment, these questions were displayed in Japanese. See Harada et al. (2018) for a Japanese version.

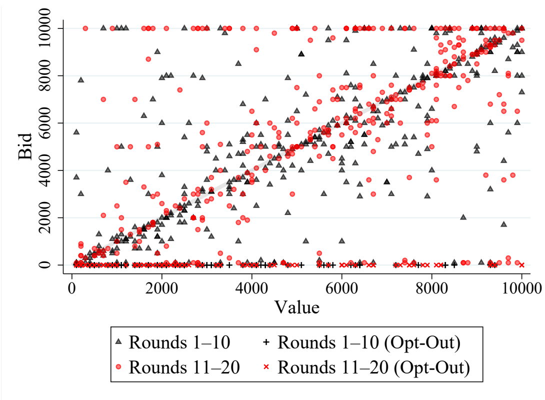
Table 4: Sessions.

Treatment	Session #	Date	# of Groups	# of Subjects	University
FE	1	January 23, 2023	2	12	Osaka
	2	January 24, 2023	4	24	Osaka
	3	January 24, 2023	3	18	Osaka
	4	March 7, 2023	3	18 (22)	Osaka
	5	January 26, 2023	4	24 (25)	Kansai
	6	January 26, 2023	4	24	Kansai
	7	January 30, 2023	3	18 (23)	Kansai
	8	January 30, 2023	4	24	Kansai
PE	1	January 23, 2023	2	12 (15)	Osaka
	2	January 25, 2023	2	12 (17)	Osaka
	3	March 6, 2023	3	18 (22)	Osaka
	7	January 27, 2023	4	24 (25)	Kansai
	8	January 27, 2023	3	18 (22)	Kansai
2P	1	January 26, 2023	3	18 (20)	Osaka
	2	January 26, 2023	2	12 (16)	Osaka
	3	March 6, 2023	3	18 (22)	Osaka
	7	January 27, 2023	2	12	Kansai
	8	January 30, 2023	4	24	Kansai

Note: This table shows the date, the numbers of (human-only) groups and subjects, and the university for each session. The numbers of subjects in parentheses include those who joined groups with computerized bidders. The two sessions with machine troubles mentioned in footnote 16 of the paper are sessions 5 and 6 for FE.

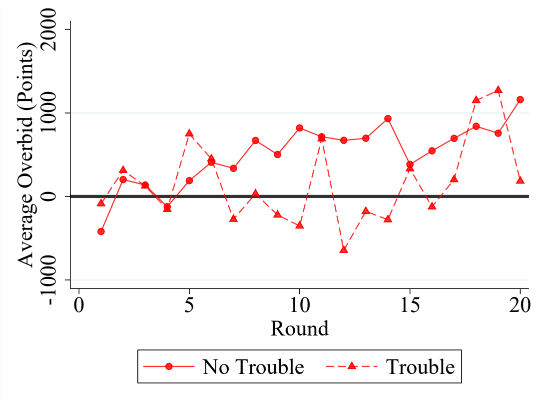


(a) No Trouble

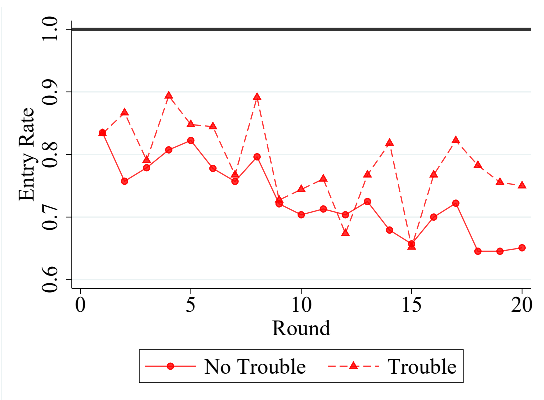


(b) Trouble

Figure 5: Bids by Value in FE. *Note*: The left and right scatter plots show bids by value in the FE sessions without and with the troubles in the instruction (footnote 16 of the paper), respectively.



(a) Average Overbid Amount Given Entry



(b) Entry Rate

Figure 6: Time Series of Average Overbid Amounts Given Entry and Entry Rates in FE. *Note*: The solid and dashed lines illustrate the time series in the FE sessions without and with the troubles in the instruction (footnote 16 of the paper), respectively.

Table 5: Regression models with the trouble dummy for FE.

Dependent Variable	Bid (in Entry Round)	Deviation (in Entry Round)	Entry Dummy × 100
Value / 10^2	74.76 (3.39)		1.17 (0.10)
Value × Value / 10^4			−0.006 (0.0009)
Round	53.62 (17.75)	12.52 (10.94)	−0.81 (0.13)
Last Opponent's Bid / 10^4	693.70 (149.34)	170.38 (118.41)	−8.12 (1.73)
Opponent's Payment Round	−588.78 (405.31)	1153.56 (326.62)	7.14 (3.16)
# of Safe Choices (Gain)			−2.38 (0.94)
# of Safe Choices (Loss)			−1.18 (1.60)
# of Mistakes in Quiz	−71.64 (54.20)	22.21 (29.79)	0.56 (0.54)
CRT Score	54.39 (91.25)	94.93 (52.06)	−0.31 (1.20)
Male	79.45 (328.23)	46.49 (200.42)	−3.03 (3.58)
Trouble in Instruction	−593.96 (431.51)	241.99 (274.99)	4.80 (3.71)
Constant	1517.30 (586.08)	671.57 (339.74)	69.52 (14.48)
Var. of Group Random Effect	285235	53507	1.7 / 10^7
Var. of Subject Random Effect	3739201	1319039	412
Var. of Residual	5477131	3243698	1177
Log Pseudolikelihood	−17452	−20361	−13796
# of Observations	2267	2267	2755

Note: We introduce a trouble dummy (footnote 16 of the paper) into the models in Tables 3–5 of the paper. The three models in this table use data in FE since the trouble occurred only in this treatment. Bid = $b_{it} > 0$. Value = v_{it}. Deviation = $|b_{it} - v_{it}|$ with $b_{it} > 0$. Entry Dummy = $I(b_{it} > 0)$. Round = $t > 1$. Last Opponent's Bid = b_{jt-1}, where subject j is i's opponent in round $t-1$. Opponent's Payment Round = 1 if t is a payment round for the current opponent, and 0 otherwise. Male = 1 if subject i is male, and = 0 otherwise. Trouble in Instruction = 1 if the session has a trouble in the instruction, and = 0 otherwise. Robust standard errors clustered at the group level are shown in parentheses. Var. denotes a variance.

