

# An Experimental Study of Sponsored-Search Auctions\*

Yeon-Koo Che<sup>†</sup>

Columbia University

Syngjoo Choi<sup>‡</sup>

Seoul National University and UCL

Jinwoo Kim<sup>§</sup>

Seoul National University

September 2015

## Abstract

We study the Generalized Second Price auctions—a standard method for allocating online search advertising—experimentally, considering both the static environment assumed by the prevailing theory and a dynamic game capturing the salient aspects of real-world search advertising auctions. Subjects of our experiment bid consistently with the leading equilibrium notions, but exhibit significant overbidding relative to the Vickrey-Clarke-Groves (VCG) outcome favored as an equilibrium selection in the literature. The observed bidding behavior is well explained by a model that explicitly accounts for the strategic uncertainty facing a bidder, which suggests strategic uncertainty as a source of the observed departure from the VCG outcome. Meanwhile, the observed bidding behavior in the dynamic game resembles the behavior in the static

---

\*We are grateful to Jacob and Michelle Goeree, John Kagel, Dan Levin, Michael Ostrovsky, Michael Schwarz, Yan Chen, and the participants at the AMMA 2011, Monash University, Ohio State University, University of East Anglia, University of Michigan, Universitat Pompeu Fabra, University of Zurich, and the WCU Market Design Conference held at Yonsei University in August 2010, for valuable comments, and to Brian Wallace for writing the experimental program, and to Tom Rutter for helping us run the experiment. Che acknowledges the support by NSF (#1023818). Che and Kim acknowledge the support by the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology through its grant of Global Research Network (2013S1A2A2035408). Choi acknowledges the Economic and Social Research Council (UK) Grant No. RES-061-25-0348 and via the ELSE for financial support.

<sup>†</sup>Department of Economics, Columbia University, 420 West 118th St, 10161AB, New York, NY 10027, USA (Email: yc2271@columbia.edu, URL: <http://www.columbia.edu/~yc2271>).

<sup>‡</sup>Department of Economics, Seoul National University, 1 Gwanak-ro Gwanak-gu, Seoul 151-742, South Korea, UK (Email: syngjoo.choi@ucl.ac.uk, URL: <http://www.homepages.ucl.ac.uk/~uctpsc0>).

<sup>§</sup>Department of Economics, Seoul National University, 1 Gwanak-ro Gwanak-gu, Seoul 151-742, South Korea (Email: jkim72@gmail.com, URL: <https://sites.google.com/site/jkim72/home>).

game. Our analysis thus lends support to the use of a static game as modeling proxy, but calls into question the prevailing equilibrium selection.

JEL Classification: C92, D44, M3.

Keywords: online advertising, sponsored search auction, generalized second price auction, experiment.

## 1 Introduction

Search engines such as Google, Yahoo! and Microsoft sell online search spaces to advertisers. In comparison with conventional advertising, online search advertising is highly targetable, and thus is an effective means for finding buyers. Naturally, the sponsored search auctions have become a major revenue source for search firms. In 2007, search advertising accounted for more than \$21 billion of revenue for search firms in US.<sup>1</sup> The auction format used for selling ad spaces has evolved, with a few adjustments along the way, to what is now known as the *generalized second price* (GSP) auctions.

Under the GSP, advertisers bid per-click prices, and these bids are converted into per-impression bids—their per-click bids multiplied by the estimated click through rates—to determine the assignment of ad positions. Specifically, the highest bidder (in per-impression bid) is assigned the top position, the second-highest bidder is assigned the next best, and so on. A winner of each ad position then pays the smallest price per click that would have won that position. If the number of clicks depends only on one’s position, as is often assumed, per-impression bids essentially coincide with per-click bids, so each winning bidder simply pays a per-click price that equals the bid submitted by the next-highest bidder.

The prevailing theory considers the GSP in a static model in which advertisers bid *simultaneously with complete information about others’ preferences* (Edelman et al., 2007, henceforth EOS; and Varian, 2007). EOS and Varian then focus on a class of Nash equilibria, called *locally envy-free* or *symmetric*, in which no bidder wishes to exchange his winning position and the associated price with others’ positions and the prices they are paying for them. The Symmetric Nash Equilibrium (SNE) concept predicts efficient allocation of ad positions but admits a plethora of equilibrium prices, including those that would obtain if the Vickrey-Clarke-Groves (VCG) mechanism were employed. This VCG equilibrium is the most preferred by bidders among all SNE’s, and is suggested as the most plausible.<sup>2</sup>

---

<sup>1</sup>See [http://www.iab.net/media/file/IAB.PwC.2007\\_full\\_year.pdf](http://www.iab.net/media/file/IAB.PwC.2007_full_year.pdf). This revenue includes over 90% of Google’s revenue and 50% of Yahoo! and MS’s revenues (more precisely, the revenue of Online Services Division of MS) according to each firm’s report of annual revenue.

<sup>2</sup>See Section 2 for several arguments that have been made in the literature in support of selecting the VCG outcome.

While the theory provides useful insights on GSP auctions, it raises two issues. First, unlike the theory, sponsored-search auctions in practice take place continuously in real time—in principle, whenever a user types in a search query—and also bidders are unlikely to have complete information about one another’s preferences. This means that advertisers face complex dynamic interactions which may provide them with opportunities to learn and adjust their behavior over time. The actual practice is therefore best described by a dynamic game in which bidders with incomplete information play repeatedly over time. It is unclear whether the static complete information model can adequately represent this rich dynamic environment. Second, the theory lacks a sharp prediction due to the multiplicity of equilibria. Although the literature suggests the VCG outcome as the most salient, there is no compelling theoretical argument or empirical evidence supporting the selection of this outcome.

The current paper investigates these issues via a laboratory experiment. At its core, our experiment induces two environments. The basic control environment is the **static complete-information** game (henceforth **SC**) used in theory, wherein *subjects play one-shot GSP game with complete information about one another’s preferences*. The main treatment environment is a **dynamic incomplete-information** game (henceforth **DI**) that captures the salient features of the GSP game in practice, wherein *subjects play the GSP games repeatedly, with possible feedback and learning, but without complete information about their opponents’ preferences*. Since the dynamic game differs from the static game with respect to *both* timing and information, we also consider a **static incomplete-information** game (henceforth **SI**) in which *bidders play one-shot GSP game under incomplete information about their opponents’ preferences*. This additional game serves as a bridge between the static complete information game and dynamic incomplete information game.

Specifically, our experiment considers three bidders competing to obtain one of the two bundles,  $A$  and  $B$ , each containing  $c_A$  and  $c_B$  units of a (fictitious) homogeneous commodity, respectively. The two bundles represent two advertising positions, and the units of the commodity in a bundle represent the number of clicks an ad position receives for a given period.<sup>3</sup> Bundle  $A$  contains more units of the commodity than does bundle  $B$ , i.e.,  $c_A > c_B > 0$ , and the ratio of the units in bundle  $B$  and  $A$ ,  $\frac{c_B}{c_A}$ , captures “clicks decays” across ad positions. In fact, the nature of strategic environment depends crucially on the magnitude of click decays. When the ratio  $\frac{c_B}{c_A}$  is close to zero (i.e., the unit difference is large),  $A$  becomes so much more attractive than  $B$ , so the competition becomes essentially about winning bundle

---

<sup>3</sup>We are thus assuming that the number of clicks an ad receives depends only on the position it is placed in. This assumption may not be realistic, but it serves to simplify the strategic environment for the experimental subjects, and more importantly, to facilitate the testing of the theory, which makes the same assumption for the most part. EOS and Athey and Nekipelov (2010) extend the theory to introduce the advertiser-specific factor in click generation.

A. The game thus becomes close in nature to standard second-price auctions. By contrast, if the ratio  $\frac{c_B}{c_A}$  is close to one (i.e., the unit difference is small), both bundles are almost equally good, so bidders compete to win *either* bundle at a cheaper price per unit. The strategic environment thus resembles that of Bertrand game. The difference in the strategic interaction provides an opportunity to test whether players respond strategically to the environments, and this difference provides testable restrictions on the subjects' strategic responses, which is exploited in our experimental design.

Our experiment yields several results. First, the GSP auctions turned out reasonably good efficiency performances in our experiment, attaining 76-93% of the maximum possible surplus improvement over random assignment of the positions. Second, the GSP in the lab yields revenue that is within the upper bound of the symmetric Nash equilibrium, but exceeds consistently the revenue corresponding to the particular selection, namely the VCG outcome, specifically at the median by 3-4% in the static complete information game, by 30-40% in the static incomplete information game, and by 9-18% in the dynamic incomplete information game. The findings in efficiency and revenue, while suggestive, may not provide detailed account of the behavior of the subjects, so our main analysis focuses on the subjects' bidding behavior.

A closer inspection of the subjects' bidding behavior shows a broad consistency with Nash equilibrium and symmetric (locally envy free) Nash equilibrium, but a systematic departure from its particular selection, the VCG outcome. Specifically, the majority of bidders with the lowest per-unit values bid close to their true values, as predicted by the theory, while a significant fraction of subjects also bid above their values. This latter pattern of behavior is in common with the overbidding patterns documented in the experimental literature of standard, single-unit, second-price auctions (see Kagel et al., 1987; Kagel and Levin, 1993; Andreoni et al., 2007; and Cooper and Fang, 2008). On the other hand, the bidders with the intermediate per-unit values—whose behavior is crucial for testing the theory as will be seen later—bid significantly higher than the VCG level in all treatments: the median subjects overbid the VCG prediction by 19-22% in the SC game, by 42-81% in the SI game, and by 22-25% in the DI game. When we compare the cumulative distributions of observed bids and the VCG prediction, this overbidding holds even in the sense of stochastic dominance in all treatments. We then check if the overbidding found in our data is due to subjects' non-strategic behavior or insincere bidding. While observed bids do not conform to the VCG prediction, a large fraction of bidders played best responses to one another at least in the static sense. Subjects' bidding behavior also responded to a change in the unit ratio of the two bundles in a way qualitatively consistent with the theory, indicating that subjects understood the strategic environments they operated in.

We next find that the subjects' behavior in the dynamic game of incomplete information (DI game) resembles the behavior in the static game of complete information (SC game).

Subjects in our dynamic treatment are seen to behave adaptively optimally. The resulting learning and feedback cause them to respond to the values of their opponents in the DI treatment qualitatively similarly to the way they do in the SC treatment, *even though the subjects do not directly observe their opponents' values*. This last finding suggests that the use of static complete information game may be a reasonable modeling proxy for the more realistic dynamic game. This has an important implication for analyzing complex games such as sponsored search auctions. Such games are often difficult to analyze in a descriptively accurate form due to complex dynamics and informational incompleteness, forcing analysts to rely on a very stylized static complete information game as a modeling framework. Our finding validates such an approach by showing that not much empirical realism would be lost.

At the same time, the support for theoretical predictions on the bidding behavior is less clear cut. While the observed behavior cannot reject the predictions based on the symmetric Nash equilibrium concepts, the set of symmetric Nash equilibria is very large. More importantly, the particular selection proposed by EOS, namely the VCG, is clearly rejected by the current experimental data. The current study, along with the lack of compelling theoretical justification for the VCG outcome, thus calls this equilibrium selection into question.

In light of this finding, the observed bidding behavior calls for a more precise account. Hence, we consider an alternative model in which the mid-value bidder faces strategic uncertainty about opponents' strategies and responds rationally to it. The mid-value bidder's behavior, if rational, should be explained by the beliefs he holds on how the lowest-value and the highest-value bidders will bid. In this regard, we first show in the current setup of three bidders and two positions that bidding above the VCG benchmark is never optimal for the mid-value bidder, if he believes that the lowest-value bidder bids no higher than her own value (regardless of the highest-value bidder's behavior). We investigate whether the extent of overbidding relative to the VCG exhibited by the mid-value bidder can be rationalized by the extent of overbidding by the lowest-value bidder relative to their values. More specifically, we construct a model in which the mid-value bidder bids optimally given his beliefs, and his beliefs are in turn consistent with the observed behavior of the other bidders. When confronted with the experimental data, this model of strategic uncertainty does a fairly good job of explaining the overbidding outcome in the data. We thus conclude that strategic uncertainty, combined with the possibility of the lowest-value bidder's bidding above her own value, can explain the observed departure from the VCG prediction.

The current paper has several broad contributions. First, to our knowledge, the current paper constitutes a first systematic experimental treatment of the GSP auctions, providing a comprehensive account of its revenue and efficiency performance and of bidder behavior,

as well as the testing of the prevailing theory.<sup>4</sup> Considering the significance of the GSP auctions in the increasingly important internet auctions, the research on the subject matter has been surprisingly sparse, particularly on the empirical front. Two important exceptions are Börgers et al. (2013) and Athey and Nekipelov (2010). The former authors fit the static complete information framework to Yahoo data and bound the set of per-click values consistent with Nash equilibria via the revealed preferences methodology. The latter authors estimate a richer model that avoids the multiplicity of equilibria by introducing asymmetric information on the part of advertisers about quality ratings of their opponents.<sup>5</sup> Analyzing real auction data is obviously very useful, but an experimental study in the lab can serve an important purpose, not easily fulfilled by field studies. There are often no direct observations on advertisers' preferences for ad positions or on their information regarding their opponents' preference. Theory often closes the data gap, but in the case of GSP auctions, this role of theory is limited due to the high level of abstraction and non-unique predictions. Lab experiment, with the ability to control subjects' preferences and information to a large degree, can actually test the theory itself.

