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 102 号　2022 年 5 月
RISS Discussion Paper Series
No.102　May , 2022

An Experimental Study of Crémer–McLean Auctions

Takeshi Nishimura

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

An Experimental Study of Crémer–McLean Auctions*

Takeshi Nishimura†

May 9, 2022

Abstract

Full-surplus-extraction auctions designed by Crémer and McLean (1988) have been criticized as unrealistic mechanisms. We assess the performance of these auctions in a controlled experimental environment. The experiment has two treatments: first-price and second-price auctions supplemented with Crémer–McLean lotteries. In the experiment, subjects cannot opt out of auctions. This rule allows us to focus on their bidding behavior. The experimental evidence clearly shows that neither auction works as predicted. In each treatment, the seller's actual revenue is significantly higher than the equilibrium revenue. Panel data analysis shows that overbidding persists throughout all rounds in the second-price treatment, while it slightly diminishes through rounds in the first-price treatment.

Keywords: Crémer–McLean auctions; Auction experiment

JEL Classification C90 · D44 · D82

*I am extremely grateful to Kazuhito Ogawa and Toshihiko Nanba for their valuable support for running the experiment. I gratefully acknowledge the Experimental Economics Laboratory, RISS, Kansai University for recruiting subjects. I thank the participants at the Contract Theory Workshop on Zoom for their helpful comments and suggestions. This work was supported by JSPS KAKENHI Grant Numbers 18K12747.

†Faculty of Economics, Komazawa University; 1-23-1 Komazawa, Setagaya-ku, Tokyo, 154-8525, Japan; tnishi@komazawa-u.ac.jp

1 Introduction

Full-surplus-extraction auctions designed by Crémer and McLean (1988, henceforth CM) have been criticized as unrealistic mechanisms. CM auctions implement an efficient outcome, without leaving any information rents for risk-neutral bidders, by using a commonly known statistical dependence of their valuations for an auctioned item. These mechanisms violate the "Wilson doctrine" (Wilson, 1987), which suggests that trading rules should not rely excessively on common knowledge. Börgers (2015, p. 124) argued 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] In fact, 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."

However, how do CM auctions work in controlled experimental environments? An experimental approach allows us to address difficulties in testing theoretical hypotheses concerning CM auctions. Most importantly, an experimenter can design a joint probability distribution of bidders' valuations, and this common prior can be common knowledge among subjects. If it turns out that CM auctions do not work as predicted, then we can analyze experimental data to investigate the underlying causes of this failure. If it is confirmed that CM auctions perform well in these clean environments, then we can take the next step to test this performance in more practical situations. The purpose of this experimental study is to provide behavioral evidence to scrutinize whether CM auctions are impractical.

As in the framework of CM, subjects in our between-subject experiment play either a Bayesian auction or a dominant strategy auction in a correlated-private-values environment. Specifically, a pair of subjects plays either a first-price (1P) or a second-price (2P) auction supplemented with "lotteries" (or "side-bets"). These lottery prizes depend on an opponent's bid, and hence, the interim expected value of a bidder's lottery can depend on his own valuation due to correlation. Using this fact, we design lotteries such that truthful bidding is a Bayesian Nash equilibrium in the 1P-CM auction and a dominant strategy in the 2P-CM auction, and the seller extracts the full surplus on average in equilibrium. Indeed, the 1P-CM auction with these lotteries is dominance solvable. Each pair repeatedly plays 40 auction rounds. Subjects have no opt-out option. In other words, our study focuses on subjects' bidding behavior in the continuation game after they participate as in equilibrium.

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

Our experimental evidence clearly shows that neither CM auction works as predicted. Notably, in each treatment, the seller's actual average revenue is significantly *higher* than the equilibrium average revenue. In the 2P-CM treatment, this is true even in later rounds. The actual revenues in the 1P-CM and 2P-CM treatments are on average 24% and 83% above the equilibrium revenues, respectively. Of course, these results imply that subjects' participation constraints are violated, that is, the actual average surplus per bidder is significantly lower than zero. Hence, if subjects had an opt-out option, then they would, at least sometimes, opt out of these unfavorable auctions.

These results are caused by subjects' deviations from equilibrium. Their bidding behavior is far from truthful bidding. Overall, overbidding (i.e., bidding above valuations) is prevalent in *both* treatments. In the 1P-CM treatment, we observe that overbidding slightly diminishes through rounds, possibly due to its dominance solvability. To put it differently, it takes time for subjects to "learn" to bid truthfully. In general, lotteries attached to the 1P auction can be quite complicated. Therefore, how to design appropriate lotteries is a difficult problem for the use of 1P-CM auctions. In the 2P-CM treatment, we find that overbidding persists throughout all rounds. In order to ensure the strategy-proofness (i.e., truthful bidding is a dominant strategy), this treatment adopts, for each bidder, a lottery that depends only on his opponent bid. This lottery seems relatively simple. However, this prize must be *decreasing* in the opponent's bid due to a positive correlation between valuations. Hence, if subjects have spite motives (Morgan et al., 2003), then the 2P-CM auction induces more aggressive overbidding than a 2P auction with no lottery does.

The remainder of the paper is organized as follows. The next subsection discusses related theoretical and experimental literature. Section 2 presents a simplified version of the theoretical framework of CM. Section 3 describes the experimental design. Section 4 analyses the experimental data to test theoretical hypotheses. Section 5 concludes. Appendix A contains supplementary data and reports some regression analysis.

1.1 Related literature

There exists a considerable theoretical literature on full surplus extraction. By extending an insightful example of Myerson (1981), Crémer and McLean (1985, 1988) designed full-surplus-extraction mechanisms in interdependent-values and private-values environments with risk-neutral agents, respectively. Indeed, CM characterized, for each concept of Bayesian-Nash and dominant-strategy equilibria, the common prior distributions that guarantee full surplus extraction given the equilibrium concept. Their Bayesian implementation result was extended to a general environment with continuum type spaces by McAfee and Reny (1992), who indicated that paradoxical full-extraction results "cast doubt on the value of the current mechanism design paradigm as a model of institutional design."

In the framework of CM, under the common-knowledge assumption that each bidder has a fixed finite number of types, priors that guarantee full surplus extraction are generic (in the sense that the set of these priors has full Lebesgue measure and is open). With different concepts of genericity, Heifetz and Neeman (2006) and Chen and Xiong (2013a) proved nongenericity and genericity results 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 the mechanism design paradigm (McAfee and Reny, 1992). Our experiment is not designed so as to empirically study genericity problems.[2] Nevertheless, our experimental evidence suggests that the optimality of CM auctions may be impaired by bidders' bounded rationality or behavioral motives.

Apart from nongenericity results, there are several theoretical explanations why full-surplus-extraction auctions are not observed in practice. These include risk aversion or limited liability of bidders (Robert, 1991), collusive agreements between bidders (Laffont and Martimort, 2000), information acquisition by bidders about others' types (Bikhchandani, 2010), and non-robustness of full-extraction auctions to bidders' beliefs (Chung and Ely, 2007; Pham and Yamashita, 2021). We prevent these potential confounds as much as possible by applying experimental methods.

Following the literature on optimal auction design, CM considered partial implementation. Full implementation of the full-surplus-extraction outcome was 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.

There exists a large body of experimental literature on auctions, as surveyed by Kagel and Levin (2015). In 1P-auction experiments, subjects tend to bid higher than risk-neutral equilibrium bids, but lower than their valuations. This bidding

[2] As suggested by CM, it may be impossible to test their generic implementation results (Theorems 1 and 2 of CM) even experimentally. However, we can design an experiment to examine how the degree or sign of value correlation changes subjects' behavior in CM auctions. See Section 5 for this possibility.

behavior has been ascribed to subjects' risk aversion (Cox et al., 1982; Füllbrunn et al., 2019; Rietz, 1993; Walker et al., 1990). Other explanations include biased probabilistic beliefs (Armantier and Treich, 2009) and regrets (Engelbrecht-Wiggans and Katok, 2008; Filiz-Ozbay and Ozbay, 2007). In our 1P-CM treatment, subjects' bids are significantly higher than their *valuations*. This result is in stark contrast with results in 1P-auction experiments.

In 2P-auction experiments, subjects tend to bid higher than their valuations. This overbidding behavior has been observed in environments with affiliated private values (Breitmoser and Schweighofer-Kodritsch, 2022a; Garratt et al., 2012; Harstad, 2000; Kagel et al., 1987; Li, 2017) and with independent private values (Andreoni et al., 2007; Bartling and Netzer, 2016; Cooper and Fang, 2008; Kagel and Levin, 1993; Schneider and Porter, 2020; Tan, 2020). For this overbidding behavior, several explanations have been proposed: spite (Morgan et al., 2003), joy of winning (Cooper and Fang, 2008), and cognitive limitation (Li, 2017). These three factors can also cause overbidding in 2P-CM auctions. In particular, as already mentioned, subjects' spite motives might cause more aggressive overbidding in the 2P-CM treatment.