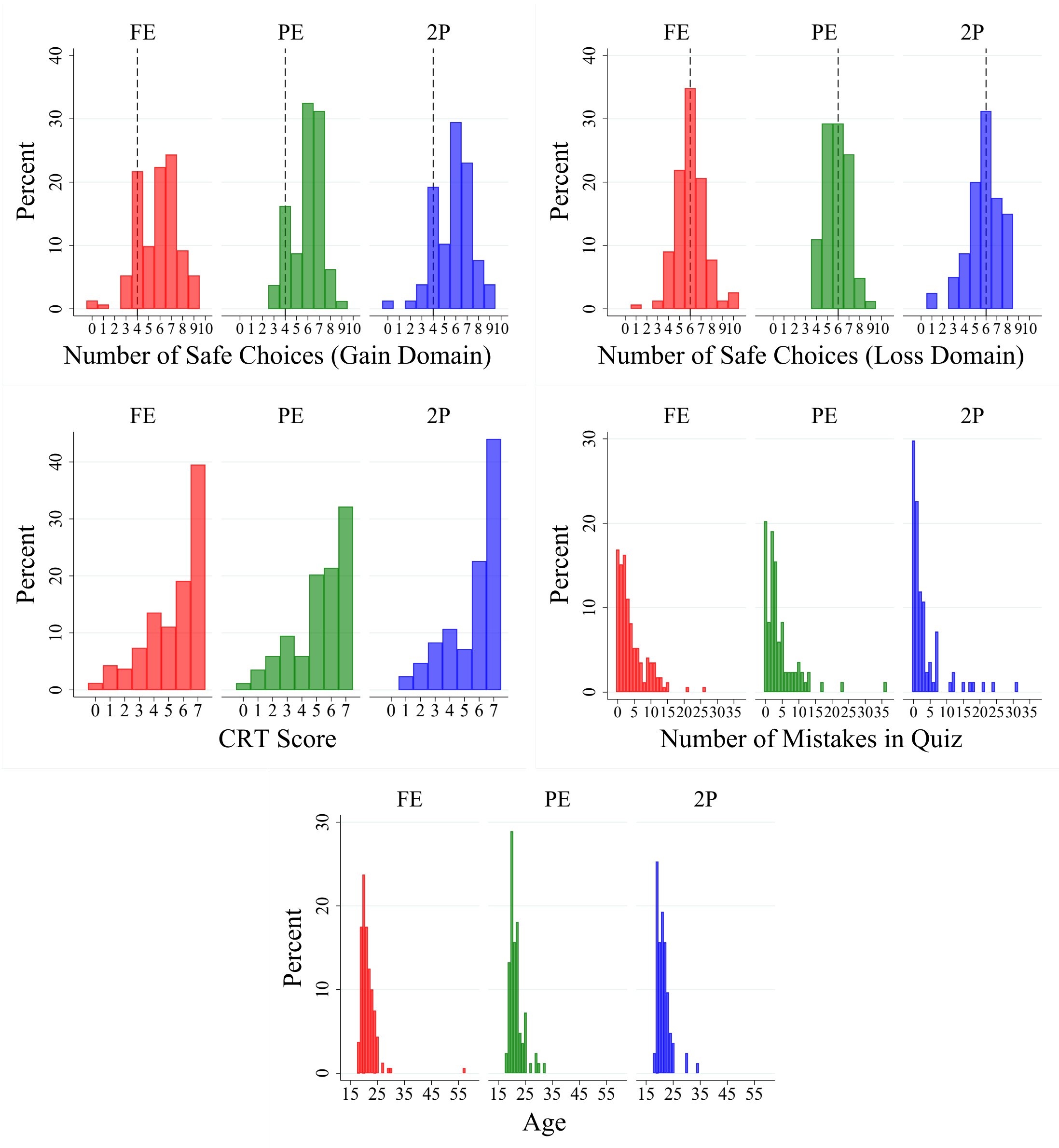


Figure 7: Histograms for the Numbers of Safe Choices in the MPL, CRT Score, the Number of Mistakes in the Comprehension Quiz, and Age. *Note*: In the histograms for the MPL, the dashed lines denote the numbers of safe lotteries chosen by a risk-neutral decision maker (4 and 6 safe choices in the gain and loss domains respectively).