Second, our study employs distinct experimental treatments to study whether the stylized static full information model can approximate the dynamic incomplete information environment of the real GSP auction and whether the bidders behave according to the equilibrium predictions under the chosen model. Our approach allows us to study whether the selection of a particular modeling framework is justified separately from whether the equilibrium prediction from the chosen model is empirically valid, and this approach can be useful beyond the current setting. In market design research, some design option entails a strategic environment that is too complicated to tractably analyze, and this forces analysts to settle on

---

<sup>4</sup>The only other experimental work we are aware of is Fukuda et al. (2013). Their main purpose is different from ours; they compare the performances between the GSP and VCG auctions (that is, not just the VCG outcome). They found that the revenue and efficiency performance is similar between the GSP and VCG auctions and that NE and SNE are more frequently observed in the VCG auction than in the GSP auction. Their GSP auction shares some common elements with our dynamic environment but there are some important differences in the design. Specifically, due to the fact that their primary interest is in the comparison between VCG and GSP auctions, they do not vary the GSP auctions in terms of information (complete vs. incomplete information), timing (static vs. dynamic game), and decay parameters ( $c_A = 20$  vs. 11). Further, they focus on a single profile of values for all groups and for all rounds of experiment. These limitations of their design make it difficult to compare their results in the GSP auction with ours. Nonetheless, they found that subjects largely bid above the lower bound of SNE but below its upper bound. They also found some evidence of learning over time.

<sup>5</sup>A related work is Jeziorski and Segal (2009) which estimates a rich demand model that allows for the interdependence of advertisers' preferences for ad positions. Their paper focuses on the user's click behavior, and thus does not study advertisers' bidding behavior. Ostrovsky and Schwarz (2009) conducted a large-scale field experiment using the keywords in Yahoo! search engine and measured the impact of introducing reserve prices on the GSP auction revenue.

a much stylized abstraction of the true environment. In particular, full information static game has been adopted as a tractable modeling shortcut in problems such as combinatorial auctions and school choice analyses.<sup>6</sup> Experimental studies may be performed to study the original environment, but without separating the two aspects of the theory—its modeling framework and its equilibrium prediction under the chosen model—one would not be able to evaluate the adequacy of the theory. We would expect our current methodology to be useful in such an environment.

Last, our paper contributes to the empirical understanding of Bayesian learning in dynamic games. Theory suggests that Bayesian learning achieved by myopic players through repeated interaction (with little information on their opponents) would lead them to play a full-information Nash equilibrium of the stage game (see Jordan (1991) for instance).<sup>7</sup> Our analysis provides some empirical support for this theory, albeit in a special context. Further, it provides some empirical understanding on how players would select a particular Nash equilibrium in case there are multiple, an issue on which theory does not provide much guide.

The rest of the paper is organized as follows. Section 2 discusses the theoretical framework about GSP auctions and our research questions. Section 3 describes the experimental design and the details of the procedures. The experimental results are gathered in Section 4. Section 5 provides an alternative model of strategic uncertainty and presents the goodness of fit for this model. Section 6 concludes by summarizing the findings.

## 2 Theoretical framework and research questions

### 2.1 Leading theory of GSP auctions

EOS and Varian consider a static game of complete information in which bidders submit simultaneously and independently their per unit bids with complete information about other bidders’ per-click values. We describe the theoretical predictions of this model using a setup that we shall utilize in our experimental design.

Suppose there are three bidders,  $i = 1, 2, 3$ , and two bundles (or “positions”),  $A$  and  $B$ , containing respectively  $c_A$  and  $c_B$  units of a commodity (“clicks”), where  $c_A > c_B > 0$ . The value per unit for bidder  $i$  is denoted by  $v_i$ . If a bidder  $i$  with per unit value  $v_i$  wins a bundle

---

<sup>6</sup>Full-information Nash equilibrium analysis has been quite standard in the spectrum auction analyses, including the analysis of the FCC’s Simultaneous Ascending auctions (Milgrom, 2000) and core-selecting auctions (Day and Milgrom, 2008). In the school choice literature, the Boston mechanism is often analyzed in the full information Nash framework (see Ergin and Sönmez, 2006; Pathak and Sönmez, 2008).

<sup>7</sup>Cox et al. (2001) provides an experimental test of Jordan (1991)’s result, showing that the convergence tends to occur in a game with unique equilibrium but not in a game with multiple equilibria.

of  $c_k$  units of the good and pays  $p$  per unit, his payoff is  $c_k(v_i - p)$ . The allocation and the payment are determined under the GSP auction as follows. The highest bidder wins bundle  $A$  and pays the second-highest bid per unit (so his total payment is that bid multiplied by  $c_A$ ), the second-highest bidder wins bundle  $B$  and pays the lowest bid per unit, and the lowest bidder wins nothing and pays nothing.<sup>8</sup>

EOS and Varian assume that bidders make bids simultaneously with full information about the entire value profile  $(v_1, v_2, v_3)$ . Consider first Nash equilibria of this game. Without loss of generality, suppose the bidders bid  $b_1 > b_2 > b_3$ . (No assumption has been made on the relative rankings of  $v_1, v_2, v_3$  so far.) This bid profile forms a Nash equilibrium (henceforth referred to as NE) if and only if

$$c_A(v_1 - b_2) \geq \max\{c_B(v_1 - b_3), 0\}, \quad (1)$$

$$c_B(v_2 - b_3) \geq \max\{c_A(v_2 - b_1), 0\}, \quad (2)$$

$$b_2 \geq v_3. \quad (3)$$

Condition (1) means that the highest bidder—the winner of position  $A$ —has no incentive to deviate in order to win  $B$  instead at the price the current winner of  $B$  is paying or to win nothing. Condition (2) means that the second-highest bidder—the winner of position  $B$  has no incentive to deviate to win position  $A$  at a price equal to the highest bid or to win nothing. Condition (3) similarly requires that the lowest bidder has no incentive to deviate. Note that the Nash equilibrium allocation need not be efficient (i.e., assortative). In fact, the only allocations Nash equilibria rule out are the ones in which the lowest-value bidder wins bundle  $A$  and the ones in which the highest-value bidder wins neither bundle.<sup>9</sup>

EOS and Varian introduce a refinement of the set of Nash equilibrium, called “symmetric” or “locally envy-free” (henceforth referred to as SNE). Formally, bids  $b_1 > b_2 > b_3$  form a SNE if and only if

$$c_A(v_1 - b_2) \geq c_B(v_1 - b_3), \quad (4)$$

$$c_B(v_2 - b_3) \geq \max\{c_A(v_2 - b_2), 0\}, \quad (5)$$

$$b_3 \geq v_3. \quad (6)$$

Notice that these inequalities follow from requiring local envy-freeness: *a bidder should never wish to exchange her winning bundle at the per-unit price she pays with another, adjacent,*

---

<sup>8</sup>Ties are broken randomly.

<sup>9</sup>If the lowest-value bidder wins bundle  $A$ , then he will need to pay the second-highest bid, which by the last condition must be no less than the lowest value, so he will lose money. Also, if the highest-value bidder wins no position, then the second-highest bid cannot be smaller than the highest value. This means that the winner of position  $A$  must be paying more than his or her value, which obviously cannot happen in equilibrium.



bundle (including possibly a null bundle) at the per-unit price the winner of that bundle pays.<sup>10</sup> It can be readily shown that the SNE requirements, i.e. inequalities (4) to (6), imply the NE requirements, i.e. inequalities (1) to (3), but not the other way around, while the allocation in any SNE is efficient, i.e.,  $v_1 > v_2 > v_3$ . However, the SNE does not pin down the equilibrium bids, or equivalently the bidders' payments. In fact, there is a continuum of SNE bid profiles. Rearranging the above inequalities yields testable bounds on the equilibrium bids, respectively, for the mid-value bidder and lowest-value bidder:

$$\left(1 - \frac{c_B}{c_A}\right) v_2 + \left(\frac{c_B}{c_A}\right) b_3 \leq b_2 \leq \left(1 - \frac{c_B}{c_A}\right) v_1 + \left(\frac{c_B}{c_A}\right) b_3,$$

$$v_3 \leq b_3 \leq v_2 \text{ and } b_3 < b_2 < b_1.$$

Among the set of SNEs, EOS suggest the one with the lowest bid profile (henceforth, “lowest equilibrium”) as most plausible.<sup>11</sup> Not only is this equilibrium most preferred by all bidders among all SNEs, but it also implements the Vickrey-Clarke-Grove (VCG) prices. Further, EOS show that this particular SNE emerges as a unique (perfect Bayesian) equilibrium outcome in an ascending version of GSP auction. In our setup, the lowest SNE corresponds to the following bid profile:

$$b_2 = \left(1 - \frac{c_B}{c_A}\right) v_2 + \left(\frac{c_B}{c_A}\right) v_3, b_3 = v_3, \text{ and } b_1 > b_2.$$

Finally, one can invoke a dominance argument to further refine SNE. Note it is weakly dominated for a bidder to bid above his own value since doing so will either make no payoff difference or entail a loss relative to bidding one's own value. Throughout, we focus on SNE in undominated strategies, labeled SNEU in short, which is characterized as follows:

$$\left(1 - \frac{c_B}{c_A}\right) v_2 + \left(\frac{c_B}{c_A}\right) v_3 \leq b_2 \leq \min \left\{ \left(1 - \frac{c_B}{c_A}\right) v_1 + \left(\frac{c_B}{c_A}\right) v_3, v_2 \right\}, \quad (7)$$

$$v_3 = b_3 < b_2 < b_1 \leq v_1. \quad (8)$$

Notice the lowest SNEU coincides with the lowest SNE, the VCG bids. (Hence, the adoption of SNEU leaves intact possible selection of the VCG outcome.)

Next, notice that the nature of competition as well as equilibrium prediction changes as the ratio  $\frac{c_B}{c_A}$  varies. Observe from (7) that the lower bound [resp. upper bound] of the SNEU for the mid-value bidder is expressed as a convex combinations of the lowest-value bidder's value  $v_3$  and her own value  $v_2$  [resp. highest value  $v_1$ ], with  $\frac{c_B}{c_A}$  serving as the weight for

---

<sup>10</sup>To explain the inequalities, note that by (4), bidder 1 who wins  $A$  at the per-unit price  $b_1$  should not envy bidder 2 who wins  $B$  at the per-unit price  $b_3$ . By (5), bidder 2 in turn should not envy either bidder 1 or bidder 3 (who wins no position at zero price). By (6), bidder 3 should not envy bidder 2.

<sup>11</sup>As is typical with assignment games (whose core allocations are identical to the SNE allocations of the current game), the equilibrium bids form a lattice, with the lowest and the highest equilibria.

the former. As  $\frac{c_B}{c_A}$  falls to zero, the SNEU bid  $b_2$  converges to  $v_2$ . In this extreme case, the competition is essentially about winning bundle  $A$ , the only one worth having, so the GSP resembles a standard second-price auction. On the other hand, as the ratio  $\frac{c_B}{c_A}$  rises to one, the equilibrium bid  $b_2$  collapses to  $v_3$ . In this other extreme case, the two bundles become indistinguishable, so the bidders compete to win *either* bundle at a cheaper price per unit. In other words, the GSP auction becomes like a Bertrand competition wherein the mid-value bidder bids close to the lowest value  $v_3$ , thereby leaving no room for the highest-value bidder to undercut him. These comparative statics can be used as an additional testable restriction on the theory, which we exploit in our experimental design.

More specifically, in the experimental design we consider two different sizes of bundle  $A$  while keeping the size of bundle  $B$  constant:  $(c_A, c_B) = (20, 10)$  and  $(11, 10)$ . We shall call these two cases as “20-unit” and “11-unit” games, respectively. Substituting these parameter values into the SNEU condition, the SNEU bounds for mid-value bid  $b_2$  in the 20-unit game (i.e., with  $c_A = 20$ ) are given by

$$\frac{1}{2}v_2 + \frac{1}{2}v_3 \leq b_2 \leq \min \left\{ v_2, \frac{1}{2}v_1 + \frac{1}{2}v_3 \right\}. \quad (9)$$

On the other hand, in the 11-unit game, the equilibrium bounds for mid-value bid  $b_2$  are

$$\frac{1}{11}v_2 + \frac{10}{11}v_3 \leq b_2 \leq \min \left\{ v_2, \frac{1}{11}v_1 + \frac{10}{11}v_3 \right\}. \quad (10)$$

These bounds help us confront the theory to the data. More importantly, observing how human subjects respond to these two different environments helps us learn about the strategic sophistication of the human subjects; namely, whether they understand the nature of the game that they are asked to play. For instance, one naive rule of thumb a bidder may employ is to simply bid his own value. Since the SNEU bounds sometime include value bidding as an equilibrium strategy, it is in principle difficult to identify whether subjects simply follow value-bidding heuristics or play strategically. Yet, the extent to which value-bidding is consistent with SNEU for the mid-value bidder differs across the two scenarios; value-bidding is more likely to fail the bounds condition when  $c_A = 11$  than when  $c_A = 20$ .<sup>12</sup> How often (mid-value) subjects bid close to their values, and whether they do so when value-bidding is consistent with SNEU condition (8) will thus help us tell whether the subjects “get” the strategic environment they face.

---

<sup>12</sup>When bidders’ values are independently and uniformly drawn from  $[0, 100]$  as in the experiment, value bidding fails to meet the equilibrium restriction with 50% probability when  $c_A = 20$ , whereas it fails with about 91% probability when  $c_A = 11$ .

## 2.2 Research questions

The leading theory of the GSP auctions described above raises two questions. The first is the apparent gap between the theoretical abstraction and reality: the model assumes one-shot interaction among bidders who have complete information about opponents’ preferences, whereas real GSP auctions are dynamic and the bidders are unlikely to have complete information about their opponents. While the dynamic interaction provides bidders with opportunities to learn and adjust their behavior, whether that process will lead them to behave “as if” they play the complete information static game assumed by the theory is far from clear.<sup>13</sup> We thus pose the following question:

**Question 1** Does the static complete information game approximate adequately the dynamic game of sponsored search auctions?