To the best of my knowledge, there has been no previous experimental study that tests the performance of CM auctions. However, a related mechanism-design problem was experimentally studied by Krajbich et al. (2009) and Krajbich et al. (2017). For public-goods problems, they tested the performance of "neurometrically informed mechanisms." In this novel mechanism, each agent's tax payment depends on an ex-post signal correlated with his 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.[3] If this signal technology satisfies a generic CM condition, then we can solve the free-rider problem by using CM lotteries. Indeed, their neurometrically informed mechanisms, which are efficient, strategy-proof, and interim individually rational, performed surprisingly well in experiments. In contrast with their mechanisms, a subject's lottery prize in CM auctions depends on his/her opponent's bid, not on his/her own neural signal. This additional strategic interaction between subjects, together with their bounded rationality and behavioral motives, might cause non-truthful bidding behavior in our experiment.

[3] In their experiments, 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

We consider a special case of the general model of CM. In particular, we employ two auction mechanisms designed by them, focussing on an environment with two ex-ante symmetric bidders.

2.1 Auction environment

An indivisible item is auctioned off by a seller who attaches no value to it. There are two risk-neutral bidders ($i \in \{1, 2\}$) whose reservation payoffs are zero. Each bidder i's valuation (i.e., willingness to pay) for the item is given by a random variable v_i. The set of possible valuations $V \subset \mathbb{R}_+$ is finite, and common to the two bidders. We denote by π a common prior distribution on the set V^2 of valuations. The joint distribution π is symmetric. Each bidder privately knows his realized valuation, but the prior π is common knowledge among the three parties.

For each valuation v_i of bidder i, we denote by $\pi(v_j \mid v_i) \equiv \pi(v_i, v_j)/\sum_{v'_j} \pi(v_i, v'_j)$ the interim belief about the other's valuation v_j, $j \neq i$. Let Γ denote the square matrix of size $|V|$ whose row vectors are $\{\pi(\cdot \mid v_i) \mid v_i \in V\}$. As a statistical dependence, we assume that the matrix Γ has full rank. Note that if the bidders' valuations were statistically independent, the matrix Γ would have rank 1.

2.2 Second-price Crémer–McLean auction

First, let us consider a standard 2P auction. A pair of sealed bids is denoted by $b = (b_1, b_2) \in V^2$. Each bidder i wins the auction if $b_i > b_j$, and loses if $b_i < b_j$. If there is a tie (i.e., $b_i = b_j$), a winner is randomly selected with equal probability. A winner i gets the item and pays the other's bid b_j to the seller. A loser obtains nothing. In this auction, truthful bidding (i.e., $b_i = v_i$) is a (weakly) dominant strategy for each bidder i. Hence, the interim expected surplus for each bidder with valuation v_i is given by

$$h(v_i) \equiv \sum_{v_j < v_i} \pi(v_j \mid v_i)\,(v_i - v_j)\,. \tag{1}$$

We attach a symmetric lottery $l : V \to \mathbb{R}$ for each bidder to the 2P auction. We call this augmented auction a *2P-CM auction*. In this auction, after a winner is decided as above, each bidder i gets a lottery prize $l(b_j)$ from the seller. This (positive or negative) prize amount depends only on the other's bid b_j. We define this lottery l as

$$l \equiv -\Gamma^{-1} h, \tag{2}$$

where the surplus function h is identified with the $|V|$-dimensional column vector. Note that the square matrix Γ is invertible because it has full rank. In this auction, each bidder's ex-post payoff is given as follows:

$$u_i = \begin{cases} v_i - b_j + l(b_j) & \text{if bidder } i \text{ wins,} \\ l(b_j) & \text{if bidder } i \text{ loses.} \end{cases} \tag{3}$$

We observe that the 2P-CM auction is also strategy-proof, and hence, a bidder with the highest valuation wins the item. Further, the definition (2) of the lottery l immediately implies that, for each valuation v_i,

$$h(v_i) + \sum_{v_j} \pi(v_j \mid v_i) l(v_j) = 0. \tag{4}$$

In other words, the interim expected surplus for each bidder is always zero (i.e., no bidder obtains any information rent), provided that both bidders adopt truthful bidding. Thus, the 2P-CM auction (partially) implements the outcome of full surplus extraction.[4]

2.3 First-price Crémer–McLean auction

Next, let us consider a standard 1P auction. A winner is decided in the same way as the 2P auction, but the payment rule is different. That is, a winner pays his own bid to the seller. The 1P auction is not strategy-proof, because each bidder has an incentive for bid shading to reduce his payment.

We attach a symmetric lottery function $l : V^2 \to \mathbb{R}$ for each bidder to the 1P auction. We call this augmented auction a *1P-CM auction*. In this auction, after a winner is decided, each bidder i gets a lottery prize $l(b_i, b_j)$ from the seller. This (positive or negative) prize amount depends on both bids. We define this function l as follows: Since the matrix Γ has full rank, no vector $\pi(\cdot \mid v_i)$ can be a convex combination of the other $|V| - 1$ vectors $\{\pi(\cdot \mid v_i') \mid v_i' \neq v_i\}$. Then, using the

[4] This auction game also has Bayesian Nash equilibria with dominated strategies (e.g., one bidder always bids the amount $\max V$, and the other always bids the amount $\min V$).

separating hyperplane theorem, we can find a function $g : V^2 \to \mathbb{R}$ such that

$$\sum_{v_j} \pi(v_j \mid v_i) g(v_i, v_j) = 0, \tag{5}$$

$$\sum_{v_j} \pi(v_j \mid v_i') g(v_i, v_j) < 0 \tag{6}$$

for each $v_i, v_i' \in V$ with $v_i \neq v_i'$. Given this function with a positive number $\gamma > 0$, the lottery function is defined by $l \equiv \gamma g$. That is, l is a proper scoring rule such that each bidder breaks even (Krajbich et al., 2017). In this auction, each bidder's ex-post payoff is given as follows:

$$u_i = \begin{cases} v_i - b_i + l(b_i, b_j) & \text{if bidder } i \text{ wins,} \\ l(b_i, b_j) & \text{if bidder } i \text{ loses.} \end{cases} \tag{7}$$

Using inequalities (6), we can choose γ so large that the expected value of the lottery $l(v_i, \cdot)$ for bidder i with valuation $v_i' \neq v_i$ is sufficiently negative, and hence, truthful bidding constitutes a Bayesian Nash equilibrium. Further, equation (5) implies that, for each valuation v_i,

$$\sum_{v_j < v_i} \pi(v_j \mid v_i)(v_i - v_i) + \sum_{v_j} \pi(v_j \mid v_i) l(v_i, v_j) = \gamma \sum_{v_j} \pi(v_j \mid v_i) g(v_i, v_j) = 0. \tag{8}$$

In other words, the interim expected surplus for each bidder is always zero, provided that both bidders adopt truthful bidding. Thus, the 1P-CM auction (partially) implements the outcome of full surplus extraction.

2.4 Specification

In the experiment, parameter values are specified as follows: The set of possible valuations is given by $V = \{0, 10, 20, 30, 40\}$. The set of possible bids is the same as the valuation set V. This means that CM auctions are considered as direct mechanisms. The common prior is specified as

$$\pi(v_i, v_j) = \begin{cases} 8/100 & \text{if } v_i = v_j, \\ 3/100 & \text{if } v_i \neq v_j. \end{cases} \tag{9}$$

It is easy to check that the vectors $\{\pi(\cdot \mid v_i) \mid v_i \in V\}$ are linearly independent, that is, the matrix Γ has full rank. Moreover, this specification implies that the bidders' valuations are positively correlated.

In 2P-CM, the symmetric lottery l is uniquely determined by (2) given the common prior π. Due to the positive correlation, the lottery prize is decreasing in the other's bid, as shown in Table 1.

b_j	0	10	20	30	40
$l(b_j)$	$+18$	$+12$	0	-18	-42

Table 1: Lottery prizes $l(b_j)$ in 2P-CM

In contrast to 2P-CM, 1P-CM allows some freedom to design lotteries. Indeed, there are infinitely many lotteries that satisfy equation (5) and inequalities (6). In the experiment, the lottery function l is specified as in Table 2. The reason for this specification is that the 1P-CM auction with these lotteries is dominance solvable. More specifically, in this Bayesian game, truthful bidding is the unique strategy that survives the interim iterative deletion of strictly dominated strategies (Fudenberg and Tirole, 1991). Some tedious calculations show that the order of deletion is given by Table 3. For example, if bidder i's valuation is $v_i = 0$, then both $b_i = 30$ and $b_i = 40$ are deleted at first, and $b_i = 20$ is deleted after some strategies are deleted at first. As a result, truthful bidding is the unique Bayesian Nash equilibrium strategy in this auction game.