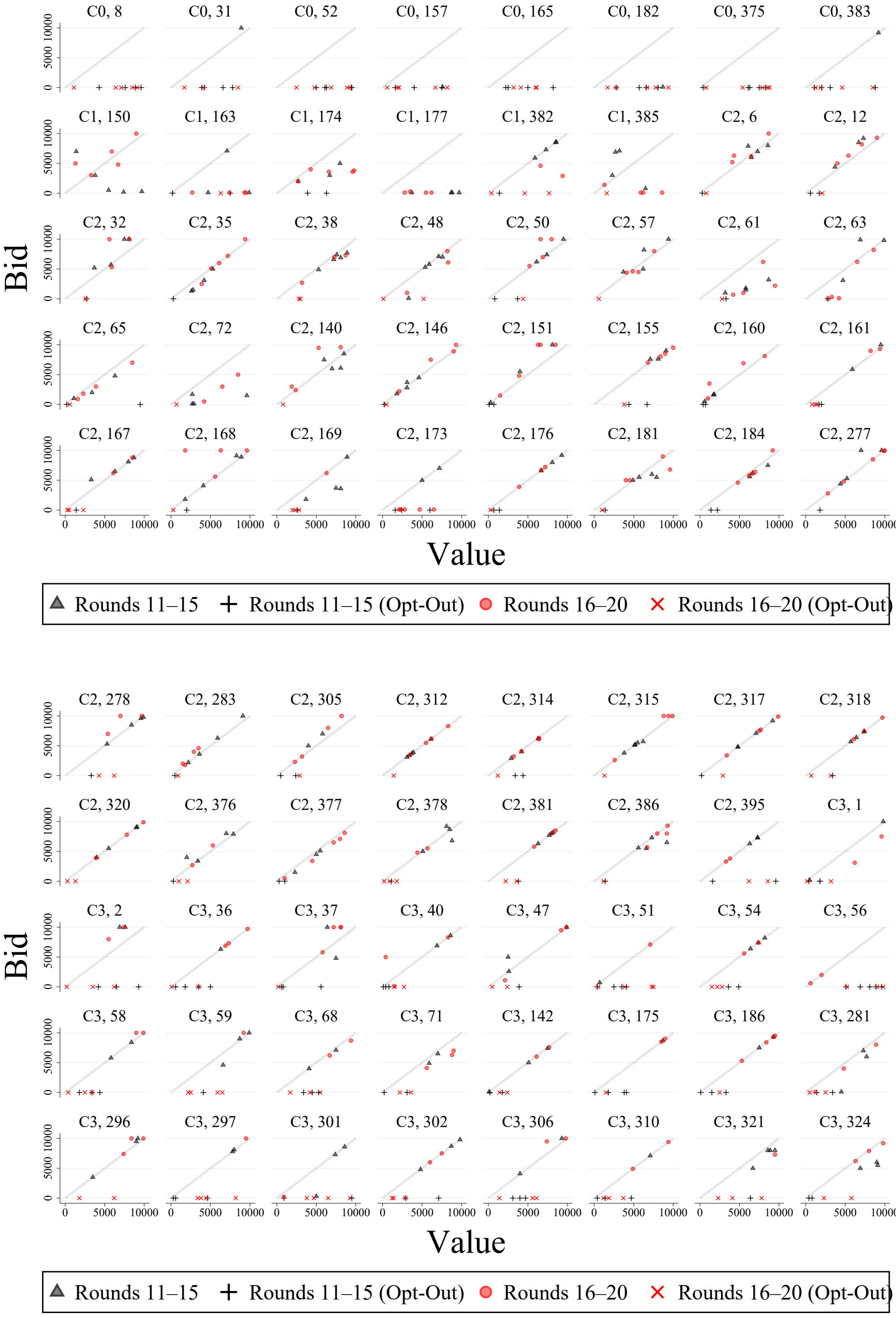


Figure 8: Bids by Subjects in FE. *Note*: Each cluster ID is followed by a participant ID number.

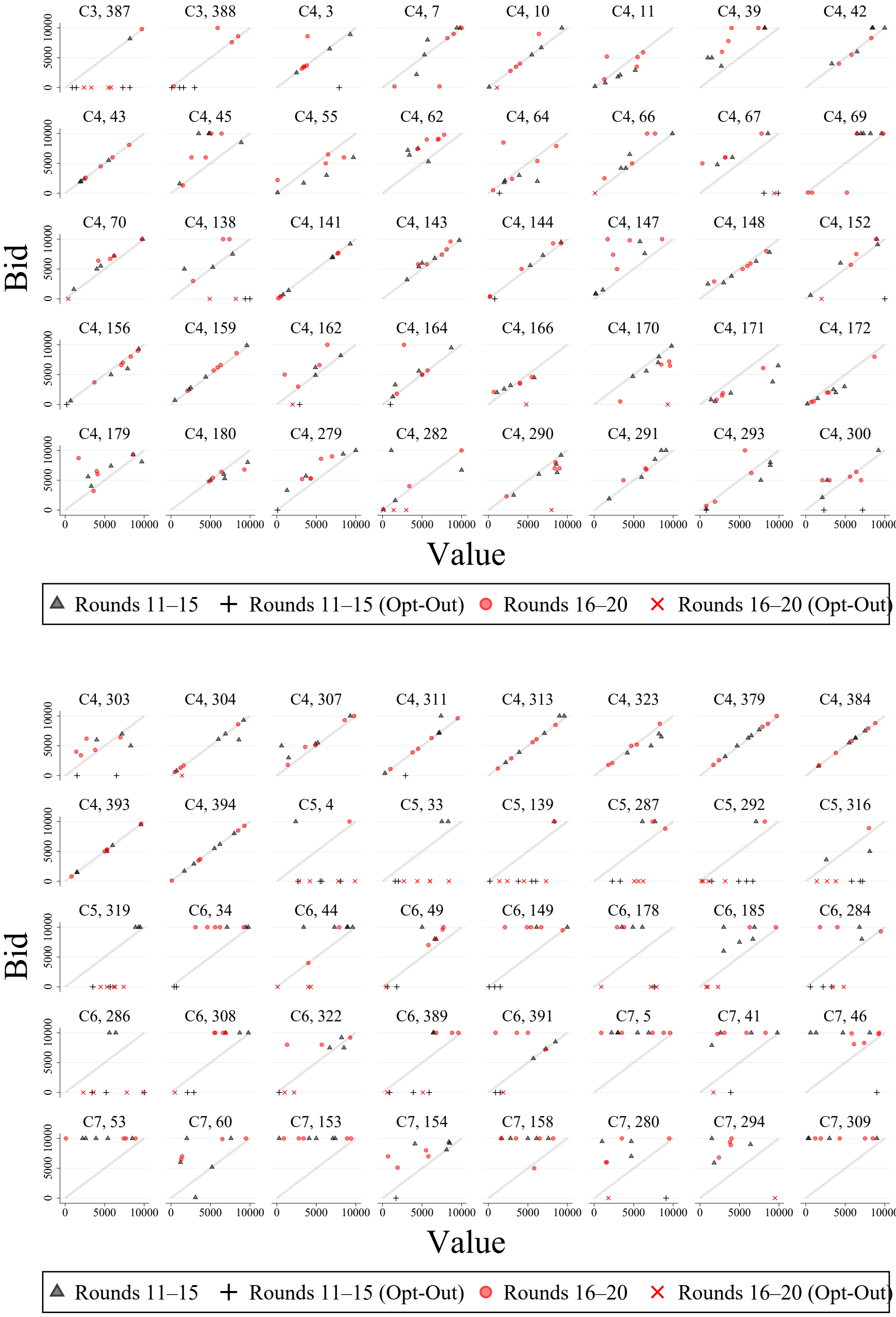


Figure 9: Bids by Subjects in FE. *Note*: Each cluster ID is followed by a participant ID number.