We seek to address this question via a lab experiment. Specifically, we shall design an experiment that replicates the static complete information game and an experiment that captures salient features of the realistic dynamic Bayesian game, and test the hypothesis that the behavior of the subjects in the former setting would resemble that exhibited by the subjects in the latter setting through repeated interaction. In other words, we are primarily concerned with the *empirical equivalence* between the subject behaviors in the DI and SC setups.

The next question deals with multiplicity of equilibria. As mentioned earlier, Nash equilibria have a very weak predictive power. Symmetric Nash or locally envy-free refinement pins down the allocation but leaves a plethora of bid profiles consistent with the requirement. Among them, the lowest SNE—the VCG bids—is often suggested as most plausible. EOS show that if bidders play an ascending-auction version of GSP, then the unique perfect Bayesian equilibrium implements the VCG outcome. Cary et al. (2014) consider a dynamic game in which bidders play GSP repeatedly over time and show that the VCG outcome can be attained if bidders behave adaptively optimally—that is, best respond to opponents’ bids in the previous period—in a particular way, employing so-called “balanced bidding.” The balanced bidding strategy requires that in any given period, (i) each bidder targets a position that would maximize his payoff if the others submitted the same bids as in the previous period and (ii) each bidder adjusts his bid to such a level that if he were to pay that bid for moving one position up, then he becomes indifferent to winning the target position.<sup>14</sup>

---

<sup>13</sup>There is a theoretical argument suggesting that Bayesian learning in repeated play would result in players behaving according to a full-information Nash equilibrium (see Jordan (1991), for instance), but the theory does not suggest how the players would “select” a Nash equilibrium, in case there are multiple Nash equilibria.

<sup>14</sup>To be formal,  $k$  and  $p$  denote the target position for bidder  $i$  with value  $v_i$  and the price he must pay to obtain that position (given others’ bids from the previous period), respectively. Then, the balanced bid for

The arguments supporting SNE and further refinements are insightful and have some appeal.<sup>15</sup> Yet, their behavioral foundations are not compelling. For instance, it is unclear why ascending bid auctions describe the actual GSP auctions (which are non-ascending), and why bidders would adopt “balanced bidding” strategies. It is even less clear whether human subjects would behave according to suggested heuristics. The lack of a clear justification for the equilibrium selection leads to the following question:

**Question 2** Are SNEUs, particularly the lowest SNEU, empirically valid?

In order to assess the empirical validity of SNEUs, we start by testing the null hypothesis of the lower bound of SNEU, that is, the VCG outcome. This will tell us a statistical validity of the VCG outcome. We then move on to test the null hypothesis of the upper bound of SNEU. Combining the statistical tests of the lower bound and upper bound of SNEU together, we assess whether the SNEU bounds are consistent with empirically observed behavior.

## 3 Design and procedures

### 3.1 Experimental design

We chose our experimental design to answer the two research questions in the previous section. For Question 1, we use three distinct game treatments: static complete-information (SC), static incomplete-information (SI), and dynamic incomplete-information (DI) treatments. The SC treatment is used to replicate the environment assumed by the leading theory of GSP auctions. The DI treatment is motivated to mimic the environment that captures the salient features of the actual practice. The SI treatment serves a bridge between

---

bidder  $i$ , denoted as  $b'$ , satisfies

$$c_k(v_i - p) = c_{k-1}(v_i - b'),$$

where  $k - 1$  denotes the position that is one position above  $k$ . (If  $k$  is the highest position, then  $b'$  is set at  $\frac{v_i + p}{2}$ .) Since the same indifference condition is used to define the lower bound of SNE, and since the lower bound coincides with VCG, it is not surprising that the resulting dynamics converge to the VCG outcome.

<sup>15</sup>Edelman and Schwarz (2010) take a slightly different approach. By invoking revenue equivalence, they show that the long-term average of the equilibrium revenue can never exceed that obtainable under the ascending version of GSP with an optimal reserve price (which coincides with the VCG revenue with the optimal reserve price). In fact, one can extend this argument to show that, without a reserve price, the revenue can never exceed that of the VCG level. It is not clear, however, that this theory necessarily supports the VCG prediction or even the SNE prediction. With the dynamic feedback, the game is rife with signaling and bid manipulation, so the equilibrium allocation is unlikely to be efficient. Given an inefficient allocation, the long term average of the revenue will be strictly below the VCG revenue, departing from the predicted range. We present this result in Appendix I.

the two games and is used to distinguish the differential impacts of information and timing on subjects' behavior. Comparing the outcomes among these three games will help answer the first question.

In order to understand subjects' bidding behavior and thus to answer Question 2, we investigate how subjects react to two strategic environments differentiated by the click decay rate. As shown in Section 2.1, the variation of click-decay rate changes the nature of competition in the GSP auction and the predictions of SNE. We exploit this comparative static as an instrument for testing theory, consider two treatments based on the 20-unit and 11-unit games.

In sum, we have in total six distinct treatments: two SC games with  $c_A = 11$  and 20 (called SC-11 and SC-20, respectively); two SI games with  $c_A = 11$  and 20 (called SI-11 and SI-20); two DI games with  $c_A = 11$  and 20 (called DI-11 and DI-20).

Each game in the experiment has three participants competing for bundles  $A$  and  $B$ , consisting of  $c_A = 20$  or 11 units and  $c_B = 10$  units of a hypothetical commodity, respectively. Obviously, the bundles correspond to ad positions and the units correspond to clicks in the sponsored search auctions, but the subjects were presented with a neutral context that avoids a possible framing effect.

In the beginning of each game, each participant is assigned a value per unit, drawn at random from the set of integers between 1 and 100. In the DI game comprising 15 periods of GSP auctions, the assigned values per unit remain constant throughout the game. Each participant is then asked to submit a single bid per unit. A per-unit bid is allowed to be any integer number between 0 and 999. Submitted bids are ranked from highest to lowest with ties being broken randomly. A participant who submits the highest per-unit bid wins bundle  $A$  and pays the second-highest bid per unit. A participant who submits the second-highest per-unit bid wins bundle  $B$  and pays the lowest bid per unit. A participant who submits the lowest bid wins nothing and pays nothing. The random generation of individual per-unit values enables us to test the theory in a wide range of value profiles.

The information structure about participants' per-unit values differs across the three games. In the SC games, a per-unit value assigned to each participant is made publicly known to *all* three participants. In the SI games, the per-unit value for each participant is his or her private information and thus is not observed by any other participants. Finally, in the DI games, each participant's per-unit value assigned in the beginning of the game is his or her private information, as in the SI games, and remains constant throughout all 15 periods.

The dynamics of GSP auctions in the DI treatments proceed as follows. For given realized per-unit values, the first decision period starts with participants playing the GSP auction game. At the end of the first period, they are informed about the results of the first period, which include all participants' per-unit bids, the allocation of two bundles, and his or her

own price and earnings in the period.<sup>16</sup> Note that per-unit *values* of other participants in the group are not revealed. Each period, from the second period on, lasts for 20 seconds. Within 20 seconds, each participant is asked to revise his or her decision from the previous one. If the participant does not submit a new bid, his or her per-unit bid in the previous period is taken as the bid in the current period. If the participant submits a new bid, that bid replaces the previous bid. After the elapse of 20 seconds, the GSP auction is run with (possibly) revised bids, and the subjects are informed of the results of that period as in the first period.<sup>17</sup> This process is repeated until all 15 decision periods are completed. Throughout the 15 periods, the value profile remains fixed and private to the subjects; they of course may learn about their opponents’ values through feedback. Note that the “opt-in” feature of our design—i.e., a bidder’s inaction in a specified time (20 seconds) causes his or her previous bid to stand—accords well with actual practice of sponsored-search auctions.<sup>18</sup>

**Remark 1.** *Recall that we are taking the theoretical prediction of SC game as the null hypothesis for the DI and SI games, since our central purpose is to test the validity of using the SC game as a modeling shortcut abstracting the features present in these games. One can nevertheless ask about the equilibrium predictions of these games themselves. On DI, Cary et al (2014) predict the VCG outcome invoking the balancedness assumption discussed earlier. On the SI game, Gomes and Sweeney (2014) provide a general analysis of Bayesian Nash equilibrium of this game and show that if unit differences between bundles are sufficiently small (as in our  $(c_A, c_B) = (11, 10)$  case), there exists no symmetric equilibrium. But we show in Appendix I that when  $(c_A, c_B) = (20, 10)$ , there exists an efficient equilibrium with the symmetric bidding strategy that is nonlinear and increasing in value. Given our null hypothesis, however, we will not focus on the empirical validity of that theory, although we shall address it briefly.*

---

<sup>16</sup>Informing each participant in the DI treatment about all participants’ bids after each period is meant to help subjects learn how much to bid and pay in order to win a desirable position. This feature is also not without realism; the Google’s AdWords has some tools such as *bid simulator* and *traffic estimator* that enable the similar type of learning for the advertisers.

<sup>17</sup>The 20 seconds limit for bid revision is reflective of the real world practice of the sponsored search auctions, which are held quite frequently — as often in principle as a user types in a query. Hence, bidders not having enough time to respond to the latest change is not particularly problematic in terms of capturing the nature of dynamic interactions advertisers face in sponsored search auctions. On the other hand, there is no evidence that our subjects found 20 seconds too short. Our data on their response time shows that those who revised their bids did so relatively early, in median time of 10 (8) seconds and mean time of 10.3 (8.9) seconds in the DI-11 (DI-20) treatment, well within the 20 seconds limit. About 26 percent of bidders did not revise their bids, but we suspect that many of them chose not to; 75% (87%) of them in the DI-11 (DI-20) treatment were already behaving (adaptively) optimally.

<sup>18</sup>Sponsored search auctions take place every time a new query is submitted and queries usually arrive more quickly than advertisers can change their bids.

## 3.2 Experimental procedures

The experiment was run at the Experimental Laboratory of the Centre for Economic Learning and Social Evolution (ELSE) at University College London (UCL) between November 2009 and March 2010. The subjects in this experiment were recruited from an ELSE pool of UCL undergraduate students across all disciplines. Each subject participated in only one of the experimental sessions. After subjects read the instructions, the instructions were read aloud by an experimental administrator. Each experimental session lasted between one and a half hours and two and a half hours. The experiment was computerized and conducted using the experimental software z-Tree developed by Fischbacher (2007). Sample instructions are reported in Appendix II.

We conducted six experimental sessions, one session for each game treatment. There were 21 subjects in each session. Each subject received £10 as an initial balance in the beginning of the session, including a £5 show-up fee. Any gains or losses incurred during the session were added into or subtracted from this balance, and the resulting earnings were paid in private at the end of the session.<sup>19</sup>

The static game sessions had 40 rounds, while the dynamic game sessions had 15 rounds. Each **round** consisted of a single auction in the SC and SI treatment, and of 15 auctions held across 15 **periods** for the DI treatment. The following process was repeated in all 40 rounds in static game treatments and in all 15 rounds in dynamic game treatments. Each round started with the computer randomly forming three-person groups and selecting per-unit values for participants. The groups formed in each round depended solely upon chance and were independent of the groups formed in any of the other rounds. Once each group played a one-shot GSP auction in the SC and SI treatments or 15 periods of GSP auctions in the DI treatments, each round ended with subjects observing the results of that round and their earnings. The results of each round in the static game treatments include per-unit values, per-unit bids, per-unit prices, allocation of bundles, and earnings in that round. In the DI treatments, a participant's earnings in each round were determined by the sum of his or her earnings in the three decision periods randomly selected out of 15 periods. At the end of each round, the computer informed each participant of the per-unit values assigned, choices and earnings made by all participants in the group in each selected decision period.

The total earnings for each participant, which were the sum of earnings in 40 rounds in the SC and SI treatments and in 15 rounds in the DI treatments, were calculated in terms of tokens (experimental currency) and then exchanged into British pounds, where 100 tokens were worth £0.10. A £10 initial balance and subsequent earnings, which averaged about

---

<sup>19</sup>We chose the initial balance large enough to make sure that there is little chance of subjects going bankrupt during the experiment and to avoid the problem of limited liability on losses (Kagel and Levin (1991), for example). In fact, no subjects in our experiment experienced a bankruptcy during the experiment.

£19, were paid in private at the end of the session.

We have in total 126 subjects participating in experimental sessions. In order to control for potential learning effects in early rounds, we used samples after 5 rounds in static game treatments and those after 3 rounds in dynamic game treatments for data analysis. Since there were 7 groups in each round of the experimental sessions, this resulted in 245 ( $= 7 \times 35$ ) observations in each static game treatment and 84 ( $= 7 \times 12$ ) observations in each dynamic game treatment.

## 4 Experimental results

In this section we present our experimental results. We first provide an overview of the experiment data by presenting the efficiency and revenue of the GSP auction under each treatment, comparing them across treatments, and against the theoretical bounds of SNEU. We next investigate the subjects' bidding behavior in order to gain a deeper understanding of the relation between experimental data and theory. We then draw implications of our findings for the two main research questions.

### 4.1 Efficiency and revenue

We first compare the allocative efficiencies across all the treatments. Recall that an efficient allocation is not implied by Nash equilibrium but that it is implied by SNE and SNEU. Table 1 presents the frequency of efficient allocation, along with an inefficient allocation in which bundle  $A$  is assigned to the mid-value bidder and bundle  $B$  to the highest-value bidder. In the dynamic game treatments, we report the frequencies of efficient allocation based on the samples of all 15 periods and of the last 5 periods. Across all treatments, the majority of auctions produced efficient allocations. In more than 80% in the data, the two bundles were assigned to the two highest values, which is consistent with NE behavior.