$b_i \backslash b_j$	0	10	20	30	40
0	$+12$	$+8$	-3	-13	-24
10	$+18$	$+9$	-2	-13	-27
20	$+26$	$+18$	$+3$	-13	-39
30	$+32$	$+27$	$+12$	-12	-39
40	$+41$	$+37$	$+22$	-4	-36

Table 2: Lottery prizes $l(b_i, b_j)$ in 1P-CM

In each CM auction, if both bidders always bid truthfully as in equilibrium, then the seller's ex-ante revenue is equal to the expected full surplus. Simple computations show that the expected value of the full surplus is 26.

v_i \ b_i	0	10	20	30	40
0		3rd	2nd	1st	1st
10	5th		3rd	1st	1st
20	3rd	6th		2nd	1st
30	3rd	6th	7th		4th
40	3rd	6th	7th	8th	

Table 3: Order of deletion of strictly dominated strategies in 1P-CM

3 Experimental design

3.1 Procedure

This is a between-subjects experiment with two treatments, *1P-CM* and *2P-CM*. Each subject participates in only one treatment, playing the role as a bidder.

We adopt the partner matching (Tan, 2020) as a matching rule. In each treatment session, two subjects are randomly matched at first, and each pair plays 40 rounds of the auction in this treatment. Under this matching rule, each pair can be regarded as the unit of independent observation in a statistical analysis.

The experiment is programmed in oTree (Chen et al., 2016). In each round, each subject sees his/her realized valuation in 10 seconds, and then, submits his/her bid in 40 seconds.[5] At the end of this round, each subject sees the auction result: who wins, the two subjects' valuations and bids, his/her auction payoff and lottery prize amount,[6] and his/her total payoff in this round. Subjects can also recheck results in previous rounds.

Since the 2P-CM auction is strategy-proof, truthful bidding is a dominant strategy even for non-risk-neutral bidders. However, the equilibrium prediction in 1P-CM may depend on bidders' risk preferences. To induce risk-neutral preferences, the following binary lottery procedure (Roth and Malouf, 1979) is adopted as a payment scheme.[7] The unit of payoff in the experiment is called point. After a subject finishes all rounds, one round is randomly chosen as a payment round. Also, a number is randomly drawn from the set of integers from -100 to 100. If

[5] On a bidding page, subjects can push a submit button only after 20 seconds pass.

[6] To avoid framing effects, this prize was called "an increase or decrease in payoff which is irrelevant to whether you win the auction."

[7] There is mixed evidence on the performance of this procedure in experiments on 1P auctions. For example, Walker et al. (1990) provided negative evidence, while Rietz (1993) provided positive evidence. Basically, our design follows suggestions made by Rietz (1993).

his/her total payoff in the payment round is greater than the random number, then this subject receives a high reward of 4,000 JPY. Otherwise, he/she receives a low reward of 2,000 JPY. Each reward includes a show-up fee of 1,000 JPY.

3.2 Administrative details

The experiment was conducted online in May 2021.[8] In total, 150 subjects were recruited from the subject pool of the Experimental Economics Laboratory, RISS, Kansai University through the ORSEE system (Greiner, 2015). All subjects except one were undergraduate students at Kansai University.[9] Table 4 summarizes the numbers of sessions, subjects, and pairs in each treatment. Table A.1 in Appendix A also shows the numbers of subjects by year of admission. A chi-squared test cannot reject the null hypothesis that years and treatments are independent ($p = 0.331$). This balance-test result justifies the estimation and test of treatment effects in Section 4.

Treatment	# of sessions	# of subjects	# of pairs
1P-CM	6	80	40
2P-CM	6	70	35

Table 4: Sessions, subjects, and pairs in each treatment

Each session proceeded as follows: First, subjects joined an online meeting (Zoom Video Communications, Inc.) to watch an instruction video that explains experimental rules. During the experiment, subjects could not see each other and their anonymity was maintained. Any communication between subjects was prohibited. Second, subjects played 40 auction rounds on an experiment website.[10] Although there was no quiz question about auction rules, subjects could ask questions on an introductory page and check the experimental rules in every round. In particular, bidding pages had a table displaying how a subject's own lottery prize depends on the other's bid. Finally, every subject who finished all rounds identified him/herself by showing his/her student card to the experimenter. The session ended with a questionnaire. Since the experiment was conducted online, rewards were paid into subjects' bank accounts after the session.

Each session lasted approximately 70 minutes on average. The average payment per subject was around 2,800 JPY (25.7 USD at the time of the experiment).

[8] It was not conducted in a laboratory, because of the COVID-19 pandemic.

[9] One graduate student at the university was recruited by mistake.

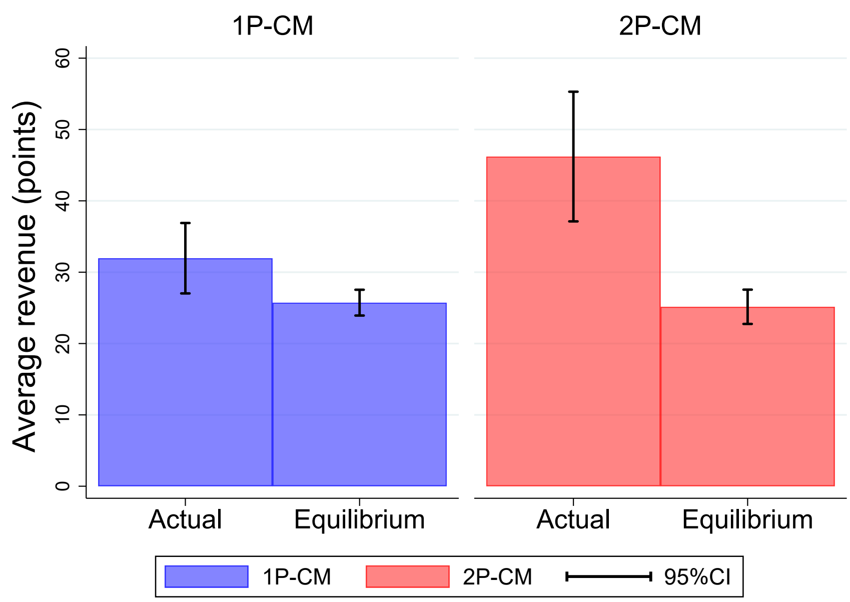
[10] Before these rounds, subjects played 2 dryrun rounds without monetary incentives.

4 Experimental results

In this section, we analyse the experimental data to test whether various measures are consistent with equilibrium predictions. In Subsections 4.1–4.3, we use average data for each pair of subjects as the unit of observation. In Subsection 4.4, we use data for each subject in each round as the unit of observation to conduct a panel data analysis.

4.1 Revenue

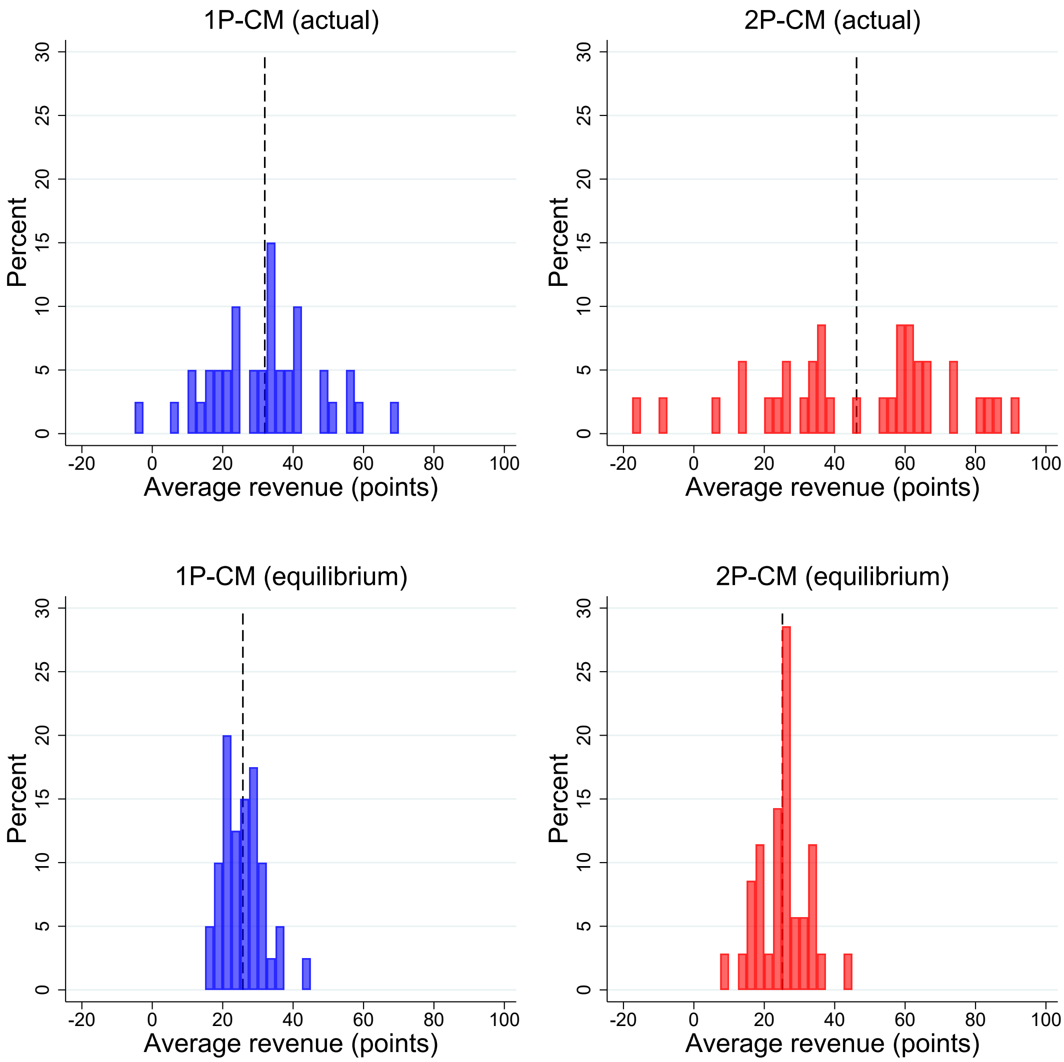
Figure 1 summarizes actual and equilibrium revenues in each treatment. In the experiment, the seller obtains 31.95 and 46.21 points on average in 1P-CM and 2P-CM, respectively. Given the realized valuations in the experiment, the seller's ex-ante revenues in equilibrium are calculated at 25.73 and 25.15 points on average in 1P-CM and 2P-CM, respectively. Note that the equilibrium average revenues from bidder pairs may be different from each other, depending on realized valuations. Figure 2 also displays histograms of average revenues from bidder pairs.