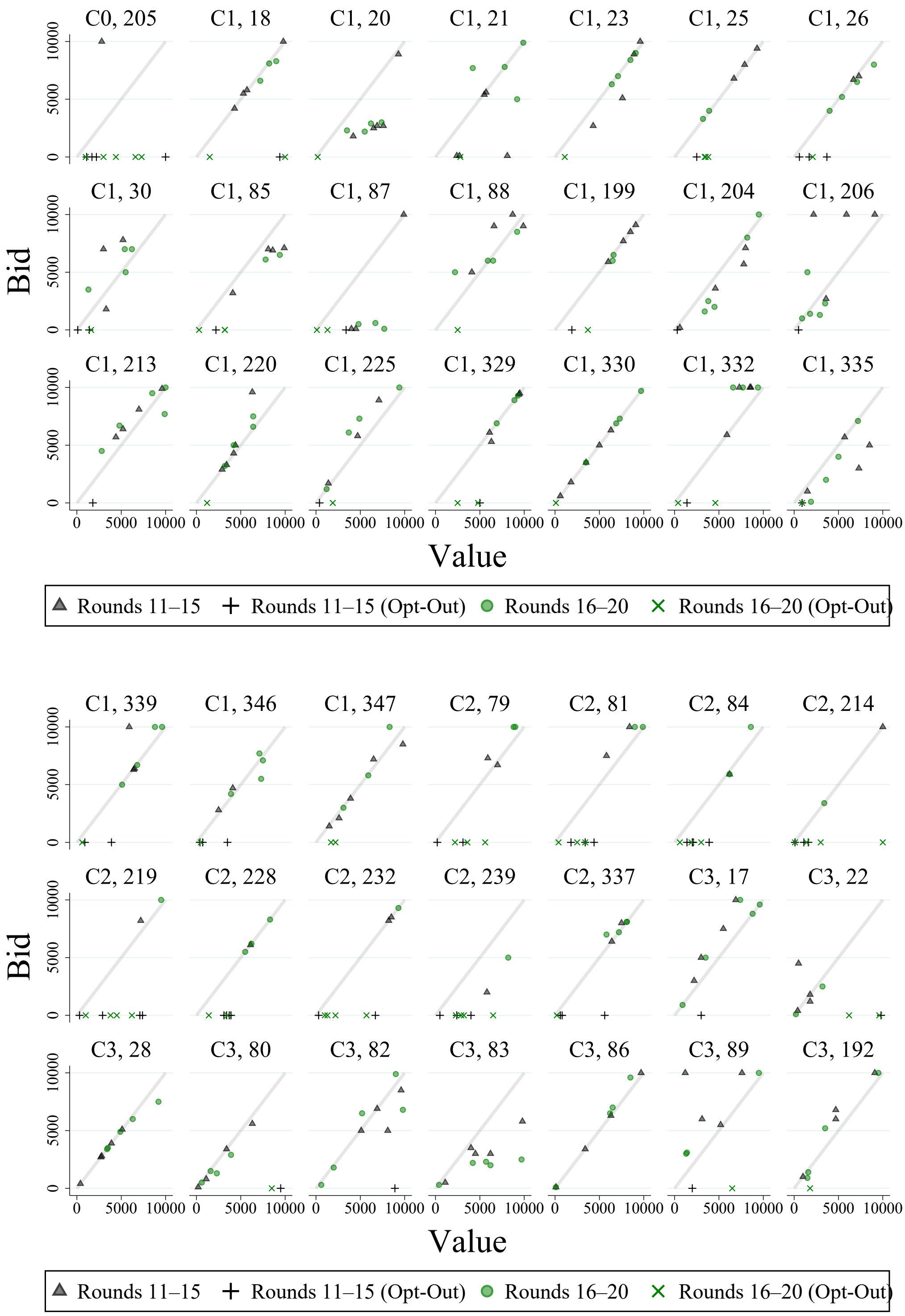


Figure 10: Bids by Subjects in PE. *Note*: Each cluster ID is followed by a participant ID number.

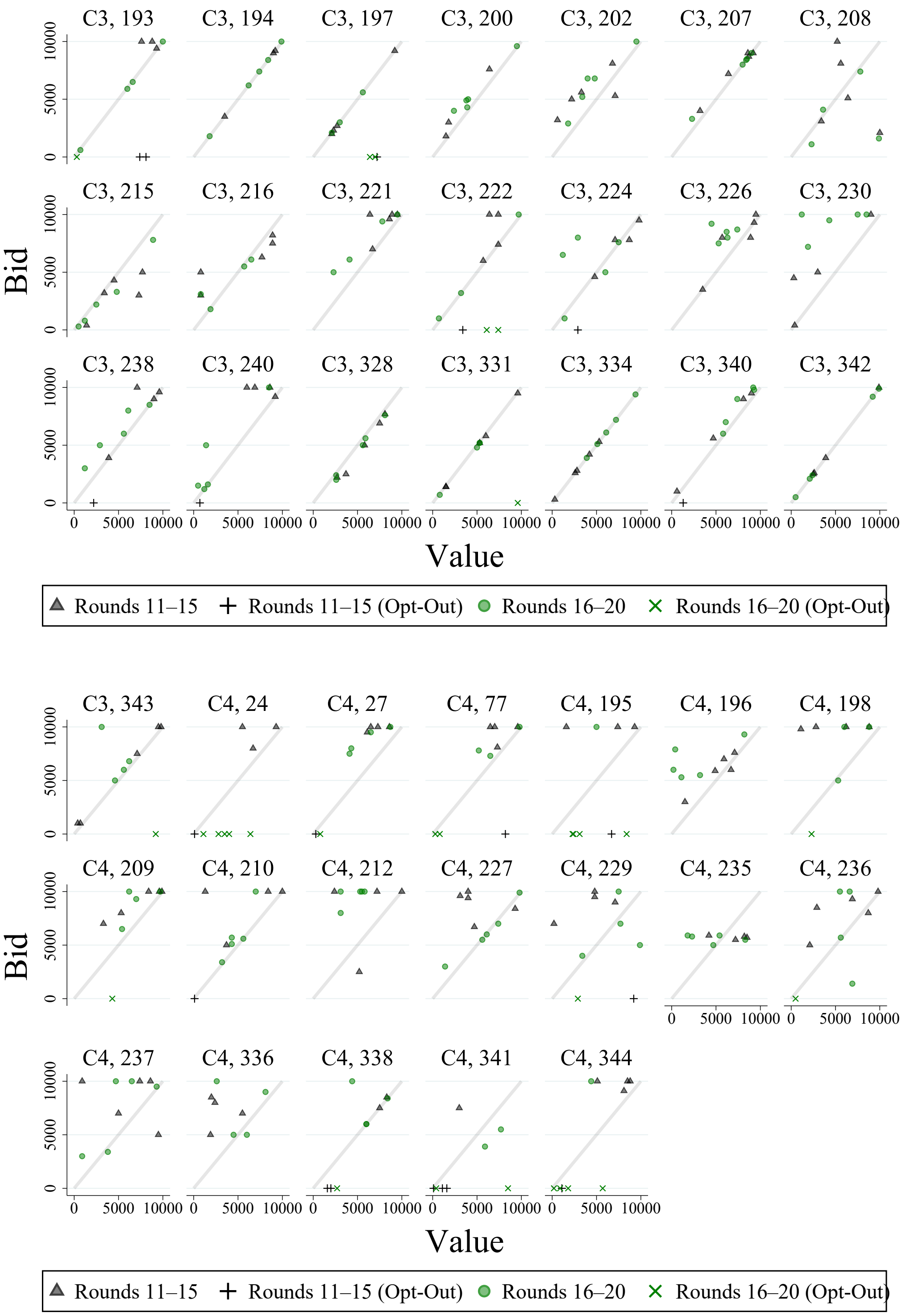


Figure 11: Bids by Subjects in PE. *Note*: Each cluster ID is followed by a participant ID number.