In order to evaluate the magnitude of inefficiencies, we also compute the *efficiency ratio*, which measures the surplus improvement over random allocation as percentage of the maximal surplus improvement (attained by full efficiency) over random assignment.<sup>20</sup> This ratio equals one if the bundles are allocated efficiently and less than one if they are allocated inefficiently; in fact, the ratio can be negative if the surplus achieved is less than that associated

---

<sup>20</sup>Note that the measure normalizes the realized surplus by taking the two surpluses from efficient and random assignments as benchmarks. This double-normalization renders the measure more robust against the rescaling of value support than a more common measure, such as the percentage of the first-best surplus realized. This latter measure would vary with the rescaling the lower bound of support; for instance, if the support were  $[\underline{v}, \underline{v}+100]$ , then as  $\underline{v}$  increases, the measure will record a very high percentage approaching 100% for all allocations, including the random assignment! The robustness of our measure makes its interpretation consistent.



with random allocation of the bundles. The last column reports the number of observations in each treatment.

- Table 1 here -

When the units of bundle  $A$  are 11, we observe higher frequencies of efficient allocation in the SI treatment than in the other treatments. In the case of 20 units of bundle  $A$ , efficient allocation is more frequently observed in the SI and DI treatments than in the SC treatment. In each game treatment, allocations are more frequently efficient in the 20-unit case than in the 11-unit case, especially in the SC and DI treatments. Interestingly, the inefficient allocation that assigns bundle  $A$  to a mid-value bidder and bundle  $B$  to a highest-value bidder occurs more frequently in 11-unit treatments than in 20-unit treatments for both SC and DI treatments. A reasonable interpretation is that the large ratio associated with the 11-unit treatment ( $\frac{c_B}{c_A} = \frac{10}{11} \approx 0.91$ ) caused two high-value bidders to undercut each other for a better deal even at the risk of not winning  $A$ . This is in line with the Bertrand-like feature of the 11-unit treatment which makes the top two bidders compete essentially for “either” bundle by undercutting each other to a level close to  $v_3$ . This effect should never create any inefficiencies in theory (which assumes that players coordinate perfectly) but in reality may manifest itself as a greater incidence of inefficiencies of the sort observed in our experiment with  $c_A = 11$  than with  $c_A = 20$ .

In fact, the differences between 11-unit and 20-unit treatments are less salient in terms of the efficiency ratio, which indicates that the payoff consequence of losing  $A$  was not so significant in the treatments with 11 units compared to those with 20 units. The efficiency ratios in our data are relatively high; the GSP auctions achieve 76-93% of the highest possible efficiency gains over the random allocation in all treatments.

Next, we examine the revenue performances of the GSP auctions in our experiment. As discussed in Section 2, the equilibrium prediction of the leading theory provides bounds on the attainable revenue. We compare the bounds of SNEU with observed revenues by computing the percentage differences of observed revenues from the theoretical bounds.<sup>21,22</sup> We also compute the percentage deviations of observed payments for each bundle,  $A$  and  $B$ , from the respective prediction from the lower and upper bounds of SNEU. Table 2 reports the summary statistics of the percentage deviations of observed revenues/payments from the lower and upper bounds of SNEU. The standard error for mean and the bootstrap standard error for median are reported. We also report the significance results of the one-sided  $t$ -test

---

<sup>21</sup>The percentage difference of observed outcome from theoretical outcome is defined by (observed outcome – theoretical prediction)/theoretical prediction.

<sup>22</sup>We have also considered two alternative measures:  $\frac{\text{observed revenue} - \text{VCG revenue}}{\text{highest SNE revenue} - \text{VCG revenue}}$  or  $\frac{\text{observed revenue} - \text{VCG revenue}}{\text{first-best surplus} - \text{VCG revenue}}$ . The results based on these measures are similar to those with the current measure.

of median and mean, respectively, against the alternative hypothesis that the median or mean percentage deviation is strictly above zero.

- Table 2 here -

There are several notable patterns in the results of revenue (bundles  $A$  &  $B$ ) and payments (bundle  $A$  and bundle  $B$ , respectively). First, in each treatment the revenues observed in the data are significantly higher than those predicted by the lower bound of SNEU (and of SNE). The median percentage deviation from the lowest SNEU is 5% and 3% in the SC-11 and SC-20 treatments, 43% and 34% in the SI-11 and SI-20 treatments, and 9% and 18% in the DI-11 and DI-20 treatments, respectively. In all treatments except for the SC-20 treatment, in which the significance result holds at the 10% level, the median percentage deviations are strictly above zero at the 1% significance level. When we use the average percentage deviations, the discrepancy becomes larger in all treatments: 16-28% in the SC treatments, 46-106% in the SI treatments, and 39-70% in the DI treatments, all of which are statistically significant with usual significance levels. This implies the empirical distributions of percentage deviations are positively skewed. Second, holding the size of  $A$  fixed, the deviation from the lower bound of SNEU is smallest in the SC and highest in the SI, and in between in the DI treatment. Third, the observed median payments for bundle  $A$  are significantly higher than those predicted by the lower bound of SNEU, whereas the same does not hold for bundle  $B$  with the exception of the DI-20 treatment. This suggests that the upward deviation of observed revenues is mainly driven by the second-highest bidders' overbidding, which causes the highest bidders to overpay (relative to the VCG level). In light of the high frequencies of efficient allocation in our data, we may conclude that the mid-value bidders tend to overbid compared to the lower bound of SNEU, which we will investigate in more details later. Last, the median revenue observed in the data does not exceed the upper bound of SNEU in most treatments, although the average percentage deviations are significantly above zero significantly in all but SC-20 treatments. The observed median revenue of the SI game still consistently and significantly exceeds the upper bound of SNEU, but this pattern is not observed in other treatments, except for the DI-20 treatment in which the median percentage deviation is 4%. We summarize the findings so far as follows:

**Result 1 (Efficiency and revenue)** *(i) The GSP auction achieves high efficiency both in frequencies and magnitude; they attain 76-93% of the highest possible surplus improvement over what can be achieved from random allocation of the bundles. (ii) The observed median revenues are significantly higher than those predicted by the lower bound of SNEU in all treatments. With regard to the upper bounds of SNEU, the observed median revenues do not differ significantly, except for the SI treatments.*

## 4.2 Bidding behavior

We now investigate subjects' bidding behavior. The upward deviation of observed revenue implies that at least some, if not all, bidders overbid compared to the lower bound of SNEU bids. It is useful to understand which bidders (in terms of the per-click values) overbid, by how much, and under what treatments. We also investigate if there exists any pattern of bid convergence over time in the DI treatments and, if so, whether the dynamic behavior tends to converge to the outcome of the SC treatments.

### 4.2.1 Overview

We begin by running simple regressions of bids on values. We median-regressed the bid of a subject with a given value ranking (i.e., highest-value, mid-value, and lowest-value) against the values of all subjects in the group.<sup>23</sup>

Table 3 reports the results of the median regression. In each regression, the independent variables are highest value ( $v_1$ ), mid value ( $v_2$ ), and lowest value ( $v_3$ ). For the DI treatments, we report the regression results in the first period and in the last 5 periods. (The last five periods of DI games are relevant for examining the long-term evolution of the subjects' behavior; and, as will be seen, the first period game of DI is relevant for studying whether subjects act differently from the way they do in a static game, which can be learned from the SI treatment.) Recall our null hypothesis that subjects would behave under DI or SI according to the equilibrium prediction of SC. Regression of their behavior on variables that the CS equilibrium prediction refers to (i.e., his value as well as those of his opponents) thus allows us to examine that hypothesis. We control for the heterogeneity of bidders by including individual subject dummies in each regression. The standard errors reported are also clustered by individual subject.

- Table 3 here -

The simple regression results reveal marked differences in the bidding behavior across treatments. In the SC treatments, the highest-value and mid-value bidders respond to their opponents' values as well as their own, as measured by the relevant coefficients, in a way consistent with the SNEU bounds. In particular, the manner in which the dependence varies with the size of  $A$  is in line with the theory. The Bertrand game feature implicit in the 11-unit game implies that the mid-value bidder will shade his/her bid more towards the lowest value in that treatment than in the 20-unit game treatment, and this is what we find. The coefficient on  $v_3$  is 0.846 in the former treatment, while it is 0.096 in the latter treatment.

---

<sup>23</sup>We employ the median-regression method instead of the least squares method since the former is more robust to outliers, which are observed in our bidding data.

The differential dependence of the median mid-value bid on  $v_3$  between the two treatments suggests that the subjects behaved strategically and responded to the varying size of  $A$  in a manner qualitatively consistent with the SNEU predictions based on (7). Compared to the lowest SNEU, the weight on  $v_3$  is somewhat close to that in the lower bound (0.91, which can be seen in the left side of (10)) in the 11-unit treatment, whereas it is significantly smaller than that in the lower bound (0.50, which can be seen in the left side of (9)) in the 20-unit treatment. The null hypothesis of mid-value bidders following the lower bound of SNEU is, based on the Wald test, rejected at the usual significance levels in the 20-unit treatment, whereas it is rejected only at the 10% significance level in the 11-unit treatment. For lowest-value bidders, the regression results suggest that they tend to bid their own values, which is the prediction of SNEU. The null hypothesis that the lowest-value bidders follow the SNEU cannot be rejected at the usual significance levels.

For the SI treatments, subjects' bidding behavior depends solely on their own values but not on other bidders' values. This stands in stark contrast to the behavior in the SC treatments and highlights the impacts of information structure on bidding behavior. In particular, the coefficients on the own values are close to one, regardless of the size of  $c_A$ , which suggest that bids in the SI-20 treatment track the values of the bidders closely in one-to-one fashion.<sup>24</sup>

The first-period behavior of subjects in the DI treatments is similar to that in the SI treatments. This result is reassuring since subjects have not yet had an opportunity to learn their opponents' values; they are informed solely of their own values just as in the SI, and are thus expected to behave similarly.<sup>25</sup> When we turn to the last five periods, the bidding behavior of the subjects depends on not only their own values but also their opponents' *unobserved* values. These two pieces of evidence thus suggest the presence of learning and feedback in DI, a point that will be investigated in a greater detail later. For the mid-value bidders in the last 5 periods, the behavior in the 20-unit treatment appears similar to that in the SC treatment, while no such similarity is apparent in the 11-unit treatment. For the lowest-value bidders, their bids appear to depend primarily on their own values but to a lesser degree also on the mid-values, especially in the 20-unit treatment. This is somewhat different from the behavior observed in both SC and SI treatments.

To further inspect subjects' behavior, we draw scatter plots between bidders' values and their own bids for mid-value and lowest-value bidders in the SC and DI treatments and for all bidders in the SI treatments. They are presented in Figure 1.

---

<sup>24</sup>Note that mid-value and lowest-value bidders tend to bid above their values, as indicated by (significantly) positive constant terms in the regression results in Table 3.

<sup>25</sup>This finding also implies that the subjects do not display any extraordinary form of experimentation—deviation from a myopically optimal behavior in an attempt to increase the amount of learning in the subsequent periods.

- Figure 1 here -

The scatter plots reveal several new insights on the subjects' bidding behavior. First, in the SC treatments (Figure 1-1 and 1-4) the majority of lowest-value bidders tend to bid their own values in both 11-unit and 20-unit treatments (72% in the SC-11 treatment and 53% in the SC-20 treatment). This is in line with the SNE prediction for lowest-value bidders. On the other hand, there is a sizable minority of these bidders who employ different types of bidding strategies: bidding close to zero (14% in the SC-11 treatment and 27% in the SC-20 treatment) and bidding above own values (13% in the SC-11 treatment and 18% in the SC-20 treatment).<sup>26</sup> When the opponents bid more than his value  $v_3$ , bidding zero and bidding one's value (as well as any amount in between) constitute best responses for the lowest-value bidder, so these patterns of bidding are optimal for him.<sup>27</sup> Nonetheless, the differences matter for the mid-value bidder's payoff and the revenue obtained by the seller.

In the DI treatments (Figure 1-3 and 1-6) with last 5 periods, we observe somewhat similar patterns for the lowest-value bidders: the majority of subjects tend to bid their own values (58% in the DI-11 treatment and 44% in the DI-20 treatment), while bidding above own values is observed frequently (32% in the DI-11 treatment and 48% in the DI-20 treatment). However, unlike the SC treatment, bidding close to zero is rarely observed (3% in the DI-11 treatment and 4% in the DI-20 treatment). The paucity of near-zero bids seems attributable to both incomplete information and the opt-in feature of the dynamic auctions: Bidders initially do not know if their values are the lowest, and the opt-in feature keeps them bidding sincerely even after they realize their values are the lowest. Second, even though some mid-value bidders bid close to their values, many of them also bid below their own values in both SC and DI treatments especially in the 11-unit game. The simple inspection of scatter plots suggests the stark difference in the behavior of mid-value bidders between the 11-unit and 20-unit cases of both SC and DI treatments. This difference conforms to the predicted differences in the responses by mid-value bidders to different click-decay factors; it also supports the view that the subjects are sophisticated enough to recognize the strategic implications of the click-decay factors. Finally, as the regression results (Table 4) indicate, the bidding behavior in the SI treatments differs significantly from those under the SC and DI

---

<sup>26</sup>When we count frequencies of the subjects bidding zero and their values, we apply a 2 token margin of error. For instance, we count a bidder's bid as value bidding if it is within a range of his own value plus/minus 2 tokens. We count bidding above own values if a bid is above by more than 2 tokens from own value.