Note: Capped spikes represent 95% confidence intervals (CI).

Figure 1: Actual and equilibrium average revenues in each treatment

We test whether these differences between actual and equilibrium average revenues are statistically significant. For each treatment, a Wilcoxon matched-pairs signed-rank test rejects the null hypothesis that both the actual and the equi-



Note: Dashed lines represent mean values.

Figure 2: Frequencies of average revenues from bidder pairs in each treatment

librium revenues come from the same distribution, at the 5% significance level ($p = 0.0326$ in 1P-CM and $p = 0.0001$ in 2P-CM, two-sided).

Result 1. In each treatment, the seller's actual revenue is significantly higher than the equilibrium revenue. On average, the actual revenues in 1P-CM and 2P-CM are around 24% and 83% above the equilibrium revenues, respectively.

Figure 3 shows how the deviation of actual average revenues from equilibrium average revenues changes over time in each treatment. As can be seen in the figure, the deviation in 2P-CM remains positive in all rounds, while the deviation in 1P-CM seems to converge to zero. Then, we focus on the last ten rounds in each treatment. For 2P-CM, the Wilcoxon matched-pairs signed-rank test rejects the null hypothesis that both the actual and the equilibrium revenues come from the same distribution, at the 1% significance level ($p = 0.0016$, two-sided). However, for 1P-CM, the same test cannot reject the null hypothesis at any conventional level ($p = 0.4638$, two-sided).

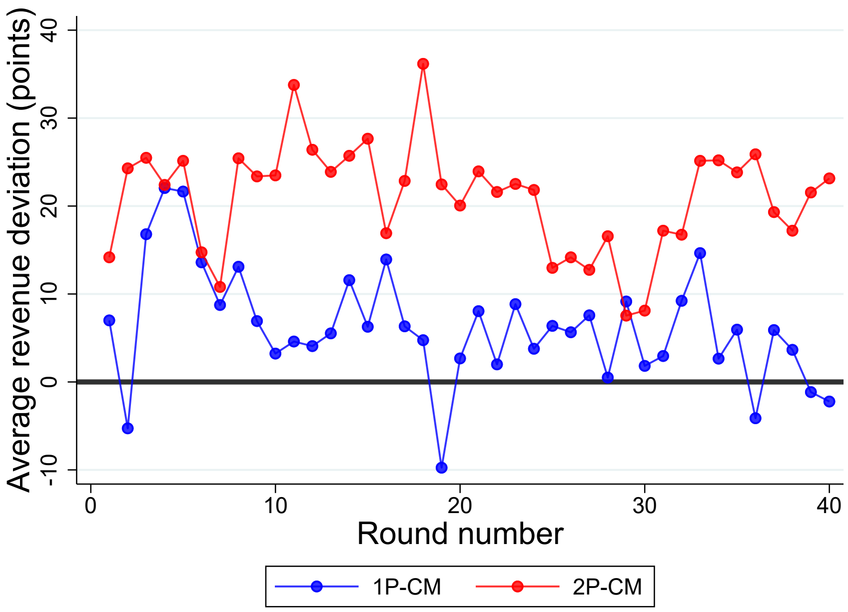


Figure 3: Time series of average revenue deviations in each treatment

Result 2. In the last ten rounds of 1P-CM, the seller's actual revenue is not significantly different from the equilibrium revenue. Even in the last ten rounds of 2P-CM, the actual revenue is significantly higher than the equilibrium revenue.

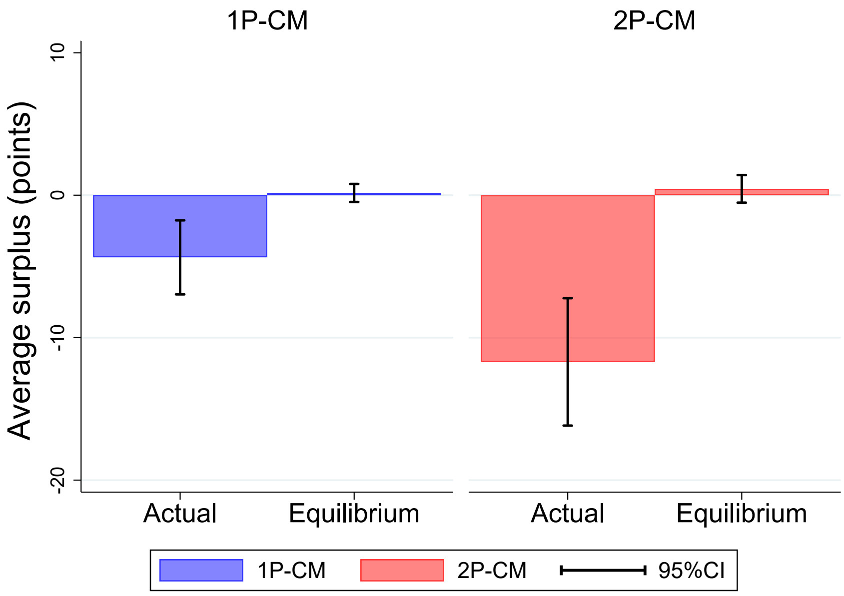
We now turn to testing for a treatment effect on revenues. As shown in Figure 1, the seller's actual average revenue is 14.26 points higher in 2P-CM than in

1P-CM. A two-sample t-test with unequal variances shows that this treatment effect is significantly different from zero at the 1% level ($p = 0.0071$, two-sided). Also, a two-sample Wilcoxon rank-sum test shows that this effect is significant at the same level ($p = 0.0055$, two-sided).

Result 3. The seller's actual revenue is significantly higher in 2P-CM than in 1P-CM. The actual average revenue in 2P-CM is around 44% above that in 1P-CM.

4.2 Bidder's surplus

Figure 4 summarizes actual and equilibrium surpluses for each bidder in each treatment. In the experiment, each bidder obtains -4.36 and -11.69 points on average in 1P-CM and 2P-CM, respectively. Given the realized valuations in the experiment, the ex-post surpluses per bidder in equilibrium are calculated at 0.15 and 0.44 points on average in 1P-CM and 2P-CM, respectively.



Note: Capped spikes represent 95% confidence intervals (CI).

Figure 4: Actual and equilibrium average surpluses in each treatment

For each treatment, the Wilcoxon matched-pairs signed-rank test rejects the null hypothesis that both the actual and the equilibrium average surpluses come from the same distribution, at the 1% level ($p = 0.0020$ in 1P-CM and $p = 0.0000$ in 2P-CM, two-sided).

Result 4. In each treatment, the actual surplus per bidder is significantly lower than the equilibrium surplus.

The two-sample t-test with unequal variances rejects the null hypothesis that the mean of the actual average surplus per bidder is equal between the two treatments, at the 1% level ($p = 0.0056$, two-sided). Also, the two-sample Wilcoxon rank-sum test shows that this treatment effect is significant at the same level ($p = 0.0044$, two-sided).

Result 5. The actual surplus per bidder is significantly lower in 2P-CM than in 1P-CM.

4.3 Total surplus

From Subsections 4.1 and 4.2, we observe that, in the experiment, the 1P-CM and 2P-CM auctions produce total surpluses of 23.23 and 22.83 points on average, respectively. These actual average total surpluses are 3 points lower than the average full surplus (around 26 points). The two-sample t-test with unequal variances cannot reject the null hypothesis that the mean of the average total surplus is equal between the two treatments, at any conventional level ($p = 0.4773$, two-sided).