C Pre-Registered Analyses

In this appendix, we conduct statistical analyses according to the pre-registered plan, and explain why and how the original plan is revised in the paper.[1]

In the pre-registration, we specified three outcome variables as the key dependent variables. The first one is an overbid amount for each subject in each round (conditional on entry). The second one is a bid deviation from value bidding for each subject in each round (conditional on entry). The third one is the entry dummy for each subject in each round.

Table 6: Treatment effects on bidding decisions.

Dependent Variable	Overbid (in Entry Round)	Deviation (in Entry Round)	Entry Dummy × 100
PE	194.39 (205.81)	114.16 (169.03)	6.36 (2.68)
2P	−618.68 (203.28)	−858.16 (153.21)	22.26 (1.73)
Constant	336.76 (140.49)	1568.75 (111.54)	74.67 (1.55)
Var. of Group Random Effect	99477	56469	$2.64 / 10^{11}$
Var. of Subject Random Effect	1540254	979837	278
Var. of Residual	3812768	2563453	1123
Log Pseudolikelihood	−45490	−44485	−30518
# of Observations	5021	5021	6132

Note: Overbid $= b_{it} - v_{it}$ with $b_{it} > 0$. Deviation $= |b_{it} - v_{it}|$ with $b_{it} > 0$. Entry Dummy $= I(b_{it} > 0)$.

[1] The experiment was pre-registered at AsPredicted.org (#119225, https://aspredicted.org/55gs-7f5r.pdf).

Table 7: Treatment effects on bidding decisions (with control variables).

Dependent Variable	Overbid (in Entry Round)	Deviation (in Entry Round)	Entry Dummy $\times$ 100
PE	127.76 (203.26)	92.06 (173.56)	7.79 (2.71)
2P	−738.19 (203.72)	−848.59 (153.75)	23.15 (1.80)
Value / 10^2			0.40 (0.03)
Round	26.21 (8.50)	−16.56 (6.96)	−0.60 (0.09)
Last Opponent's Bid / 10^4	507.13 (109.30)	153.79 (91.17)	−5.15 (1.26)
Opponent's Payment Round	−603.12 (241.07)	1052.77 (188.73)	3.78 (2.14)
# of Safe Choices (Gain)			−1.70 (0.62)
# of Safe Choices (Loss)			−0.54 (0.86)
# of Mistakes in Quiz	−48.51 (23.42)	28.27 (18.30)	0.10 (0.31)
CRT Score	−48.99 (47.99)	9.59 (33.58)	−0.77 (0.61)
Male	−164.15 (152.74)	−3.49 (128.97)	−3.25 (2.04)
Constant	438.03 (270.43)	1523.68 (224.83)	80.98 (7.97)
Var. of Group Random Effect	96952	55285	4.32 / 10^{10}
Var. of Subject Random Effect	1548973	1048596	262
Var. of Residual	4160288	2737016	965
Log Pseudolikelihood	−45689	−44641	−27666
# of Observations	5020	5020	5643

Note: Overbid = $b_{it} - v_{it}$ with $b_{it} > 0$. Deviation = $|b_{it} - v_{it}|$ with $b_{it} > 0$. Entry Dummy = $I(b_{it} > 0)$. Value = v_{it}. Round = $t > 1$. Last Opponent's Bid = b_{jt-1}, where subject j is i's opponent in round $t-1$. Opponent's Payment Round = 1 if t is a payment round for the current opponent, and 0 otherwise. Male = 1 if subject i is male, and = 0 otherwise.

Following the pre-registration, we demonstrate the estimated average treatment effects on these outcome variables in Table C.11, using the multilevel mixed-effects linear regression models with group- and subject-level random effects.[2] Here, we exclude data in opponents' payment rounds. In FE, the overbid amount given entry is 336.76 points (p = 0.0165). The overbid amount and deviation given entry are larger in FE than in 2P (p = 0.0023 for the overbid amount; p < 0.0001 for the deviation). The overbid amount and deviation given entry are smaller in FE than in PE (p = 0.3449 for the overbid amount; p = 0.4994 for the deviation). The entry rate is lower in FE than in PE and 2P (p = 0.0176 for FE vs. PE; p < 0.0001 for FE vs. 2P). The three linear models in Table C.11 have several control variables specified in the pre-registration. Here, we include data

[2]When reporting any effect in this appendix, we present a p-value from the z-test of the null hypothesis that the associated coefficient (or constant) is zero. The z-statistic is defined as the estimate divided by the robust standard error clustered at the group level.

in opponents' payment rounds to examine the effects on bidding decisions. These models demonstrate similar treatment effects as follows: The overbid amount and deviation given entry are larger in FE than in 2P ($p = 0.0002$ for the overbid amount; $p < 0.0001$ for the deviation). The overbid amount and deviation given entry are smaller in FE than in PE ($p = 0.5296$ for the overbid amount; $p = 0.5958$ for the deviation). The entry rate is lower in FE than in PE and 2P ($p = 0.0040$ for FE vs. PE; $p < 0.0001$ for FE vs. 2P).

In Section 4.1 of the paper, we have used the seller's realized revenues, instead of subjects' decision variables, to test equilibrium hypotheses. Moreover, we include the minimal-entry equilibrium as another hypothesis for each of FE and PE (Figure 3 of the paper). The reason for this inclusion is that, in FE, this equilibrium is more plausible than the full-entry equilibrium from the perspective of evolutionary game theory, as shown by Theorem 2 and Claim 1 of the paper.