<sup>27</sup>To understand better on the motive by lowest values subjects for bidding close to zero, we ran simple probit regressions on the difference between mid value and lowest value. We omit the details in the interest of space, but we find some evidence that lowest-value bidders tend to bid close to zero when this difference is large, particularly in the 20-unit treatment. This reinforces the intuition that when there is little chance of winning in terms of realized values, lowest-value bidders become pessimistic and check out of auction (by bidding close to zero).

treatments: the subjects bid close to their values.<sup>28</sup> We do not find this result unreasonable, however, since, absent the information about other bidders' values, value-bidding becomes a compelling rule of thumb.<sup>29</sup>

#### 4.2.2 Above-VCG bidding

We now compare observed bids with the theoretical predictions of SNEU for mid-value bidders. Table 4 presents the percentage differences of observed bids from the bound of SNEU (condition (10)), as well as from value bidding.<sup>30</sup> We use the data of last 5 periods for the DI treatments. The robust standard error for mean and the bootstrap standard error for median are reported, both of which are clustered by individual subject. The significance results are also reported for the one-sided  $t$ -test of median and mean, respectively, against the alternative hypothesis that the median or mean percentage deviation is strictly above zero.

- Table 4 here -

Overall, the results confirm that the mid-value bidders overbid compared to the lower bound of SNEU. First of all, the subjects bid within the upper bound of SNEU, but exhibit significant overbidding relative to its lower bound, VCG, in all treatments. The median percentage differences are 19-22% in the SC treatments, 42-81% in the SI treatments, and 22-25% in the DI treatments. The mean percentage differences are even larger as the distribution of percentage differences is positively skewed.<sup>31</sup>

---

<sup>28</sup>More precisely, bidders slightly overbid above their values, as

<sup>29</sup>Even though value-bidding is not an equilibrium of SI, the second-price feature of GSP makes it difficult for subjects to figure out an optimal deviation from value bidding (when others are bidding their values), and the gains from deviation tend to be modest. It is therefore not surprising that the Bayesian Nash equilibrium of the SI does not explain the bidders' behavior: The subjects' behavior in the SI-20 treatment (Figure 1-5)—the only case in which the Bayesian Nash equilibrium exists—exhibits a substantial overbidding compared to the Bayesian Nash equilibrium strategy (drawn in the solid line). More specifically, the Bayesian Nash equilibrium is given by  $b(v) = 100 \ln(100 + v) - 100 \ln(100)$  (see more details in Appendix I). The median regression using all the sample yields the estimated bidding function as  $\hat{b}(v) = 143.44 \times \ln(100 + v) - 664.37$ , where the bootstrap clustered standard errors are reported in parentheses. The hypothesis that subjects play the Bayesian Nash equilibrium is rejected at the usual significance levels ( $p$ -value = 0.000).

<sup>30</sup>We replicated Table 4 with the subsample of the data in which efficient allocation was achieved. The results remain similar, which we do not report in the interest of brevity.

<sup>31</sup>It is possible that the overbidding behavior mid-value bidders exhibited in our experiment may have been influenced by the fact the he/she is competing against a bidder who would not win any position, which is a necessity when only two positions are on sale. While it is difficult to predict the behavior in an auction with more than two positions, the pressure from the low value bidders (who would not win any positions) may still cascade into a similar overbidding behavior for winning bidders.

Second, when we take the upper bound of SNEU as the benchmark, overbidding becomes much less pronounced. The median differences from the SNEU upper bound are virtually none for SC-20 and DI-20, whereas the median differences are 7% and 11% for SC-11 and DI-11, respectively. The fact that we only observe the significant patterns of overbidding relative to the upper bound of SNEU for SC-11 may reflect the subjects' reluctance to carry out the aggressive shading the SNEU strategies call for in this case—not an unreasonable response in light of the strategic uncertainty actual subjects face.

Third, the median differences of the observed mid-value bids from their values are negative for SC-11 and DI-11, suggesting that mid-value bidders tend to bid below their values in these treatments. These median differences equal zero for the SC-20, DI-20 and SI-11, suggesting that the mid-value bidders bid close to their values. Recall that, for the 20-unit case, value bidding is often consistent with SNEU strategies. Meanwhile, mid-value bidders' bids tend to be higher than the values in the SI-20 treatments. For the SC and DI treatments, the median percentage differences are significantly lower in the 11-unit treatment than in the 20-unit treatment.

In order to examine further the overbidding pattern in the data, we plot in Figure 2 the cumulative distributions of observed bids (the black solid line), lower and upper bounds of SNEU (the blue and red dotted lines, respectively) and value bidding (the purple dotted line) for mid-value bidders in each treatment. The distributions of SNEU bounds and value bidding are computed based on the value profiles in the experiment data. Each panel also contains a stochastic dominance test based on the Kolmogorov-Smirnov (K-S) test, under the null hypothesis that the distribution of observed bids first-order stochastically dominates that of the lower bound of SNEU.

- *Figure 2 here* -

The inspection of Figure 2 reveals that the cumulative distribution of observed bids appears to first-order stochastically dominate that of the lower bound of SNEU in all treatments. This observation is statistically confirmed by the stochastic dominance test: we cannot reject the null hypothesis in each treatment. This is in line with the results about the lower bound of SNEU in Table 4 and even suggests that the overbidding tendency of mid-value bidders in the data affects not only median and mean but also the entire distribution of bids. In the 11-unit treatments, the distribution of the upper bound of SNEU is not much different from that of the SNEU lower bound, suggesting that the bounds of SNEU for mid-value bidders are very tight. It is interesting to observe that the empirical distributions of bids lie somewhere between those of the lower bound of SNEU and value bidding in the SC-11, SC-20, and DI-11 treatments. This seems to suggest that overbidding is not driven by the simple heuristic of value bidding. Overall, Figure 2 reinforces the evidence of overbid-

ding relative to the lower bound of SNEU found in Table 4, even in the sense of stochastic dominance.

While the findings of overbidding relative to the lower bound of SNEU are overwhelming, they should not be regarded as a rejection of these solution concepts, but rather a rejection of the particular equilibrium selection. Since that selection lacks a compelling justification, our findings should not be regarded as an anomaly or even a failure of theory. Nevertheless, it would be useful for providing a plausible explanation of the observed bidding behavior. We shall do this in Section 5. For now, we refute two obvious hypotheses as explanations of the observed behavior.

One may wonder if the observed bidding behavior is simply the result of subjects not acting rationally or acting suboptimally against their opponents. The regression results in Table 3 partly discredit the hypothesis of irrational behavior, showing that subjects responded to the changes in strategic environments in a manner qualitatively consistent with the theoretical predictions. Table 5 provides a more direct evidence against irrational behavior. The table counts the frequency with which each subject plays a best response to his opponents' observed bids, and the frequency with which a subject bids within the SNEU range.

- Table 5 here -

The frequency of subjects playing their best responses against their opponents is quite high for mid-value and lowest-value bidders in all treatments: 60% and 88% in the SC-11; 80% and 91% in the SC-20; 70% and 86% in the DI-11; and 81% and 95% in the DI-20 treatment. By comparison, the frequency of subjects playing SNEU strategies are much lower: 49% and 72% in the SC-11; 40% and 53% in the SC-20; 40% and 58% in the DI-11; 52% and 44% in the DI-20 treatments. The difference in frequency between best response and SNEU for mid-value bidders ranges from 11% (in the SC-11) to 40% (in the SC-20). When we consider the SI treatments, the difference grows even larger. Overall, the results seem to suggest that subjects behaved optimally in the sense of Nash equilibrium, even though their bids exceeded the lower bounds of SNEU.

Another possible hypothesis explaining the above-VCG bidding is that subjects adopt value bidding as a rule of thumb. In order to investigate this possibility, we partition the samples for mid-value bidders between when value bidding is in the set of SNEU and when it is not. If some subjects were to simply employ the value-bidding heuristics, the frequency of value bidding should not differ based on whether it satisfies the requirements of SNEU. Table 6 reports the frequency of value bidding (with 2 token margin) in each of SC and DI treatments conditional on whether mid-value lies in the bounds of SNEU. In identifying whether value belongs to the set of SNEU, we round up the upper bound and down the lower bound of SNEU. The  $p$ -values of Chi-square test for the equivalence of two distributions are reported in parentheses below each table.



- Table 6 here -

In the SC treatments, the frequency of value bidding when it lies in the bounds of SNEU is significantly higher than when it does not: 0.94 vs. 0.24 in the SC-11; and 0.55 vs. 0.42 in the SC-20 treatment. The Chi-square test rejects the null hypothesis in both cases at the 5% significance level. The frequency results in the DI treatments are quite similar to those in the SC treatments: 0.71 vs. 0.24 in the DI-11; and 0.59 vs. 0.39 in the DI-20 treatments. Thus, we conclude that the adoption of value bidding heuristics is not the main source of overbidding in the SC and DI treatments.

**Result 2 (Bidder behavior)** *(i) In all treatments, mid-value bidders tend to overbid significantly relative to the lower bound of SNEU in the sense of stochastic dominance. When compared to the upper bound of SNEU, the median bids are significantly higher in the SC-11 and DI-11 treatments, whereas they are not in the SC-20 and DI-20 treatments. (ii) The majority of lowest-value bidders in the SC treatments bid close to their own values, consistent with the SNEU prediction. Nevertheless, they also tend to bid above their own values with around 25% frequency.*

#### 4.2.3 Does the behavior of SC game approximate that of DI game?

One important question of this paper is whether the SC game is a good approximation of the DI game. The hypothesis is that, even though the bidders in the DI game do not observe directly their opponents' values, dynamic interaction among bidders leads them to end up behaving "as if" they observe each other's values like in the SC game. We test this hypothesis here by comparing the behavior of SC treatments with that of DI treatments in each period.

Table 7 reports the temporal evolution of percentage differences of observed bids from the lower bound of SNEU in the DI treatments, along with the corresponding outcomes in the SC and SI treatments. Our interest is in examining how closely the DI game behavior matches that of the SC game. The normalization via the lower bound of SNEU is meant to control for the differences in the realized value profiles between the two games. We report the  $p$ -value of equality tests of medians between the SC and DI in each period, using the  $t$ -test with bootstrap clustered standard errors, and the  $p$ -value of the Kolmogorov-Smirnov test of equality of two distributions of bids between the SC and the DI in each period.

- Table 7 here -

In the 11-unit treatments, there is a clear pattern that the median percentage deviation in the DI game converges toward that in the SC game: the median percentage differences start around 50% in the first 3 periods, go down around 30-37% in the next 4 periods, and fluctuate between 15% and 28% in the remaining periods. When we conduct the test of equality of

medians between the SC-11 and the DI-11 in each of 15 periods, we cannot reject the null hypothesis from period 4 on at the usual significance levels. Since the median comparison may not fully capture the similarities/differences in the behavior of the two games, we also inspect the probability (or empirical frequency) that the percentage deviation is above zero. These probabilities in the DI-11 treatment begin low with 0.74 in the first period and reach close to that of SC-11 treatment (0.86) in later periods. We use the K-S test for the equality of two distributions between SC and DI in each period. We cannot reject the null hypothesis in most periods after period 7 at the usual significance levels. Based on these two results, we conclude that the behavior in the DI-11 treatment tends to converge toward that in the SC-11 treatment.

On the other hand, in the 20-unit treatments, the median percentage deviation of the bids of DI game (from the lower bound of SNEU) stays close to that of the SC game from the beginning to the end: the median deviations lie between 16% and 26% in the first 5 periods; between 14% and 26% in the next 5 periods; and 16% and 29% in the last 5 periods. Using the median tests, we cannot reject the null hypothesis in any periods at the usual significance levels. This finding remains the same if we use the K-S test for the equality of two distributions. Combining the findings in the 11- and 20-unit treatments, we conclude that the behavior of the SC games approximates that of the DI games relatively well.<sup>32</sup>

Recall that the 11-unit case requires a relatively more demanding coordination and a narrower bandwidth of equilibrium plays than the 20-unit counterpart. This may explain why it takes some time for the subjects to converge on the outcome that the SC game exhibits. By contrast, the coordination in the 20-unit case is much easier and the equilibrium (Nash and SNE) conditions are much more permissive, which is consistent with the convergence to the SC outcome even from the beginning.

We next investigate a possible mechanism by which this “approximation” is achieved. A reasonable process by which the bidders in the DI game act as if they play the SC game is via an adaptive process; if the bidders respond optimally to their opponents’ plays in the previous periods, they may end up responding to opponents’ true values in a way resembling the behavior in the SC game. Formally, *adaptive optimality* in period  $t$  ( $> 1$ ) is defined as individual bidders playing best response against their opponents’ bids in the previous period  $t - 1$ . The regression results of Table 3 already suggest that such a “learning” process is in operation for the DI game. Here, we look for a more direct evidence of the adaptively optimal behavior. The first four columns of Table 8 report the frequencies of adaptive optimality for all bidders and according to the value ranking.

- Table 8 here -

---

<sup>32</sup>By contrast, there is no such resemblance between SI and SC. Table 7 shows that the median percentage deviation in the SI games differs significantly from that in the DI games in both 11- and 20-unit treatments.

Notice that bidders in both the DI-11 and DI-20 treatments tend to behave more adaptively optimally in the later periods than in the earlier periods, while the level of compliance to adaptive optimality starts relatively low in the DI-11 treatment and high in the DI-20 treatment. The frequencies of adaptive optimality begin from 0.68 (periods 2-5) to 0.74 (periods 11-15) in the DI-11 treatment and from 0.80 (periods 2-5) to 0.87 (periods 11-15) in the DI-20 treatment. When we examine this more finely according to the value ranking, the highest-value bidders exhibit the lowest level of adaptive optimality and the lowest-value bidders show the highest level of adaptive optimality in each period. For instance, those frequencies in the first four periods are 0.50 (0.67) for the highest-value, 0.71 (0.81) for the mid-value, and 0.83 (0.93) for the lowest-value bidders in the DI-11 (DI-20) treatment. In summary, subjects behave more adaptively optimally over time.