Result 6. The actual total surpluses in 1P-CM and 2P-CM are not significantly different from each other. On average, these total surpluses amount to around 90% of the full surplus.

4.4 Bidding behavior

Finally, we conduct a panel data analysis of bidding behavior. This analysis helps us to find more fundamental causes for results in Subsections 4.1–4.3. Following the literature on auction experiments, we introduce two variables to measure how far subjects' bidding behavior is from truthful bidding. We define an *overbid amount* as a bid minus a valuation, and an *absolute bid deviation* as the absolute value of this overbid amount.

Figures 5 and 6 show how overbid amounts and absolute bid deviations change over time in each treatment, respectively. As shown in Figure 5, the average overbid amounts remain positive through all rounds of 2P-CM, but these amounts approach zero from above in later rounds of 1P-CM. Figure 6 shows that, in 1P-CM, the average absolute bid deviation is slightly decreasing over rounds. However, these deviations stay high in later rounds of each treatment. This implies that, in both treatments, there are many subjects who do not bid truthfully even with some learning and experience. In the last round of 1P-CM (2P-CM), 27.5%

(52.8%) of subjects overbid, 32.2% (17.1%) of subjects underbid, and the other subjects bid truthfully.

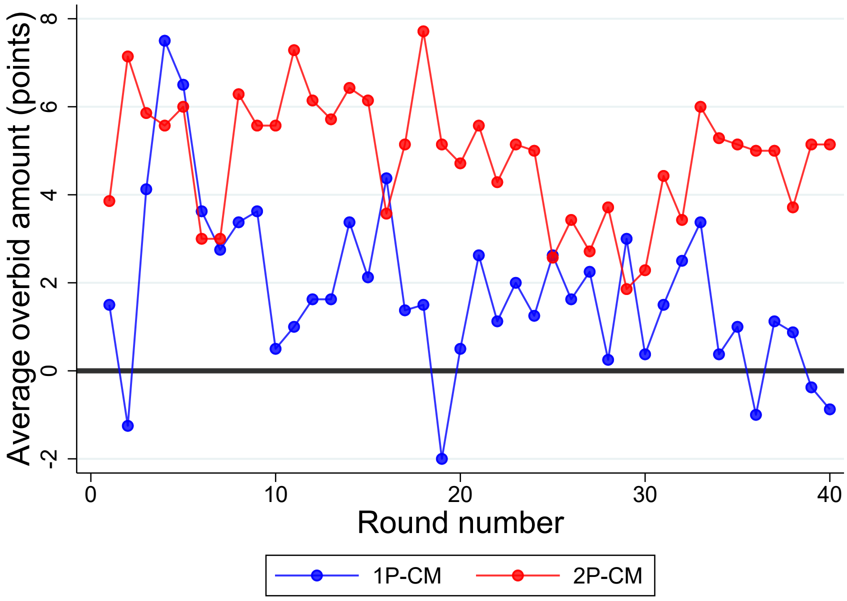


Figure 5: Time series of average overbid amounts in each treatment

These visual patterns are confirmed by multi-level mixed-effects regressions (Moffatt, 2015) reported in Table 5. The dependent variables are an overbid amount and an absolute bid deviation for each subject in each round. Each regression model has pair- and subject-level random effects. Regression results (1) and (3) show that the average overbid amount in 1P-CM (2P-CM) is 1.83 (4.86) points, and the average absolute bid deviation in 1P-CM (2P-CM) is 9.85 (11.10) points. For the overbid amount, these treatment coefficients are significantly different from zero at the 5% level ($p = 0.018$ in 1P-CM and $p = 0.000$ in 2P-CM, two-sided).

Result 7. In each treatment, the overbid amount is significantly higher than zero. On average, these overbid amounts in 1P-CM and 2P-CM account for 9% and 24% of realized valuations (around 20 points), respectively.

An F-test cannot reject the null hypothesis that the absolute bid deviations in both treatments are equal, at the 5% level ($p = 0.0918$, two-sided). However, the F-test rejects the null hypothesis that the overbid amounts in both treatments are equal, at the 5% level ($p = 0.0170$, two-sided).

Result 8. The overbid amount is significantly higher in 2P-CM than in 1P-CM. On average, the amount in 2P-CM is 2.6 times that in 1P-CM

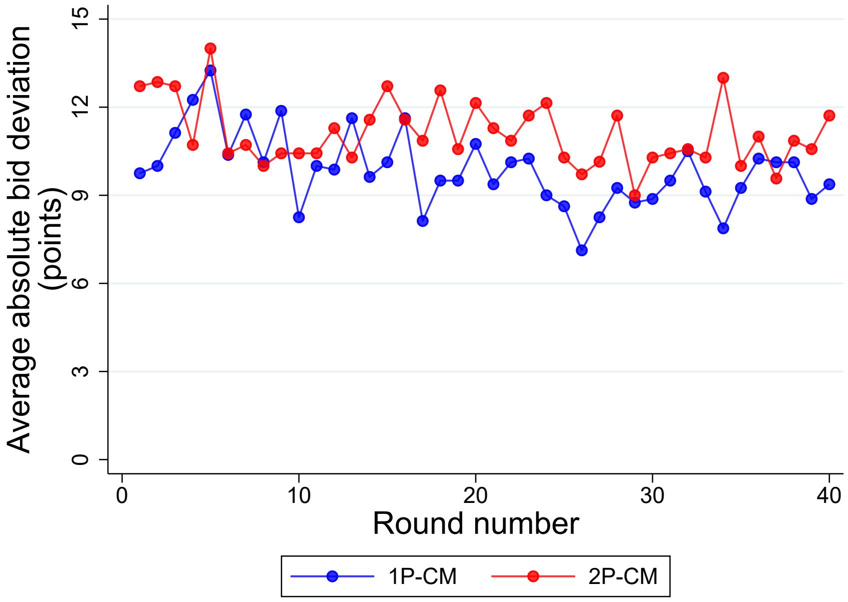


Figure 6: Time series of average absolute bid deviations in each treatment

Moreover, regression results (2) and (4) show that, in 1P-CM, the overbid amount and the absolute bid deviation decrease by 0.07 and 0.05 points per round ($p < 0.05$ and $p < 0.01$), respectively. In this treatment, some pairs of subjects might slowly learn to bid truthfully because the 1P-CM auction is dominance solvable as shown in Table 3.

Result 9. In 1P-CM, both the overbid amount and the absolute bid deviation are decreasing over rounds significantly. In 2P-CM, neither variable exhibits such a significant time trend.

Dependent variable	$bid - value$		$\lvert bid - value \rvert$	
	(1)	(2)	(3)	(4)
1P-CM	1.83***	3.28***	9.85***	10.88***
	(0.77)	(0.91)	(0.53)	(0.52)
2P-CM	4.86***	5.61***	11.10***	11.72***
	(1.00)	(1.16)	(0.51)	(0.76)
$(round\ number\ -1) \times$ *1P-CM*		−0.07**		−0.05***
		(0.03)		(0.01)
$(round\ number\ -1) \times$ *2P-CM*		−0.03		−0.03
		(0.05)		(0.03)
S.D. of the pair random effect	4.49	4.49	2.41	2.41
S.D. of the subject random effect	3.62	3.63	2.51	2.51
S.D. of the residual	13.31	13.29	10.09	10.08
Observations	6000	6000	6000	6000

Note: This table reports the results of multi-level mixed-effects regressions with pair and subject random effects. The models have no constant term. Standard errors with clustering at the pair level are shown in parentheses. *,** , and *** denote significance at 10%, 5% and 1%, respectively. S.D. means a standard deviation.

Table 5: Multi-level mixed-effects regressions for overbid amounts and absolute bid deviations

As shown by regression models (1) and (2) for overbid amounts, the between-pair standard deviations are larger than the between-subject standard deviations. Hence, the variation across pairs is substantial in magnitude. Indeed, Figure 7 shows that, in each treatment, the average overbid amount of each subject is positively correlated with the amount of the other subject in the same pair. This positive correlation may be partly due to "repeated-game effects" under the rule of partner matching. On the one hand, some pairs succeed in tacit collusion.[11] On the other hand, some pairs seem to engage in negative reciprocal behavior. To examine repeated-game behavior in each pair, it is useful to see Figures A.1 and A.2 in Appendix A. These figures show how pair-averaged overbid amounts change over time. Then, we run a linear regression with pair-averaged overbid amounts being the dependent variable and round numbers being the independent

[11] For example, in 2P-CM, there was a pair of subjects who always submitted the lowest bid (i.e., zero) from rounds 18 to 40. This pair corresponds to the leftmost point in the 2P-CM diagram of Figure 7. In a post-session questionnaire, a subject in this pair indicated that he made the zero bid in round 16 or 17 in order to establish cooperation with his opponent. Also, the latter subject answered that he tried to go along with his opponent.

variable. Using t-tests, we can show that linear time trends are downward for 8 pairs (pairs 12, 22, 31, 42, 61, 67, 69, and 72), and upward for 7 pairs (pairs 9, 36, 44, 46, 48, 66, and 73), at the 1% significance level.