In Section 4.2 of the paper, we have changed the models in Table C.11 into those in Tables 3–5 of the paper. The first change is that we use models by treatment. This is because the effects of some control variables are essentially different between the treatments (e.g., the round effects on the deviation in Table IV). Second, we replace the overbid amount with the bid, and adopt tobit models in which the bids in entry rounds, censored at the upper limit of 10000, are regressed on values with other control variables (Table III). As shown in the paper (e.g., Figures 5 and 6a), both FE and PE had overbidders who bid the maximum amount 10000. This suggests that the upper limit might be constraints for them. Hence, we use the tobit models to obtain less biased estimates. Third, we introduce the squared term of values for entry decisions (Table V) since the entry probabilities in partial-entry equilibria for risk-averse bidders are nonmonotonic in values (Figures 1 and 2 of the paper).

D English Translation of the Instruction for FE

We now explain the rules of the experiment. Please watch the video until the end. This is an "auction" experiment. Participants will be asked to play an "auction game" with others under the experimental rules. The experiment will be conducted on a website. The items up for auction are hypothetical. There will be two bidders, competing as a pair. Monetary rewards will be provided according to the experimental results. The initial reward is 3,000 yen, and the reward amount will increase or decrease depending on the results. The profits and losses that occur in the game are referred to as "points."

Now, we explain the rules for each round of the auction. First, the "value" of each item for each bidder will be determined stochastically. A bidder who wins an auction will earn points based on their value of the item. The value ranges from 100 to 10,000 points, in increments of 100 points. The probability of each combination of values for the two bidders is shown on the next page.

In this table,[3] the numbers on the left represent Bidder A's values, while the numbers on the top represent Bidder B's values. The values range from 100 to 10,000 points, in increments of 100 points. Some parts of the table are omitted. The probability of each cell along the diagonal is 51 out of 15,000, while the probability of any other cell is 1 out of 15,000. This means that the likelihood of both bidders having the same value is relatively high. For example, the probability that both bidders have a value of 100 is 51 out of 15,000. Meanwhile, the probability that Bidder A has a value of 100 and Bidder B has a value of 10,000 is 1 out of 15,000. The sum of the fractions in the white cells adds up to 1, or 100%.

Each participant is only informed of their own value and not the opponent's value. Each participant will then decide whether to opt in the auction and, if they choose to opt in, the amount they wish to bid. Bids can range from 100 to 10,000 points, in increments of 100 points. If a participant chooses to opt out, their bid will be 0 points. In that case, their profit will also always be 0 points.

After both participants have submitted their bids, the price will automatically rise from 0 points, stopping at the lower of the two bids. This final price is the "winning price." The participant who submitted the higher bid becomes the "winner." In the event of equal bids, the winner is determined randomly. If both participants choose to opt out, there is no winner. The winning bidder earns a "winning profit." The winning profit is determined by subtracting the winning price from the winner's value.

Finally, if both participants opt in the auction, each participant's profit will also be adjusted depending on the opponent's bid, regardless of the participant's own bid or whether the participant won the auction. The adjustment is made according to the formula displayed on the screen:

$$\text{Your profit change} = 2223 - \frac{\text{Opponent's bid}}{100} \times \left(\frac{\text{Opponent's bid}}{100} - 1 \right)$$

Although the formula is complicated, you can check it during the experiment. However, if either you or your opponent opts out of the auction, this profit ad-

[3] The table is shown in Figure B.0.

justment will be 0 points.

Additionally, the profit adjustment is represented by a graph.[4] In this graph, the horizontal axis represents the opponent's bid, while the vertical axis represents the change in your profit. As you can see from the graph, the higher the opponent's bid, the fewer points you earn. You can also check this graph during the experiment.

Now, we explain how groups and competing pairs are determined. In the beginning, each participant will be randomly assigned to a group of six. These groups will remain fixed for the duration of the experiment. If there are not enough participants, computers will join to form groups of six. Participants will be informed if computers are included in their group. If a computer participates, its bidding behavior will be displayed on the screen. Each group will play 20 rounds of auction games. The competing pair for each auction will be randomly selected within the group for each round.

Next, we explain how rewards from the auction are determined. For each participant, one of the 20 rounds will be randomly selected as the round that determines their reward. The reward from the auction is calculated as follows: The total profit from the selected round is multiplied by 0.1 to determine the performance-based reward. However, this performance-based reward can be negative. Fractions will be rounded up. In addition, an initial reward of 3,000 yen will be given. The total auction reward is the sum of the performance-based reward and the initial reward. You will not know which round is selected for your reward until the end of the experiment, so please make decisions seriously in all rounds.

Now, we explain the layout of the screens for each round. Each round of the auction consists of two screens. The first is the bidding screen.[5] The upper left of the page displays the time limit. After the time limit passes, the system automatically moves to the next page. The page title indicates that this is the third round of bidding. Below that, you can confirm whether this round is the reward-determining round for your opponent. However, you will not know if this round is the reward-determining round for you until the end of the experiment. Below that, you can check your value for this round. In this example, your value is 9,000 points. Next, please focus on the blue box. Here, you will choose whether to opt in and, if so, how much to bid. Use the radio buttons to select whether to

[4] The graph is shown in Figure B.0.

[5] The screen is shown in Figure B.0.

opt in. If you choose to opt in, use the slider to select your bid. You can operate the slider with your mouse or keyboard. The amount you select will be displayed below the slider. After the time limit passes, your selection will be submitted automatically. You can change your selection as many times as you like before the time limit.

There is more to the bidding screen. By clicking this menu, you can check the results of previous rounds, the experimental rules, and the profit adjustment rules. Now, let's explain how to use a calculator for the profit adjustment in detail.[6] First, enter Bidder B's bid amount. Then, click the "calculation" button. The change in Bidder A's profit will then be displayed below. Please use this calculator if you wish to check the profit adjustment rules.

The second screen is the auction result screen.[7] The price will rise from 0 points in an animation, stopping at the lower of the two bids. This final price is the winning price. In this example, your bid is 5,000 points, and your opponent's bid is 4,000 points, so the winning price is 4,000 points. Since you submitted the higher bid, you are the winner. The winning profit is calculated by subtracting the winning price (4,000 points) from your value (9,000 points), resulting in a profit of 5,000 points. Additionally, according to the profit adjustment rules explained earlier, your profit increases by 663 points according to your opponent's bid. As a result, your total profit in this round is the sum of your winning profit and the profit change, which is 5,663 points. This concludes one round of the auction.