Of course, the adaptively optimal behavior does not imply a convergence to a particular SNE such as the VCG outcome. All it means is that the process converges to some Nash equilibrium outcome. This latter fact is consistent with the high frequency of (contemporaneous) best response play, shown in Table 5. To investigate the dynamic behavior further, we look for the evidence that bidders followed the so-called balanced-bidding strategies. Recall that the balanced-bidding strategy is a *particular* type of adaptively optimal behavior which Cary et al. (2014) show would lead to the VCG outcome. Whether subjects use the balanced-bidding strategy in the experiment and, if not, how the behavior differs from it will provide an opportunity to enhance further our understanding of the dynamic behavior.

The second part of Table 8 reports the summary statistics of percentage differences of observed bids from the balanced-bidding strategy in each period. Since the balanced-bidding strategy is well defined only for bidders whose adaptively optimal bundle is  $B$ , we restrict our attention to this case. As evident from the table, the observed bids display a systematic deviation from the balanced-bidding strategy. The median bids for those whose adaptively optimal bundle is  $B$  are significantly higher than the “balanced bids”: they are 6-73% higher in the DI-11 treatment and 6-24% higher in the DI-20 treatment. This finding parallels the above-VCG bidding pattern found in the previous analysis. Interestingly, there is a pattern in the DI-11 treatment that the percentage difference decreases over time but remains significantly above zero even in later periods: it starts from 34% (16%) in periods 2-5 to 17% (9%) in periods 6-10 and to 8% (8%) in periods 11-15 in the DI-11 (DI-20) treatment. Combining these results with the evidence supporting adaptive optimality, we conclude that the subjects in our data tend to bid adaptively optimally but at levels higher than predicted by the balanced bidding strategy. This is another way of showing why the behavior of DI game approximates that of SC game, which exhibits overbidding relative to the lower bound of SNEU, but not the VCG outcome.

Finally, we investigate the factors that affect the adaptive process. One would expect the amount of loss caused by a suboptimal behavior in the past to be a factor in causing

bid revision. The larger the loss is in the previous period, the more likely the bidders will adjust their bids to be adaptively optimal. We define the loss in period  $t - 1$  as the difference between optimal earnings and actual earnings at  $t - 1$ . The loss at  $t - 1$  is zero if the bidder behaved optimally at  $t - 1$  and positive if the bidder behaved suboptimally at  $t - 1$ . We next pool the data across periods and run probit regressions of adaptively optimal behavior on the loss and other variables by controlling for individual subject and period fixed effects. The regression results are collected in Table 9. In each DI treatment, column (1) reports the results of probit regression with all samples, while controlling for the earnings the players had at  $t - 1$  and for whether they played the best response in period  $t - 1$ . Column (2) reports the results of probit regression with the subsample of the data in which bidders behaved non-optimally at  $t - 1$ . The robust standard errors clustered by individual subject are reported in parentheses.

- Table 9 here -

In both DI treatments with all samples, it is more likely that bidders behave adaptively optimally at  $t$  if their bids are optimal at  $t - 1$ . The higher the actual earnings are in the previous period, the less likely the behavior of bidders is adaptively optimal. The effect of actual earnings at  $t - 1$  is similar even in the regressions with the subsample of non-optimal behavior at  $t - 1$ . More important, the likelihood of adaptive optimality is positively and significantly related to the loss incurred in the previous period. The larger the period  $t - 1$  loss is, the more likely it is for the mid-value bidders to behave adaptively optimally. This tendency is similar for highest-value bidders in both treatments and for lowest-value bidders in the 20-unit treatment. Overall, the regression results confirm that subjects respond to the loss incurred in the previous period and tend to play adaptively optimally when the loss incurred in the previous period gets larger.

**Result 3 (Dynamic behavior)** *(i) The behavior of the SC treatment approximates well the late-period behavior of the DI treatment. (ii) There is evidence suggesting that subjects tend to bid adaptively optimally, particularly in response to the loss incurred in the previous period, but at levels higher than predicted by the balanced-bidding strategy.*

## 5 Explaining the observed behavior

In this section, we explore more deeply into the cause(s) of overbidding relative to the VCG benchmark observed in the data. We do so by focusing on the strategic uncertainty facing the subjects as a basis for explaining the behavior. Motivated by the evidence of bidding rationality (Table 5), we seek to explain subjects' behavior as a rational response to the strategic uncertainty they face. More precisely, we shall investigate whether the observed

bidding behavior can be explained by the beliefs the subjects form based on their opponents' behavior. For clarity of interpretation, we shall focus on the SC experiment, in which subjects bid once with full information about their opponents' values. DI and SI treatments involve many other elements that make it difficult to understand the role of strategic uncertainty.

## 5.1 A model of strategic uncertainty

Our main interests lie in the behavior of the mid-value bidders since their bidding above VCG is largely responsible for higher than expected revenue performance. In order to operationalize the strategic uncertainty facing the mid-value bidders, we assume that the mid-value bidder forms non-degenerate beliefs about the behavior of the highest-value bidder ( $i = 1$ ) and the lowest-value bidder ( $i = 3$ ) conditional on a profile of values  $v = (v_1, v_2, v_3)$ , represented by cumulative distribution functions  $F_i(\cdot|v)$  for  $i = 1$  or  $3$ . Given the beliefs, the mid-value bidder maximizes his expected payoff.<sup>33</sup>

Specifically, the mid-value bidder ( $i = 2$ ) believes bidder  $i = 1, 3$  employs a bid distribution  $F_i(\cdot|v) : [0, \infty) \rightarrow [0, 1]$ . We shall later relate how such beliefs can be estimated from the actual bids made by these bidders.

Given the beliefs, the payoff for the mid-value bidder from bidding  $b \in \{0, 1, 2, \dots, 500\}$  can be expressed as:

$$\pi_2(b, v) = c_A \left\{ \sum_{s=0}^{b-1} (v_2 - s) [F_3(s|v)f_1(s|v) + F_1(s|v)f_3(s|v)] \right\} \quad (11)$$

$$+ c_B \left\{ \sum_{s=0}^{b-1} (v_2 - s) [(1 - F_3(b|v)) f_1(s|v) + (1 - F_1(b|v)) f_3(s|v)] \right\} \quad (12)$$

$$+ (c_A + c_B) \frac{(v_2 - b)}{2} \{F_3(b - 1|v)f_1(b|v) + F_1(b - 1|v)f_3(b|v)\} \quad (13)$$

$$+ c_B \frac{(v_2 - b)}{2} \{(1 - F_3(b|v)) f_1(b|v) + (1 - F_1(b|v)) f_3(b|v)\} \quad (14)$$

$$+ (c_A + c_B) \frac{(v_2 - b)}{3} f_3(b|v)f_1(b|v), \quad (15)$$

where  $f_i(s|v) := F_i(s|v) - F_i(s-1|v)$ , i.e. the probability that bidder  $i$  bids  $s$ . Note that (11) and (12) are the bidder 2's payoffs when his bid is tied with no others while (13) and (14) are his payoffs when his bid is tied with one other bidder at the top and bottom, respectively. Lastly, (15) is the bidder 2's payoff when his bid is tied with two other bidders. The optimal bidding strategy  $b^*(v)$  simply specifies the bid that maximizes the expected payoff:

$$b^*(v) \in \arg \max_b \pi_2(b, v).$$

---

<sup>33</sup>We do not attempt to develop an equilibrium model of strategic uncertainty in which the behavior and beliefs of all three bidders are consistently determined. Bidders' beliefs in our setup are high-dimensional. Thus, such an equilibrium model of strategic uncertainty is unnecessarily complex for our purpose.

Before proceeding with estimation, it is instructive to understand how the strategic uncertainty affects the mid-value bidder's incentive. In particular, it is useful to make the following observation. Let us fix a value profile  $v$ , and let  $\text{supp } F_i$  denote the support of cdf  $F_i(\cdot|v)$  and  $b_2^* = \left(\frac{c_B}{c_A}\right) v_3 + \left(1 - \frac{c_B}{c_A}\right) v_2$  denote the VCG bid for bidder 2.

**Proposition 1.** *If  $\text{supp } F_3 \subset [0, v_3]$ , then any bid  $b > b_2^*$  is (weakly) suboptimal for bidder 2, regardless of bidder 1's strategy.*

*Proof.* Let  $\bar{b}_3 = \int_0^{v_3} s dF_3(s|v)$ . We choose any  $b > b_2^*$  and compare the (ex-post) payoff from such  $b$  to that from  $b_2^*$  against all possible bidder 1's bid  $b_1 \in \text{supp } F_1$ . If  $b_1 > b$  or  $b_1 < b_2^*$ , then bidder 2 is clearly indifferent between bidding  $b$  and  $b_2^*$ . Consider thus  $b_1 \in (b_2^*, b)$ .<sup>34</sup> By bidding  $b$ , bidder 2 obtains  $c_A(v_2 - b_1)$  while, by bidding  $b_2^*$ , he obtains  $c_B(v_2 - \bar{b}_3)$ . The latter payoff is greater than the former since

$$\begin{aligned} c_B(v_2 - \bar{b}_3) - c_A(v_2 - b_1) &= c_A \left[ b_1 - \left(1 - \frac{c_B}{c_A}\right) v_2 - \left(\frac{c_B}{c_A}\right) \bar{b}_3 \right] \\ &\geq c_A \left[ b_1 - \left(1 - \frac{c_B}{c_A}\right) v_2 - \left(\frac{c_B}{c_A}\right) v_3 \right] = c_A [b_1 - b_2^*] > 0, \end{aligned}$$

where the weak inequality holds since  $\bar{b}_3 \leq v_3$ . The above argument shows that bidder 2 is always weakly (strictly) better off with  $b_2^*$  than with  $b > b_2^*$  (if  $b_1 \in (b_2^*, b)$ ), which gives us the desired result. ■

The proposition means that bidding above the VCG level cannot be optimal for the mid-value bidder if he believes the lowest-value bidder will never bid above her value. For strategic uncertainty to explain the higher-than-VCG bidding, therefore, his belief must put a positive probability on the event that the lowest-value bidder bids above her own value with positive probability. Even though this latter strategy (bidding above one's value) is in turn weakly dominated for the lowest-value bidder,<sup>35</sup> its possibility could loom "real" in the mind of the mid-value bidder and thus could lead him to bid above the VCG benchmark. In fact, the possibility of overbidding by the low-value bidders is borne out by the data; they bid above their values about 25% of the time. This latter phenomenon can be related to a commonly observed behavior in auction experiments. The lowest-value bidder in the GSP auction with two items and three bidders faces incentives similar to those faced by the second-highest-value bidder in the single-item second-price sealed-bid auctions. It is documented that the subjects in second-price auction experiments often bid above their values (see Kagel et al., 1987; Kagel and Levin, 1993 for instance) while it is more likely

<sup>34</sup>A similar argument, which is omitted, applies in the case  $b_1 = b$  or  $b_2^*$  so there is a tie if bidder 2 bids  $b$  or  $b_2^*$ , respectively.

<sup>35</sup>It therefore follows that bidding above the VCG benchmark for the mid-value bidder can be ruled out by iteratively deleting the weakly dominated strategies.

for bidders who are more certain that their values are the second-highest (see Andreoni et al., 2007; Cooper and Fang, 2008).<sup>36</sup> One can deduce the forces that cause bidders in the second-price auction experiments to bid above their values may cause lowest-value bidders to do the same in our experiment. Our hypothesis is therefore that this overbidding of the lowest-value bidders pushes the mid-value bidders to bid above the VCG benchmark, as will be seen next.

## 5.2 Estimation

In order to estimate the model, the beliefs must be linked in a meaningful way to the observed data. In this regard, we invoke the natural requirement of consistent beliefs, namely that the mid-value bidder correctly infers the actual distributions of the highest and lowest value bidders. Further, since the mid-value bidder’s beliefs about each opponent,  $F_i(\cdot|v)$  for  $i = 1, 3$ , are high dimensional, we make a simplifying assumption that the profile of observed values  $v$  affects the mid-value bidder’s belief only through the shifting of a distribution function: for any  $v \neq v'$ , there exists a location parameter  $\mu_i^{v,v'}$  such that  $F_i(b - \mu_i^{v,v'}|v) = F_i(b|v')$  for any  $b \in \mathbb{R}_+$ . We use a linear regression framework to model a location parameter  $\mu_i^v$  and the distribution  $F_i(b_i|v)$  for a given value profile  $v = (v_1, v_2, v_3)$  such that, for  $i = 1, 3$ ,

$$b_i = \alpha_{i1} + \beta_{i1}v_1 + \beta_{i2}v_2 + \beta_{i3}v_3 + \varepsilon_i,$$

where the error term  $\varepsilon_i$  follows the distribution  $F_i^\varepsilon$  for  $i = 1, 3$ . The distribution  $F_i(\cdot|v)$ , conditional on a value profile  $v$ , is then a shift of  $F_i^\varepsilon$  such that  $F_i(b_i|v) = F_i^\varepsilon(b_i - \mu_i^v)$  where  $\mu_i^v = \alpha_{i1} + \beta_{i1}v_1 + \beta_{i2}v_2 + \beta_{i3}v_3$ .

The empirical analysis proceeds in the following two stages. We first estimate the distribution of  $\varepsilon$  by taking the (empirical) distribution of residuals from regressing bidder  $i$ ’s bids on the value profile  $v$ . This gives us the estimated distribution  $\widehat{F}_i(\cdot|v)$  for each profile  $v$ . Next, we use the estimated distribution,  $\widehat{F}_i(\cdot|v)$ , to construct the mid-value bidder’s expected payoff as given above.