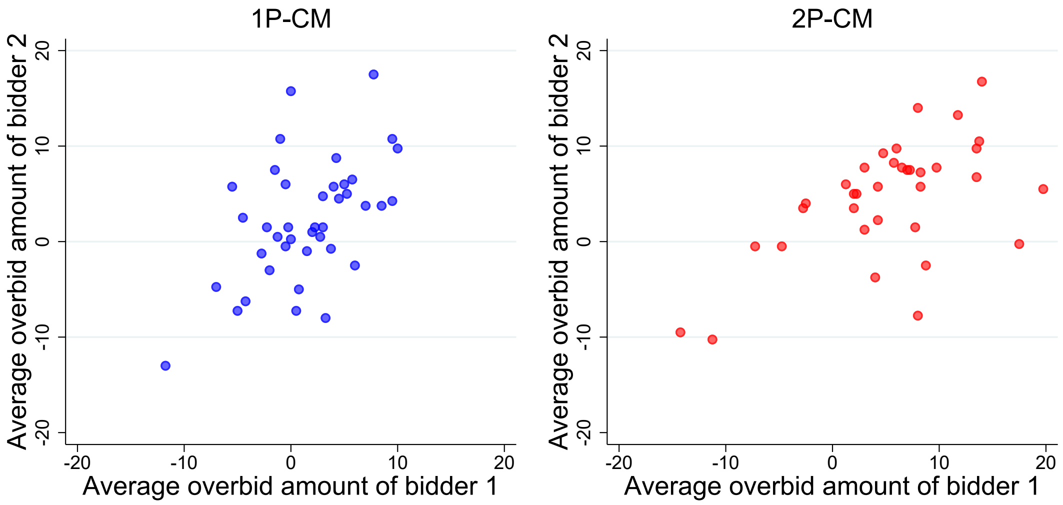


Figure 7: Average overbid amounts of two subjects in each pair

In standard 1P-auction experiments, subjects tend to bid higher than risk-neutral equilibrium bids, but lower than their valuations to obtain positive rents. However, in 1P-CM, subjects' bids are significantly higher than their *valuations*. This clearly shows that subjects respond to the presence of CM lotteries to some extent.

Overbidding in standard 2P-auction experiments is a well-ducumented phenomenon. Table 6 summarizes subjects' bidding behavior in previous 2P-auction experiments and the 2P-CM experiment.[12] In 2P-CM, 46.2% of subjects overbid, 18.9% of subjects underbid, and the other subjects bid truthfully. However, this frequency of overbidding is not so high compared to the frequencies in the other experiments. To test whether subjects overbid more frequently (or more aggressively) in a 2P-CM auction than in a 2P auction with no lottery, we should run a randomized controlled experiment with these two auction treatments. This question is left for future work.

[12] Overbidding is less frequent than underbidding only in Garratt et al. (2012). In their experiment, subjects were experienced participants in eBay auctions. Their evidence shows that bidders with experience as eBay sellers tend to underbid more frequently than those without this experience.

Study	Matching rule	# of bidders	# of rounds	Frequency (%) Truthful	Underbid	Overbid
KL (1993)	Stranger	5	23, 24	27.0	5.7	67.2
KL (1993)	Partner	10	12	32.5	9.6	57.9
ACK (2007)	Stranger	4	30	77.3	8.8	14.5
CF (2008)	Stranger	2	20	44.0	16.0	40.0
GWW (2012)	Stranger	5	2	21.2	41.3	37.5
BN (2016)	Stranger	2	24	46.1	14.0	39.8
Li (2017)	Partner	4	10	17.8	40.6	41.6
Tan (2020)	Stranger	2	15	30.0	18.8	51.2
Tan (2020)	Partner	2	15	26.1	30.1	43.9
BSK (2022b)	Partner	4	10	42.5	27.6	29.9
This study	Partner	2	40	34.9	18.9	46.2

Note: The frequency data comes from Table 2 of Kagel and Levin (1993, KL), Table 5 of Andreoni et al. (2007, ACK), page 1582 of Cooper and Fang (2008, CF), Table 1 of Garratt et al. (2012, GWW), raw experimental data of Bartling and Netzer (2016, BN) and Li (2017), Table 5 of Breitmoser and Schweighofer-Kodritsch (2022b, BSK), and Fig. 2 of Tan (2020).

Table 6: Bidding behavior in 2P-auction experiments and the 2P-CM experiment

5 Conclusion

This study is only a first step toward assessing the performance of CM auctions in experimental environments. There are many possible extensions. Perhaps most importantly, we should examine the performance in a setting where subjects can opt out of auctions. If subjects are risk-averse, then they may require some risk premia to participate in CM auctions. It is interesting to know how much a designer should pay risk premia to encourage their full participation. Even if subjects have (innate or induced) risk-neutral preferences, they may not participate in CM auctions. On this point, Milgrom (2004, p. 166) cautioned as follows:

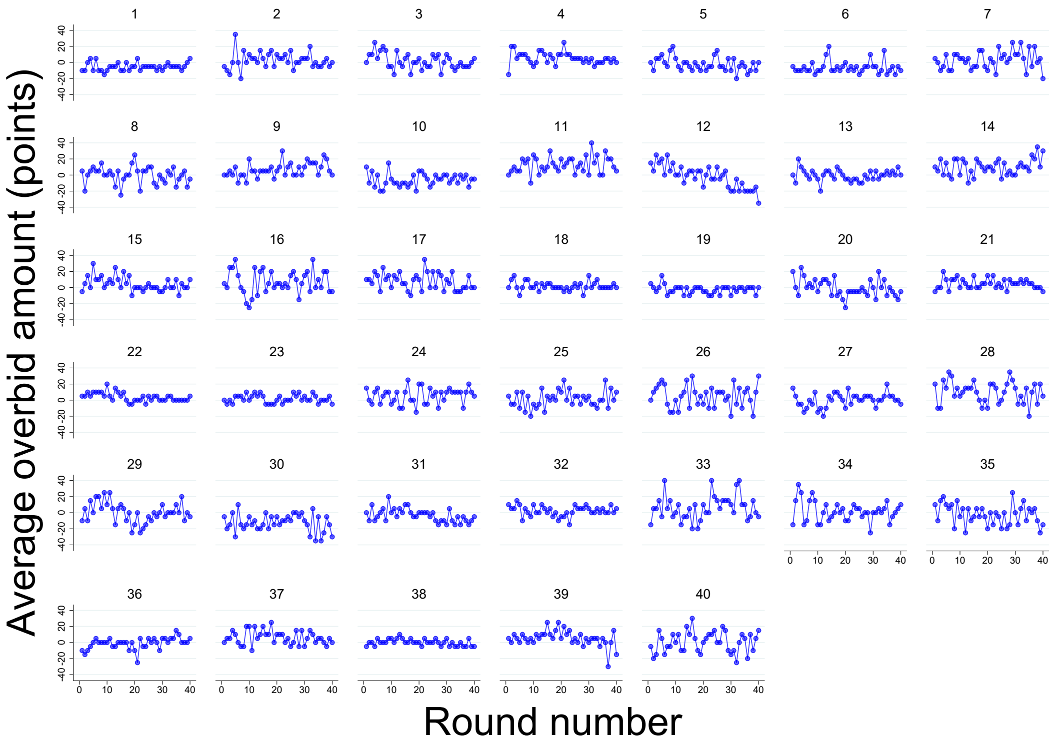
> *More generally, in real auctions, bidders frequently refuse to participate if the proposed mechanism seems strange or unfair. Many might apply these adjectives to a mechanism that links bids to side bets. Precedent and familiarity often limit the set of practically feasible designs.*

Next, it is important to examine how subjects' behavior is affected by the degree or sign of value correlation. When a common prior is close to an independent distribution, CM lotteries become extremely risky. Further, in 2P-CM auctions, a bidder's lottery prize can be *increasing* in his opponent's bid if their valuations are *negatively* correlated. In this case, subjects with spite motives have incentives to decrease their bids. These questions are left for future research.