Finally, we explain the schedule of the experiment. First, you will log in to the experiment website. After logging in, you will be asked to answer a brief questionnaire for trait assessment. The questionnaire will be shown on the screen later. The results of the questionnaire will affect your rewards. Afterward, you will participate in practice auctions. These practice rounds are intended to help you get familiar with the rules and operations, and two rounds of practice auctions will be conducted. The practice rounds will not count toward the rewards. After the practice auctions, you will take a quiz to confirm your understanding of the rules. Then, the actual auction will begin. There will be 20 rounds of auctions in total, and they will count toward your rewards. After the main rounds, you will take another quiz. The quiz content will be shown on the screen later. This quiz is unrelated to the auction. There is an additional reward for completing this quiz. Finally, after a brief survey and receipt of your participation payment,

[6] The calculator is shown in Figure B.0.
[7] The screen is shown in Figure B.0.

the experiment will conclude. The payment will be made in increments of 10 yen. Any amount less than 10 yen will be rounded up.

This concludes the explanation of the experimental rules.

E Reinforcement Learning Model

In this appendix, we apply the reinforcement learning model of Börgers and Sarin (1997) to our auction game. Since the auction game is a Bayesian game, we need some modification. We also borrow the idea of perturbation from Gale et al. (1995, Section 6) .

Fix any step size $\kappa \in (0, 1]$. Let $\tilde{\sigma}(t)$ and $\tilde{\varsigma}(t)$ denote bidder i's and j's *states* at each time $t \in \{0, \kappa, 2\kappa, ...\}$ respectively. Their states are random variables. We assume that $\tilde{\sigma}(0) = \sigma(0)$ and $\tilde{\varsigma}(0) = \varsigma(0)$ with probability one, and bidders have the same initial state $\sigma(0) = \varsigma(0) \in \Delta(B)^V$. At each time t, bidders independently choose bids according to their states given their realized values. We denote by $\sigma(t) = E[\tilde{\sigma}(t)]$ and $\varsigma(t) = E[\tilde{\varsigma}(t)]$ their expected states at t. We call the component $\tilde{\sigma}^v(t) \in \Delta(B)$ the substate for bidder i with value v at t.

We assume that, at each time, each bidder observes only his or her own previous bids, values, and payoffs. Bidders update their states as follows: Following Börgers and Sarin (1997), we interpret payoffs as "strengths of reinforcement." Fix any CARA parameter $r \in \mathbb{R}$. We apply a positive affine transformation to the CARA utility function u so that $u(x) \in (0, 1)$ for all profit levels x in the experiment. Let $\tilde{u}_b^v(t)$ denote the payoff for bidder i with value v bidding b at time t. This payoff is realized according to the opponent's state $\tilde{\varsigma}(t)$ and value $\tilde{v}_j$ conditional on $\tilde{v}_i = v$. The interim expected payoff is

$$E\left[\tilde{u}_b^v(t)\right] = U^{\varsigma(t)}(v, b).$$

First, we define a learning process with no perturbation. We suppose that bidder i with value v has chosen a bid b at time t. At time $t + \kappa$, bidder i with value v updates his or her substate by taking a weighted average of the old substate $\tilde{\sigma}^v(t)$ and the unit vector that puts all probability on bid b. The weight on the unit vector is given by $\kappa\tilde{u}_b^v(t)$. Bidder i with another value $w \neq v$ retains the old substate. The learning process is thus defined as

$$\tilde{\sigma}_b^v(t+\kappa) = (1 - \kappa\tilde{u}_b^v(t))\tilde{\sigma}_b^v(t) + \kappa\tilde{u}_b^v(t),$$
$$\tilde{\sigma}_a^v(t+\kappa) = (1 - \kappa\tilde{u}_b^v(t))\tilde{\sigma}_a^v(t)$$

for each $a \in B \setminus \{b\}$, and $\tilde{\sigma}^w(t+\kappa) = \tilde{\sigma}^w(t)$ for each $w \in V \setminus \{v\}$. Taking expectation with respect to bidder i's value, bid, payoff, and substate at time t yields

$$\frac{\sigma_b^v(t+\kappa) - \sigma_b^v(t)}{\kappa p(v)} = \left(U^{\varsigma(t)}(v,b) - U^{\varsigma(t)}(v,\sigma(t))\right)\sigma_b^v(t).$$

Next, we introduce perturbations into the learning process. At each time $t+\kappa$, bidder i with value v follows the original process with probability $1-\delta(t)$, and updates his or her substate by taking a weighted average of $\tilde{\sigma}^v(t)$ and θ^v with the remaining probability $\delta(t)$. The weight on the exogenous substate θ^v is κ. Bidder i with another value $w \neq v$ retains the old substate. The process is thus modified as follows:

$$\tilde{\sigma}_b^v(t+\kappa) = (1-\delta(t))\left((1-\kappa\tilde{u}_b^v(t))\tilde{\sigma}_b^v(t) + \kappa\tilde{u}_b^v(t)\right) + \delta(t)\left((1-\kappa)\tilde{\sigma}_b^v(t) + \kappa\theta_b^v\right),$$
$$\tilde{\sigma}_a^v(t+\kappa) = (1-\delta(t))(1-\kappa\tilde{u}_b^v(t))\tilde{\sigma}_a^v(t) + \delta(t)\left((1-\kappa)\tilde{\sigma}_a^v(t) + \kappa\theta_a^v\right)$$

for each $a \in B \setminus \{b\}$, and $\tilde{\sigma}^w(t+\kappa) = \tilde{\sigma}^w(t)$ for each $w \in V \setminus \{v\}$. Taking expectation with respect to bidder i's value, bid, payoff, and substate at time t yields

$$\frac{\sigma_b^v(t+\kappa) - \sigma_b^v(t)}{\kappa p(v)} = (1-\delta(t))\left(U^{\varsigma(t)}(v,b) - U^{\varsigma(t)}(v,\sigma(t))\right)\sigma_b^v(t) + \delta(t)\left(\theta_b^v - \sigma_b^v(t)\right).$$

Since $\sigma(t) = \varsigma(t)$ due to $\sigma(0) = \varsigma(0)$ and symmetry between bidders, we obtain equation (8) in the paper by replacing the step size κ with τ.[8]

F Proofs of Lemmas in the Paper

Lemma 1. *A function $l_0 : V \to \mathbb{R}$ satisfies equation (2) for each $v \in V$ if and only if l_0 satisfies equation (1) for each $b \in V$.*

Proof of Lemma 1. It follows from the compound distribution of $(\tilde{v}_1, \tilde{v}_2)$ that the break-even condition (2) is equivalent to

$$\rho l_0(v) + (1-\rho)E\left[v - \min\{v,\tilde{w}\} + l_0(\tilde{w})\right] = 0. \tag{16}$$