We take two different approaches to fit the data with the model of strategic uncertainty. The first approach is to compute the mid-value bidder’s optimal bid under strategic uncertainty. This approach amounts to assuming that the (mid-value) bidder makes no decision errors or faces no bid-specific idiosyncratic preferences. The second approach is to place strategic uncertainty in the random utility model (RUM). Specifically, the mid-value bidder

---

<sup>36</sup>Kagel (1995) and Kagel and Levin (2008) provide excellent surveys on the experimental literature of auctions and present the robust patterns of overbidding in the single-unit auctions. There are several competing explanations about overbidding in the single-unit second price auctions such as mistakes, “joy of winning,” and “spite motive.”

is assumed to face a preference shock

$$\tilde{\pi}_2(b, v, \varepsilon_b) = \pi_2(b, v) + \varepsilon_b,$$

where  $\varepsilon_b$  follows the Type I extreme distribution. Then the probability of choosing a bid  $b \in \{0, 1, 2, \dots, 500\}$  is described by the familiar logistic distribution

$$\Pr(b = k|v) = \frac{\exp(\lambda\pi_2(k, v))}{\sum_{s=0}^{500} \exp(\lambda\pi_2(s, v))},$$

where  $\lambda \geq 0$  is a payoff-sensitivity parameter. If  $\lambda$  goes to the infinity, the probability of choosing an optimal bid approaches one. If  $\lambda$  goes to zero, a bid choice becomes purely random. This second approach can be viewed as the optimal bid choice subject to decision errors or idiosyncratic preference shocks for  $b$ . We use the maximum likelihood method to estimate  $\lambda$  from the random utility model, given the estimated distributions of the highest and lowest value bidders from the first stage.

Table 10 collects the summary information about optimal bids under strategic uncertainty and the results of estimation of the random utility model. We report both sample median and mean percentage deviations of optimal bids and RUM-estimated bids from corresponding VCG bids, respectively.<sup>37</sup>

- Table 10 here -

Interestingly, the model of strategic uncertainty with no decision error seems to do a good job of matching the observed mean and median percentage deviations in the SC-11 treatment, while this model predicts less degree of overbidding in the SC-20 treatment than the data. Meanwhile, the RUM of strategic uncertainty appears to fit the data well in the SC-20 treatment, while over-fitting the data in the SC-11 treatment. In order to have a statistical judgement, we conduct the Wilcoxon-Mann-Whitney nonparametric tests of the equivalence of two distributions between observed bids and optimal bids (and RUM-estimated bids, respectively), whose  $p$ -values are reported in parentheses. The test results confirm that the model of strategic uncertainty without decision error fits the data very well in the SC-11 treatment, whereas the RUM of strategic uncertainty accounts well for the data in the SC-20 treatment. Finally, for the graphical presentation of the goodness of fit, we draw the kernel density estimates of percentage deviation of optimal bids (and RUM-estimated bids, respectively) from VCG bids, along with observed percentage deviations from the data. This is presented in Figure 3.

---

<sup>37</sup>When there are multiple optimal bids under strategic uncertainty, we select the lowest of them. On the other hand, for each sample observation (i.e., a given value profile), the estimated random utility model predicts the choice probability distribution over the integer set, from 0 to 500. We chose a median bid in this distribution as an RUM-estimated bid.



- Figure 3 here -

The model of strategic uncertainty without error has a remarkable fit of matching the empirical distribution of percentage deviations from VCG bids in the SC-11 treatment. In this treatment, the distribution of percentage deviation of RUM-estimated bids appears positively more skewed than that of the observed bids. On the other hand, the RUM model appears to have a better match to the data than the model of optimal bids with no error in the SC-20 treatment. This again confirms the Wilcoxon-Mann-Whitney tests in Table 10.

In sum, our model based on strategic uncertainty about opponents' behavior does a remarkably good job of fitting the pattern of mid-value bidder's bidding behavior. Thus, we conclude that the mid-value bidder's strategic uncertainty, combined with his beliefs about the lowest-value bidder's bidding above her own value, is an important ingredient explaining the empirical departures from the VCG predictions in the data.

## 6 Conclusion

The current paper has explored the behavior of bidders participating in the generalized second price auctions—the leading format of allocating sponsored search advertising. We have employed an experimental method to address the outstanding issues with the theory of GSP auctions: (i) the use of a stylized static game of complete information and (ii) the multiplicity of equilibria.

On the first issue, we have found that the static game of complete information does reasonably well in approximating the outcomes and behavior in a more realistic dynamic environment with incomplete information and feedback. This finding is important since it lends support to the prevailing theoretical approach of focusing on the full-information static game as a modeling short-cut. We believe our methodology for testing the adequacy of the stylized model to be useful beyond the current setting, for full-information Nash equilibrium is often adopted as a solution concept in many complex strategic environments.

On the second issue, the bidding data from our experiment displays significant overbidding relative to the leading prediction of the theory, although it is consistent with the weaker predictions of symmetric (or envy-free) Nash equilibria. The departure of the observed bidding from the particular equilibrium selection, namely the VCG outcome, is striking, but this finding should not be regarded as an experimental anomaly or a result of subjects' lack of sophistication. On the contrary, there is an extensive evidence suggesting that subjects' behavior is consistent with rational Nash behavior and reflects sound awareness of the underlying strategic environment. In particular, the mid-value bidder's behavior—a key element of the higher-than-the-VCG revenue—can be explained fairly well as an optimal response

to the strategic uncertainty facing that bidder, suggesting it as an important source of the failure of the prevailing equilibrium selection.

## References

- [1] Andreoni, J., Che, Y., and Kim, J., (2007), “Asymmetric Information About Rivals’ Types in Standard Auctions: An Experiment,” *Games and Economic Behavior*, 59, 240-259.
- [2] Athey, S., and Nekipelov, D, (2010), “A Structural Model of Sponsored Search Advertising Auctions,” Working Paper, Microsoft Research.
- [3] Börgers, T., Cox, I., Pesendorfer, M., and Petricek, V., (2013), “Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence,” *American Economic Journal: Microeconomics*, 5, 163-187.
- [4] Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, Karlin, A., Kominers, S.D., Mathieu, C., and Schwarz, M., (2014), “Convergence of Position Auctions under Myopic Best-Response Dynamics,” *ACM Transactions on Economics and Computation*, 2, 1-20.
- [5] Cooper, D.J. and Fang, H. (2008), “Understanding Overbidding in Second Price Auctions: An Experimental Study,” *Economic Journal*, 118, 1572-95.
- [6] Cox, J.C., J. Shachat, and M. Walker (2001), “An Experiment to Evaluate Bayesian Learning of Nash Equilibrium Play,” *Games and Economic Behavior*, 34, 11-33.
- [7] Cox, J.C., V. L. Smith, and J. M. Walker (1988), “Theory and Individual Behavior of First-Price Auctions,” *Journal of Risk and Uncertainty*, 1, 61-99.
- [8] Day, R. and Milgrom, P. (2008), “Core-Selecting Package Auctions,” *International Journal of Game Theory*, 36, 393-407.
- [9] Edelman, B., Ostrovsky, M. and Schwarz, M., (2007), “Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords,” *American Economic Review*, 97, 242-259.
- [10] Edelman, B. and Schwarz, M., (2010), “Optimal Auction Design and Equilibrium Selection in Sponsored Search Auctions,” *American Economic Review Papers and Proceedings*, 100 (2), 597-602.

- [11] Ergin, H., and Sönmez, T. (2006), “Games of School Choice under the Boston Mechanism,” *Journal of Public Economics*, 90, 215-237.
- [12] Fischbacher, U., (2007), “z-Tree: Zurich Toolbox for Ready-made Economic Experiments,” *Experimental Economics*, 10, 171-178.
- [13] Fukuda, E., Kamijo, Y., Takeuchi, A., Masui, N., and Funaki, Y. (2013), “Theoretical and Experimental Investigations of the Performance of Keyword Auction Mechanisms,” *Rand Journal of Economics*, 45, 47-56.
- [14] Gomes, R., and Sweeney, K. (2014), “Bayes-Nash Equilibria of the Generalized Second Price Auction,” *Games and Economic Behavior*, 86, 421-437.
- [15] Goeree, J. K., C. A. Holt, and T. R. Palfrey (2002), “Quantal Response Equilibrium and Overbidding in Private-Value Auctions,” *Journal of Economic Theory*, 104, 247-272.
- [16] Jeziorski, P. and Segal, I., (2015), “What Makes them Click: Empirical Analysis of Consumer Demand for Search Advertising,” *American Economic Journal: Microeconomics*, 7, 24-53.
- [17] Jordan, J. S. (1991), “Bayesian Learning in Normal Form Games,” *Games and Economic Behavior*, 3, 60-81.
- [18] Kagel, J.H., (1995), “Auctions: A Survey of Experimental Research,” In: Kagel, J., Roth, A. (Eds.), *Handbook of Experimental Economics*. Princeton University Press.
- [19] Kagel, J.H., Harstad, R., and Levin, D. (1987), “Information Impact and Allocation Rules in Auctions with Affiliated Private Values: a Laboratory Study,” *Econometrica*, 55(6), 1275 - 1304.
- [20] Kagel, J.H., and Levin, D. (1991), “The Winner’s Curse and Public Information in Common Value Auctions: Reply,” *American Economic Review*, 81, 362-369.
- [21] Kagel, J.H. and Levin, D. (1993), “Independent Private Value Auctions: Bidder Behavior in First-, Second- and Third price Auctions with Varying Numbers of Bidders,” *Economic Journal*, 103 (419), 868 - 879.
- [22] Kagel, J.H. and D. Levin, (2008), “Auctions: A Survey of Experimental Research, 1995-2008,” forthcoming in *Handbook of Experimental Economics*, Volume II, Kagel, J.H. and A.E. Roth (eds). Princeton University Press.
- [23] Milgrom, P. (2000), “Putting Auction Theory to Work: The Simultaneous Ascending Auction,” *Journal of Political Economy*, 108, 245-272.

- [24] Ostrovsky, M. and Schwarz, M. (2009), “Reserve Prices in Internet Advertising Auctions: A Field Experiment,” *SSRN Working Paper*.
- [25] Pathak, P., and Sönmez, T. (2008), “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism,” *American Economic Review*, 98, 1636-52.
- [26] Varian, H. R. (2007), “Position Auctions,” *International Journal of Industrial Organization*, 25, 1163-78.

Table 1. Frequencies and ratios of efficient allocation

Game	#(A)	Period	Allocation		Ratio	# of obs.
			(A, highest-value) (B, mid-value)	(A, mid-value) (B, highest-value)		
SC	11	1	0.56	0.33	0.83	245
	20	1	0.65	0.18	0.83	245
SI	11	1	0.76	0.15	0.91	245
	20	1	0.77	0.15	0.93	245
DI	11	All	0.53	0.27	0.76	1260
		11 ~ 15	0.55	0.27	0.77	420
	20	All	0.72	0.17	0.87	1260
		11 ~ 15	0.76	0.17	0.91	420

Note: The efficiency ratio is defined as (realized surplus minus random surplus) divided by (first-best surplus minus random surplus).

Table 2. Percentage differences of observed revenues (payments) from theoretical bounds

Lower bound of SNEU

Game	#(A)	Period	Bundle	Median	Std. Err.	Mean	Std. Err.	# of obs.
SC	11	1	A	0.13***	0.03	0.44***	0.06	245
			B	0.00	0.01	0.09	0.08	245
			A & B	0.05***	0.02	0.28***	0.06	245
	20	1	A	0.12***	0.03	0.25***	0.05	245
			B	0.00	0.01	0.16	0.22	245
			A & B	0.03*	0.02	0.16***	0.04	245
SI	11	1	A & B	0.43***	0.08	1.06***	0.09	245
	20	1	A & B	0.34***	0.02	0.46***	0.03	245
DI	11	11 ~ 15	A	0.15***	0.04	0.76***	0.07	420
			B	0.00	0.01	0.75***	0.14	420
			A & B	0.09***	0.03	0.70***	0.08	420
	20	11 ~ 15	A	0.21***	0.02	0.37***	0.04	420
			B	0.09***	0.03	2.03***	0.44	420
			A & B	0.18***	0.01	0.39***	0.04	420

Upper bound of SNEU

Game	#(A)	Period	Bundle	Median	Std. Err.	Mean	Std. Err.	# of obs.
SC	11	1	A	0.02*	0.02	0.22***	0.04	245
			B	0.00	0.01	0.09	0.08	245
			A & B	0.01	0.01	0.15***	0.04	245
	20	1	A	0.00	0.00	0.02	0.04	245
			B	0.00	0.01	0.16	0.22	245
			A & B	-0.03	0.03	-0.02	0.03	245
SI	11	1	A & B	0.37***	0.06	0.84***	0.07	245
	20	1	A & B	0.14***	0.02	0.25***	0.02	245
DI	11	11 ~ 15	A	0.04	0.03	0.51***	0.06	420
			B	0.00	0.01	0.75***	0.14	420
			A & B	0.02	0.03	0.54***	0.07	420
	20	11 ~ 15	A	0.00	0.00	0.11**	0.03	420
			B	0.09***	0.03	2.03***	0.44	420
			A & B	0.04***	0.01	0.15***	0.03	420

Notes: The bootstrap standard error is reported for median. \*, \*\*, and \*\*\* represents 10%, 5%, and 1% significance level for the one-sided  $t$ -test of median (mean) against the alternative hypothesis that the median (mean) is strictly above 0.