Appendix A

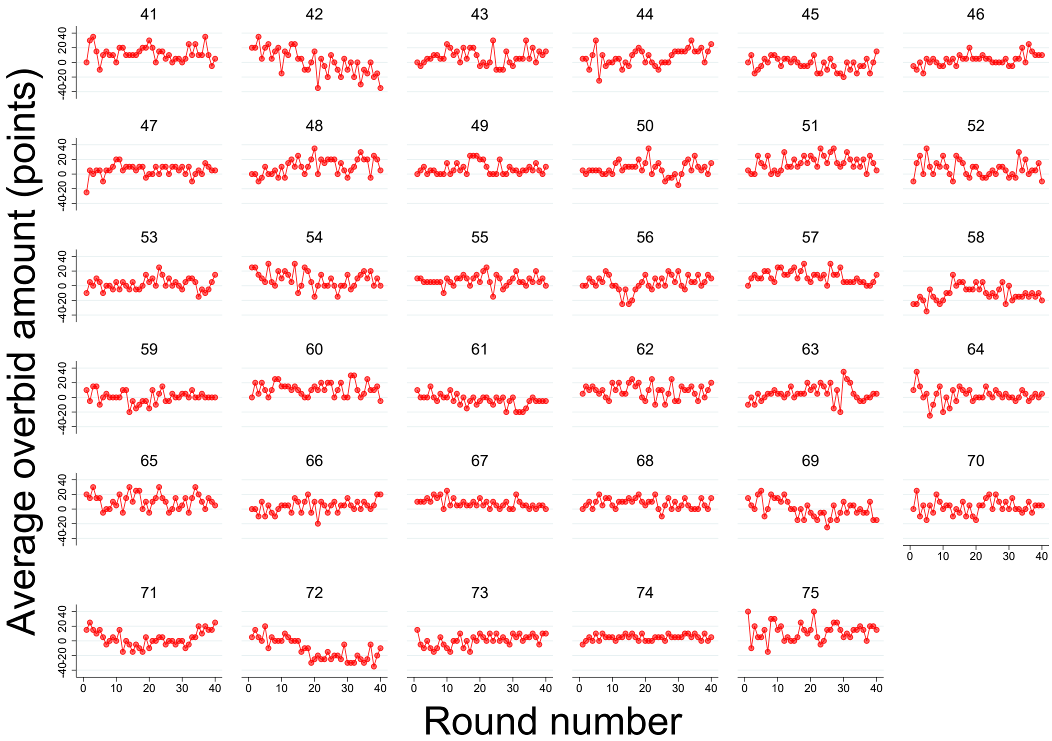
Treatment	# of subjects by year			
	2018	2019	2020	2021
1P-CM	25	19	27	9
2P-CM	15	21	21	13

Table A.1: Subjects by year of admission



Note: A pair ID number is shown on the top of each graph.

Figure A.1: Time series of average overbid amounts for each pair in 1P-CM



Note: A pair ID number is shown on the top of each graph.

Figure A.2: Time series of average overbid amounts for each pair in 2P-CM

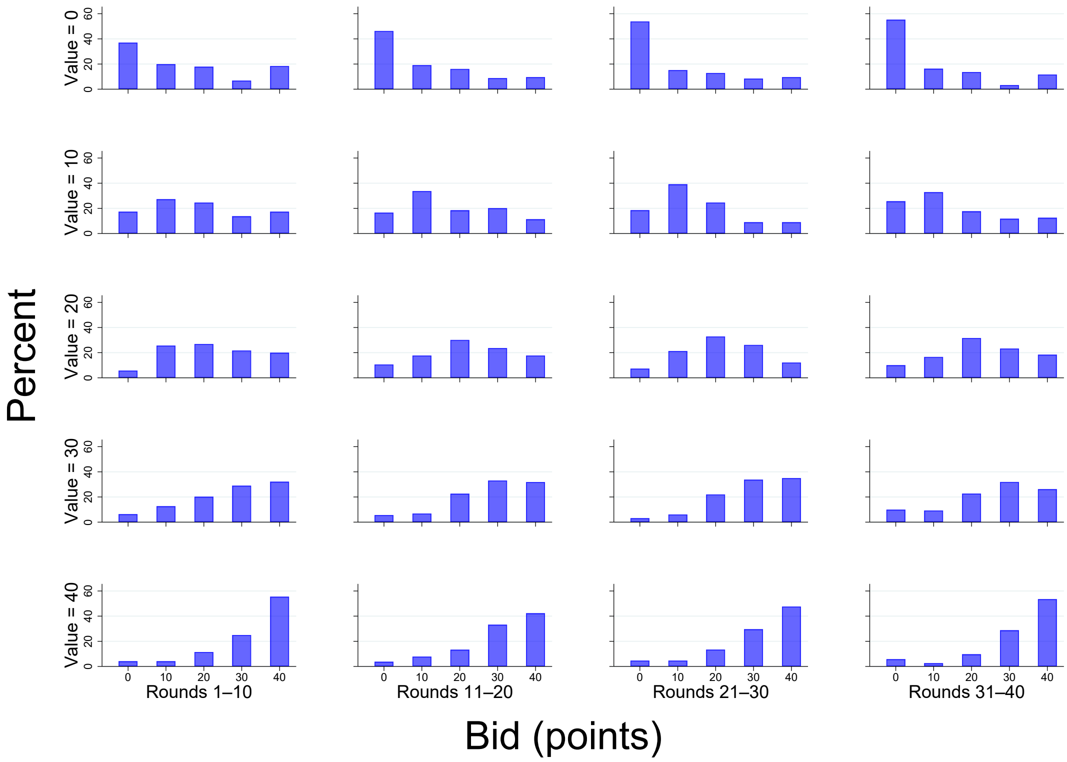


Figure A.3: Frequencies of bids conditional on valuations and rounds in 1P-CM

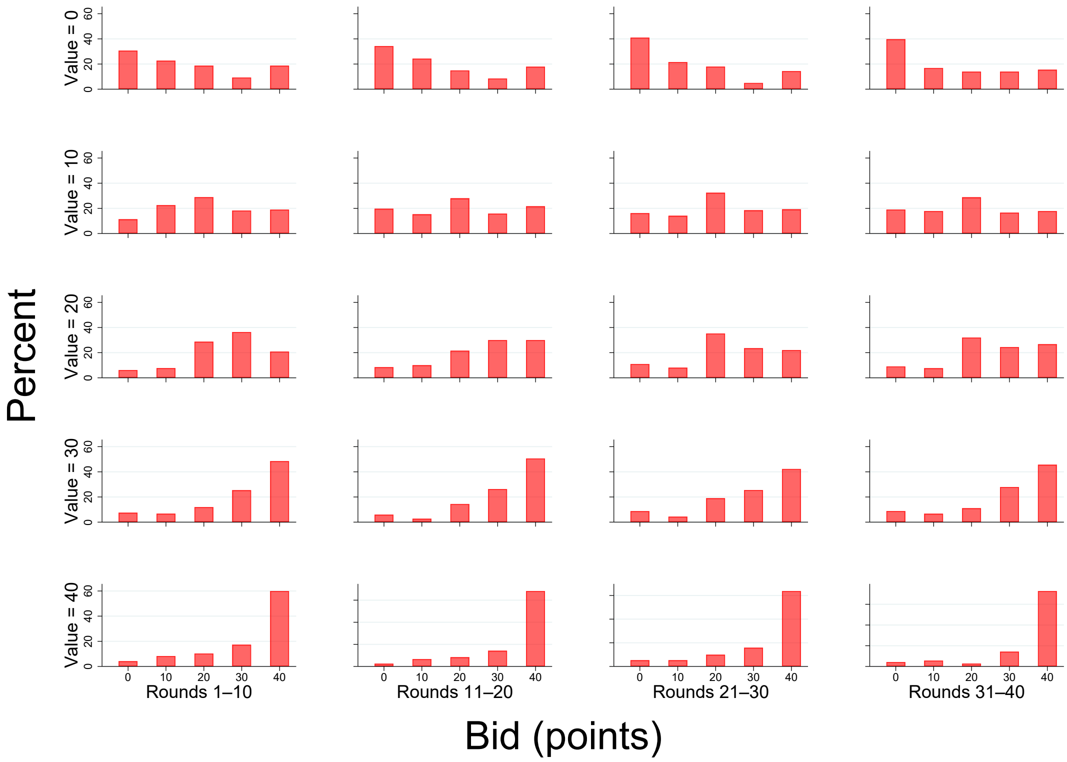


Figure A.4: Frequencies of bids conditional on valuations and rounds in 2P-CM

Dependent variable	*bid − value*			
Rounds	1–10	11–20	21–30	31–40
value = 0	15.24***	11.11***	10.19***	9.68***
	(1.48)	(1.28)	(1.46)	(1.67)
value = 10	8.16***	7.41***	4.99***	5.38***
	(1.08)	(1.22)	(1.13)	(1.74)
value = 20	2.88**	2.25**	1.64	2.16
	(1.30)	(1.13)	(1.19)	(1.34)
value = 30	−3.10***	−2.06*	−1.36	−3.86***
	(1.19)	(1.21)	(0.91)	(1.52)
value = 40	−7.85***	−9.49***	−8.18***	−8.11***
	(1.22)	(1.16)	(1.07)	(1.28)
S.D. of the pair random effect	3.72	3.47	3.72	5.83
S.D. of the subject random effect	3.77	2.99	3.92	3.85
S.D. of the residual	11.57	11.21	10.32	10.35
Observations	2800	2800	2800	2800

Note: This table reports the results of multi-level mixed-effects regressions with pair and subject random effects. The models have no constant term. Standard errors with clustering at the pair level are shown in parentheses. *,** , and *** denote significance at 10%, 5% and 1%, respectively. S.D. means a standard deviation.