Simple algebra shows that the function l_0 defined by (1) satisfies equation (16)

[8] As explained in footnote 28 of the paper, we set $\tau = 10$ to speed up computation. This is equivalent to setting $\kappa = 1$ and multiplying the right side of the above equation by 10.

for each $v \in V$. Conversely, suppose that a function l_0 satisfies equation (16) for each $v \in V$. By multiplying both sides of (16) by $p(v)$ and summing over v, we obtain $\sum_v p(v)l_0(v) = -(1-\rho)\sum_v \sum_w p(v)p(w)\max\{v-w,0\}$, and hence, $E[l_0(\tilde{w})] = -(1-\rho)E[\max\{\tilde{w}_i - \tilde{w}_j, 0\}]$. Substituting this expected value back into equation (16) gives equation (1). □

Lemma 2. *Fix any opt-out strategy z for risk-neutral bidders. (i) If $z(v')>0$ for some $v'>l(v')$, then $U^z(v')>U^z(v)$ for each $v \leq l(v)$. (ii) If $z(v)>0$ for some $v<l(v)$, then $U^z(v)<U^z(v')$ for each $v' \geq l(v')$.*

Proof of Lemma 2. (i) Fix any $v \leq l(v)$. We obtain the inequality $U^z(v')>U^z(v)$ from equation (4) with $(v'-l(v'))z(v')>0 \geq (v-l(v))z(v)$ and $v'>v$. (ii) Fix any $v' \geq l(v')$. We obtain the inequality $U^z(v)<U^z(v')$ from equation (4) with $(v-l(v))z(v)<0 \leq (v'-l(v'))z(v')$ and $v<v'$. □

Lemma 3. *Fix any symmetric equilibrium z for risk-neutral bidders. Then: (i) z is decreasing on $\{v \in V \mid v \geq l(v)\}$. (ii) $z(v^n)<1$. (iii) If $z(v^m)=0$ for some $v^m \geq l(v^m)$, then $z(v)=0$ and $U^z(v)=U^z(v^m) \geq 0$ for each $v>l(v)$. (iv) $U^z(v) \leq \pi$ for each $v \in V$.*

Proof of Lemma 3. (i) To derive a contradiction, suppose that $z(v')>z(v)$ for some $v,v' \in V$ with $v'-l(v')>v-l(v) \geq 0$. We then obtain $U^z(v')>U^z(v) \geq 0$, where the first inequality follows from equation (4) with $v'>v$ and the second from $z(v)<z(v') \leq 1$. Thus, a bidder with value v' strictly prefers value bidding to opt-out. However, $z(v')>z(v) \geq 0$, which contradicts the equilibrium condition $z(v')=0$.

(ii) To derive a contradiction, suppose $z(v^n)=1$. Then, $z(v)=1$ for each $v \in V$ from part (i) and Lemma 2 (i), that is, both bidders always opt out in this equilibrium. A bidder can then obtain the item for free by bidding a positive bid. This is a contradiction.

(iii) Suppose $z(v^m)=0$ for some $v^m \geq l(v^m)$. Since z is an equilibrium, $U^z(v^m) \geq 0$. Part (i) also implies that $z(v)=0$ for each $v \geq v^m$. Now, if $(v^{m-1}-l(v^{m-1}))z(v^{m-1})>0$, then we would obtain $U^z(v^{m-1})>U^z(v^m) \geq 0$ from equation (4), which contradicts $z(v^{m-1})>0$. It must thus hold that $z(v^{m-1})=0$ if $v^{m-1}>l(v^{m-1})$. We can proceed by induction on v to show that $z(v)=0$ if $v>l(v)$. Then, equation (4) again implies that $U^z(v)=U^z(v^m) \geq 0$ for each $v>l(v)$.

(iv) To derive a contradiction, suppose that $U^z(v')>\pi \geq 0$ for some $v' \in V$. Then, $z(v')=0$ since z is an equilibrium. This implies $U^z(v^n) \geq U^z(v')>0$ and

$z(v^n) = 0$. However, it then follows from part (iii) that $z(v) = 0$ for each $v > l(v)$. Thus,

$$\pi < U^z(v') = (1-\rho) \sum_{w \le l(w)} p(w)(w - l(w))z(w) + \pi \le \pi,$$

which is a contradiction. $\square$

Lemma 4. *Every ASS z with $z(v) = 0$ for each $v \ge l(v)$ satisfies $z(v) \in \{0, 1\}$ for each $v < l(v)$.*

Proof of Lemma 4. To derive a contradiction, suppose that an ASS z satisfies $z(v) = 0$ for each $v \ge l(v)$ while $\underline{V} := \{v \in V \mid z(v) \in (0,1), v < l(v)\}$ is nonempty. Since z is an equilibrium, $U^z(v) = 0$ for each $v \in \underline{V}$. The replicator dynamics (5) then implies that, on the subset

$$\begin{aligned}\{\sigma \in \Delta(B)^V \, |\sigma(v,v) + \sigma(v,0) = 1 \text{ and } \sigma(v,0) \in (z(v),1) \text{ if } v \in \underline{V},\\ \sigma(v,v) = 1 - z(v) \text{ and } \sigma(v,0) = z(v) \text{ if } v \notin \underline{V}\}\end{aligned}$$

of states, the proportion of value-bidders in the subpopulation of any value $v \in \underline{V}$ is strictly decreasing because $\dot{\sigma}(v,v) = p(v)U^\sigma(v,v)(1-\sigma(v,v))\sigma(v,v) < p(v)U^z(v)(1-\sigma(v,v))\sigma(v,v) = 0$, where the inequality follows from equation (3) with $\sigma(v',0) > z(v')$ for each $v' \in \underline{V}$. Thus, the state z is not an ASS. This is a contradiction. $\square$

References

Börgers, Tilman and Rajiv Sarin (1997): "Learning through Reinforcement and Replicator Dynamics," *Journal of Economic Theory*, 77 (1), 1–14.

Gale, John, Kenneth G. Binmore, and Larry Samuelson (1995): "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior*, 8 (1), 56–90.

Harada, Yuki, Etsuko T. Harada, and Satoru Suto (2018): "Examination of Cognitive Reflection Test (CRT): Homogeneity of Questions, Order Effects, and Effects of Instruction," *Tsukuba Psychological Research*, 56, 27–34.