Table A.2: Multi-level mixed-effects regressions for overbid amounts in 1P-CM

Dependent variable	*bid − value*			
Rounds	1–10	11–20	21–30	31–40
$value = 0$	15.59***	15.72***	13.80***	14.50***
	(1.71)	(1.83)	(1.87)	(1.92)
$value = 10$	11.25***	10.96***	10.05***	9.35***
	(1.29)	(1.57)	(1.77)	(1.82)
$value = 20$	5.83***	5.78***	4.48***	5.79***
	(1.31)	(1.58)	(1.56)	(1.63)
$value = 30$	−0.00	0.30	−0.95	−0.41
	(1.35)	(1.36)	(1.68)	(1.84)
$value = 40$	−7.30***	−5.59***	−7.93***	−5.46***
	(1.39)	(1.49)	(1.48)	(1.53)
S.D. of the pair random effect	4.31	5.92	6.53	7.77
S.D. of the subject random effect	4.45	3.50	4.95	3.82
S.D. of the residual	10.96	10.76	9.76	9.59
Observations	3200	3200	3200	3200

Note: This table reports the results of multi-level mixed-effects regressions with pair and subject random effects. The models have no constant term. Standard errors with clustering at the pair level are shown in parentheses. *,** , and *** denote significance at 10%, 5% and 1%, respectively. S.D. means a standard deviation.

Table A.3: Multi-level mixed-effects regressions for overbid amounts in 2P-CM

References

J. Andreoni, Y.K. Che, and J. Kim. Asymmetric information about rivals' types in standard auctions: An experiment. *Games and Economic Behavior*, 59(2): 240–259, 2007.

O. Armantier and N. Treich. Subjective probabilities in games: An application to the overbidding puzzle. *International Economic Review*, 50(4):1079–1102, 2009.

B. Bartling and N. Netzer. An externality-robust auction: Theory and experimental evidence. *Games and Economic Behavior*, 97:186–204, 2016.

S. Bikhchandani. Information acquisition and full surplus extraction. *Journal of Economic Theory*, 145(6):2282–2308, 2010.

T. Börgers. *An Introduction to the Theory of Mechanism Design*. Oxford University Press, 2015. (with a chapter by Daniel Krähmer and Roland Strausz).

Y. Breitmoser and S. Schweighofer-Kodritsch. Obviousness around the clock. *Experimental Economics*, 25(2):485–513, 2022a.

Y. Breitmoser and S. Schweighofer-Kodritsch. Supplement to "obviousness around the clock". *Experimental Economics*, 2022b. https://static-content.springer.com/esm/art%3A10.1007%2Fs10683-021-09720-z/MediaObjects/10683_2021_9720_MOESM1_ESM.pdf.

S. Brusco. Unique implementation of the full surplus extraction outcome in auctions with correlated types. *Journal of Economic Theory*, 80(2):185–200, 1998.

G. Carroll. Robustness in mechanism design and contracting. *Annual Review of Economics*, 11:139–166, 2019.

D.L. Chen, M. Schonger, and C. Wickens. oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016. https://doi.org/10.1016/j.jbef.2015.12.001.

Y.C. Chen and S. Xiong. Genericity and robustness of full surplus extraction. *Econometrica*, 81(2):825–847, 2013a.

Y.C. Chen and S. Xiong. Supplement to "Genericity and robustness of full surplus extraction". *Econometrica*, 2013b. https://www.econometricsociety.org/ecta/Supmat/10123_proofs.pdf.

K.S. Chung and J.C. Ely. Foundations of dominant-strategy mechanisms. *Review of Economic Studies*, 74(2):447–476, 2007.

D.J. Cooper and H. Fang. Understanding overbidding in second price auctions: An experimental study. *Economic Journal*, 118(532):1572–1595, 2008.

J.C. Cox, B. Roberson, and V.L. Smith. Theory and behavior of single object auctions. *Research in Experimental Economics*, 2(1):1–43, 1982.

J. Crémer and R.P. McLean. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53 (2):345–361, 1985.

J. Crémer and R.P. McLean. Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257, 1988.

R. Engelbrecht-Wiggans and E. Katok. Regret and feedback information in first-price sealed-bid auctions. *Management Science*, 54(4):808–819, 2008.

E. Filiz-Ozbay and E.Y. Ozbay. Auctions with anticipated regret: Theory and experiment. *American Economic Review*, 97(4):1407–1418, 2007.

D. Fudenberg and J. Tirole. *Game theory*. MIT press, 1991.

S. Füllbrunn, D.J. Janssen, and U. Weitzel. Risk aversion and overbidding in first price sealed bid auctions: New experimental evidence. *Economic Inquiry*, 57 (1):631–647, 2019.

R.J. Garratt, M. Walker, and J. Wooders. Behavior in second-price auctions by highly experienced eBay buyers and sellers. *Experimental Economics*, 15(1): 44–57, 2012.

B. Greiner. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125, 2015. https://doi.org/10.1007/s40881-015-0004-4.

R.M. Harstad. Dominant strategy adoption and bidders' experience with pricing rules. *Experimental Economics*, 3(3):261–280, 2000.

A. Heifetz and Z. Neeman. On the generic (im)possibility of full surplus extraction in mechanism design. *Econometrica*, 74(1):213–233, 2006.

J.H. Kagel and D. Levin. Independent private value auctions: Bidder behaviour in first-, second- and third-price auctions with varying numbers of bidders. *Economic Journal*, 103(419):868–879, 1993.

J.H. Kagel and D. Levin. Auctions: A survey of experimental research. In J.H. Kagel and A.E. Roth, editors, *The Handbook of Experimental Economics, Volume 2*, pages 563–637. Princeton University Press, 2015.

J.H. Kagel, R.M. Harstad, and D. Levin. Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica*, 55 (6):1275–1304, 1987.

I. Krajbich, C. Camerer, J. Ledyard, and A. Rangel. Using neural measures of economic value to solve the public goods free-rider problem. *Science*, 326(5952): 596–599, 2009.

I. Krajbich, C. Camerer, and A. Rangel. Exploring the scope of neurometrically informed mechanism design. *Games and Economic Behavior*, 101:49–62, 2017.

J.J. Laffont and D. Martimort. Mechanism design with collusion and correlation. *Econometrica*, 68(2):309–342, 2000.

S. Li. Obviously strategy-proof mechanisms. *American Economic Review*, 107 (11):3257–87, 2017.

E. Maskin and J.G. Riley. Auctioning an indivisible object. Technical report, Kennedy School, Harvard University, 1980.

H. Matsushima. Mechanism design with side payments: Individual rationality and iterative dominance. *Journal of Economic Theory*, 133(1):1–30, 2007.

R.P. McAfee and P.J. Reny. Correlated information and mecanism design. *Econometrica*, 60(2):395–421, 1992.

P. Milgrom. *Putting Auction Theory to Work*. Cambridge University Press, 2004.

P.G. Moffatt. *Experimetrics: Econometrics for Experimental Economics*. Macmillan International Higher Education, 2015.

J. Morgan, K. Steiglitz, and G. Reis. The spite motive and equilibrium behavior in auctions. *Contributions in Economic Analysis & Policy*, 2(1):1–25, 2003.

R.B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6 (1):58–73, 1981.

H. Pham and T. Yamashita. Auction design with approximate common prior. *SSRN*, 2021.

T.A. Rietz. Implementing and testing risk-preference-induction mechanisms in experimental sealed-bid auctions. *Journal of Risk and Uncertainty*, 7(2):199–213, 1993.

M.H. Riordan and D.E.M. Sappington. Optimal contracts with public ex post information. *Journal of Economic Theory*, 45(1):189–199, 1988.

J. Robert. Continuity in auction design. *Journal of Economic Theory*, 55(1): 169–179, 1991.

A.E. Roth and M.W. Malouf. Game-theoretic models and the role of information in bargaining. *Psychological Review*, 86(6):574, 1979.

M. Schneider and D. Porter. Effects of experience, choice architecture, and cognitive reflection in strategyproof mechanisms. *Journal of Economic Behavior & Organization*, 171:361–377, 2020.

C.H.Y. Tan. Overbidding and matching rules in second-price auctions: An experimental study. *Journal of Behavioral and Experimental Economics*, 84:101507, 2020.

J.M. Walker, V.L. Smith, and J.C. Cox. Inducing risk-neutral preferences: An examination in a controlled market environment. *Journal of Risk and Uncertainty*, 3(1):5–24, 1990.

R. Wilson. Game-theoretic analyses of trading processes. In T.F. Bewley, editor, *Advances in Economic Theory*, pages 33–70. Cambridge University Press, 1987.