Table 3. Median regression results of bids on values

<i>#(A) = 11</i>												
Variables	SC			SI			DI (first period)			DI (last 5 periods)		
	highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value
$v_1$	.232*	.058***	-.000	.881***	-.009	.007	.905***	.048	-.017	.250	-.062	.010
	(.118)	(.023)	(.000)	(.071)	(.019)	(.008)	(.119)	(.133)	(.051)	(.300)	(.101)	(.051)
$v_2$	-.028	.124**	.000	-.038	.966***	.001	.019	.870***	.010	.643**	.685***	.086
	(.119)	(.054)	(.000)	(.042)	(.039)	(.011)	(.132)	(.127)	(.042)	(.275)	(.146)	(.071)
$v_3$	.868***	.846***	1.000***	-.016	.015	.979***	-.051	-.075	.900***	.021	.271**	.872***
	(.068)	(.056)	(.000)	(.042)	(.024)	(.035)	(.088)	(.096)	(.090)	(.242)	(.119)	(.121)
<i>Constant</i>	-5.043	5.313	-.000	9.019	2.078	-.231	27.162	9.760	1.162	-14.032	22.237*	-2.953
	(7.431)	(3.837)	(.000)	(9.024)	(2.862)	(2.133)	(26.124)	(19.347)	(7.549)	(13.752)	(12.085)	(5.098)
Wald test		0.064	1.000		0.000	0.799		0.000	0.728		0.000	0.647
# of obs.	245	245	245	245	245	245	84	84	84	420	420	420
Pseudo R2	0.467	0.610	0.742	0.368	0.532	0.621	0.470	0.567	0.642	0.269	0.345	0.675
<i>#(A) = 20</i>												
Variables	SC			SI			DI (first period)			DI (last 5 periods)		
	highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value
$v_1$	0.908***	.037	.000	1.000***	-.000	.000	.933***	-.030	.014	.449***	.051*	.031
	(.081)	(.052)	(.016)	(.016)	(.002)	(.003)	(.113)	(.046)	(.035)	(.266)	(.029)	(.039)
$v_2$	-.269	.869***	-.000	-.000	1.000***	-.000	-.066	.886***	-.024	.261	.795***	.184*
	(.192)	(.131)	(.040)	(.008)	(.001)	(.004)	(.057)	(.118)	(.042)	(.193)	(.071)	(.110)
$v_3$	.318*	.096	1.000***	-.000	-.000	1.000***	-.072	.027	.920***	.225**	.195***	.727***
	(.192)	(.084)	(.200)	(.010)	(.002)	(.006)	(.072)	(.079)	(.071)	(.099)	(.075)	(.133)
<i>Constant</i>	-14.672	-9.223**	-11.000	2.000	8.000**	9.000**	10.201	15.240	49.698**	-.684	-5.273	56.483*
	(11.912)	(3.761)	(7.889)	(7.379)	(3.618)	(4.343)	(13.217)	(8.767)	(19.574)	(17.824)	(5.560)	(29.619)
Wald test		0.000	1.000		0.000	1.000		0.000	0.349		0.000	0.235
# of obs.	245	245	245	245	245	245	84	84	84	420	420	420
Pseudo R2	0.467	0.588	0.551	0.631	0.772	0.784	0.460	0.702	0.759	0.254	0.559	0.565

Notes: Each regression in the table controls heterogeneity of individual subjects by including dummies for individual subjects. The bootstrap standard errors are reported in parentheses. \*, \*\*, and \*\*\* represents 10%, 5%, and 1% significance level, respectively. The Wald test reports the p-value of testing the null hypothesis that the bids follow the lower bound of SNEU.

Table 4. Percentage differences of observed bid from theoretical bids for mid-value bidders

Lower bound of SNEU

Game	#(A)	Period	Median	Std. Err.	Mean	Std. Err.	# of obs.
SC	11	1	0.19***	0.04	0.58***	0.09	245
	20	1	0.22***	0.06	0.32***	0.09	245
SI	11	1	0.81***	0.10	1.77***	0.24	245
	20	1	0.42***	0.04	0.52***	0.05	245
DI	11	11 ~ 15	0.25***	0.08	1.74**	0.77	420
	20	11 ~ 15	0.22***	0.05	0.46***	0.16	420

Upper bound of SNEU

Game	#(A)	Period	Median	Std. Err.	Mean	Std. Err.	# of obs.
SC	11	1	0.07**	0.04	0.34***	0.07	245
	20	1	0.00	0.02	0.09*	0.07	245
SI	11	1	0.67***	0.07	1.39***	0.22	245
	20	1	0.16***	0.05	0.27***	0.04	245
DI	11	11 ~ 15	0.11*	0.08	1.34**	0.64	420
	20	11 ~ 15	0.01	0.01	0.17*	0.10	420

Value bidding

Game	#(A)	Period	Median	Std. Err.	Mean	Std. Err.	# of obs.
SC	11	1	-0.19	0.05	-0.25	0.02	245
	20	1	0.00	0.01	-0.02	0.07	245
SI	11	1	0.00	0.00	0.11	0.08	245
	20	1	0.03*	0.02	0.11***	0.03	245
DI	11	11 ~ 15	-0.10	0.05	0.21	0.34	420
	20	11 ~ 15	0.00	0.01	0.05	0.09	420

Notes: The robust standard error for mean and the bootstrap standard error for median are reported, both of which are clustered by individual subject. \*, \*\*, and \*\*\* represents 10%, 5%, and 1% significance level for the one-sided t-test of median or mean against the alternative hypothesis that the median or mean is strictly above 0.



Table 5. Frequencies of best response and SNEU

Game	#(A)	Period	Best response			SNEU		
			highest-value	mid-value	lowest-value	highest-value	mid-value	lowest-value
SC	11	1	0.49	0.60	0.88	0.56	0.49	0.72
	20	1	0.70	0.80	0.91	0.58	0.40	0.53
SI	11	1	0.23	0.80	0.93	0.54	0.13	0.58
	20	1	0.60	0.80	0.94	0.33	0.30	0.60
DI	11	11 ~ 15	0.58	0.70	0.86	0.56	0.40	0.58
	20	11 ~ 15	0.75	0.81	0.95	0.71	0.52	0.44

Notes: In checking the consistency of samples with SNEU strategies for mid-value and lowest-value, we relax the conditions with two tokens.

Table 6. Bidding behavior conditional on value (not) being in SNEU  
(Chi-square test in parentheses)

6-1: #(A) = 11

<u>SC treatment</u>	value $\notin$ SNEU	value $\in$ SNEU	<u>DI treatment (last 5 periods)</u>	value $\notin$ SNEU	value $\in$ SNEU
bid $\neq$ value	0.76	0.06	bid $\neq$ value	0.76	0.29
bid = value	0.24	0.94	bid = value	0.24	0.71
# of obs.	210	35	# of obs.	355	65

(p-value = 0.000) (p-value = 0.000)

6-2: #(A) = 20

<u>SC treatment</u>	value $\notin$ SNEU	value $\in$ SNEU	<u>DI treatment (last 5 periods)</u>	value $\notin$ SNEU	value $\in$ SNEU
bid $\neq$ value	0.58	0.45	bid $\neq$ value	0.62	0.41
bid = value	0.42	0.55	bid = value	0.39	0.59
# of obs.	113	132	# of obs.	200	220

(p-value = 0.043) (p-value = 0.000)

Note: When we count frequencies of value bidding, we relax the condition with 2 tokens. In identifying whether value belongs to the set of SNEU, we round up the upper bound of SNEU and down the lower bound of SNEU.

Table 7. Dynamics of percentage differences of observed bid from lower bound of SNEU for mid-value bidders

Game	Period	#(A) = 11					#(A) = 20				
		Median	Std. Err.	Pr(>0)	Median test (p-value)	K-S test (p-value)	Median	Std. Err.	Pr(>0)	Median test (p-value)	K-S test (p-value)
DI	1	0.53	0.19	0.74	0.08	0.00	0.26	0.09	0.81	0.72	0.41
	2	0.51	0.15	0.77	0.04	0.00	0.26	0.08	0.77	0.65	0.76
	3	0.55	0.17	0.77	0.04	0.00	0.21	0.05	0.76	0.94	0.86
	4	0.36	0.13	0.83	0.20	0.01	0.16	0.07	0.70	0.51	0.90
	5	0.30	0.10	0.80	0.32	0.05	0.25	0.05	0.81	0.72	0.36
	6	0.34	0.13	0.82	0.26	0.02	0.21	0.06	0.80	0.90	0.69
	7	0.37	0.13	0.82	0.19	0.05	0.26	0.06	0.85	0.65	0.31
	8	0.25	0.11	0.75	0.58	0.11	0.19	0.05	0.79	0.65	0.69
	9	0.28	0.11	0.77	0.44	0.11	0.14	0.05	0.76	0.31	0.59
	10	0.26	0.09	0.77	0.45	0.16	0.23	0.05	0.82	0.85	0.29
	11	0.26	0.09	0.85	0.49	0.26	0.28	0.07	0.85	0.51	0.28
	12	0.15	0.15	0.76	0.81	0.09	0.22	0.04	0.85	0.97	0.15
	13	0.25	0.10	0.85	0.55	0.28	0.16	0.04	0.81	0.43	0.63
	14	0.32	0.14	0.81	0.35	0.10	0.22	0.05	0.85	0.97	0.10
	15	0.24	0.10	0.79	0.60	0.13	0.29	0.06	0.88	0.41	0.08
SC	1	0.19	0.04	0.86	--	--	0.22	0.06	0.73	--	--
SI	1	0.81	0.09	0.96	--	--	0.41	0.04	0.97	--	--

Notes. The bootstrap standard errors for median, clustered by individual subject, are reported in the second column of each of #(A) = 11 and 20. The median test reports the p-value of equality test of medians, using the t-test, between SC and DI in each period. The K-S test reports the p-value of equality of distributions between SC and DI in each period.

Table 8. Adaptive optimality and percentage differences of observed bids from  
balanced bidding (BB) strategy

#(A)	Period	Adaptive optimality				BB strategy when targeting bundle B is adaptively optimal				# of obs.
		all	highest- value	mid- value	lowest- value	Median	Std. Err.	Mean	Std. Err.	
11	2	0.66	0.43	0.73	0.82	0.73***	0.29	1.49***	0.20	103
	3	0.67	0.51	0.65	0.86	0.38***	0.07	1.07***	0.16	101
	4	0.65	0.48	0.69	0.77	0.29***	0.05	0.94***	0.10	95
	5	0.74	0.58	0.77	0.86	0.18***	0.04	0.84***	0.17	107
	6	0.73	0.56	0.79	0.85	0.18***	0.07	0.96***	0.33	104
	7	0.67	0.50	0.73	0.79	0.22***	0.04	1.23**	0.63	91
	8	0.75	0.63	0.79	0.83	0.23***	0.05	0.68***	0.12	98
	9	0.73	0.64	0.75	0.81	0.14***	0.04	0.56***	0.13	102
	10	0.70	0.62	0.69	0.80	0.13***	0.04	0.47***	0.10	94
	11	0.69	0.58	0.70	0.80	0.10**	0.06	0.49***	0.11	85
	12	0.72	0.57	0.69	0.90	0.07***	0.02	0.46***	0.11	93
	13	0.73	0.57	0.73	0.89	0.12***	0.04	0.49***	0.11	95
	14	0.77	0.60	0.81	0.90	0.10**	0.05	0.68***	0.20	101
	15	0.79	0.67	0.80	0.90	0.06**	0.04	0.56***	0.16	98
	20	2	0.80	0.67	0.86	0.87	0.24***	0.08	0.41***	0.07
3		0.80	0.63	0.82	0.95	0.15***	0.05	0.30***	0.07	81
4		0.80	0.68	0.77	0.94	0.10**	0.05	0.35***	0.09	74
5		0.82	0.71	0.79	0.95	0.12***	0.05	0.32***	0.08	80
6		0.84	0.77	0.83	0.90	0.08**	0.04	0.23***	0.03	81
7		0.84	0.77	0.79	0.96	0.08*	0.05	0.26***	0.06	77
8		0.82	0.76	0.75	0.95	0.09***	0.04	0.26***	0.07	73
9		0.87	0.85	0.81	0.94	0.08***	0.03	0.18***	0.07	85
10		0.85	0.81	0.85	0.89	0.11***	0.03	0.18***	0.05	89
11		0.86	0.81	0.83	0.93	0.08***	0.03	0.26***	0.08	80
12		0.87	0.81	0.85	0.96	0.07***	0.03	0.21***	0.04	79
13		0.85	0.73	0.82	0.99	0.06*	0.04	0.30***	0.08	75
14		0.88	0.80	0.88	0.95	0.13***	0.05	0.94*	0.55	82
15		0.87	0.83	0.81	0.98	0.14***	0.05	0.40***	0.11	80

Notes: The robust standard error for mean and the bootstrap standard error for median are reported, both of which are clustered by individual subject. \*, \*\*, and \*\*\* represents 10%, 5%, and 1% significance level for the one-sided t-test of median (mean) against the alternative hypothesis that the median (mean) is strictly above 0.

Table 9. Probit regression results of adaptive optimality

	#(A) = 11		#(A) = 20	
	(1)	(2)	(1)	(2)
Loss <sub>(t-1)</sub>	.0021*** (.0004)	.0021*** (.0004)	.0023*** (.0005)	.0014** (.0006)
Loss <sub>(t-1)</sub> × 1 {highest-value bidder}	-.0013*** (.0004)	-.0014** (.0004)	-.0017** (.0005)	-.0010 (.0005)
Loss <sub>(t-1)</sub> × 1 {lowest-value bidder}	-.0016*** (.0006)	-.0024*** (.0007)	-.0004 (.0007)	-.0014** (.0007)
Earnings <sub>(t-1)</sub>	-.0007*** (.0001)	-.0012*** (.0002)	-.0002** (.0001)	-.0014*** (.0003)
Optimality <sub>(t-1)</sub>	1.8126*** (.0653)		1.7882*** (.0817)	
Constant	.1425 (.1762)	.2496 (.2463)	-.3808** (.1820)	.2651 (.2999)
Pseudo R2	0.331	0.144	0.247	0.173
# of obs.	3528	1220	3528	655

Notes: We report probit regression results controlling for individual subject and period fixed effects. The variables, Loss<sub>(t-1)</sub>, is defined as optimal *minus* actual earnings at t-1, which is non-negative. The regression equation (1) involves all samples and equation (2) uses a subsample of data where bidders behaved non-optimally at t-1. \*, \*\*, and \*\*\* represents 10%, 5%, and 1% significance level, respectively. The robust standard errors are reported in parentheses.

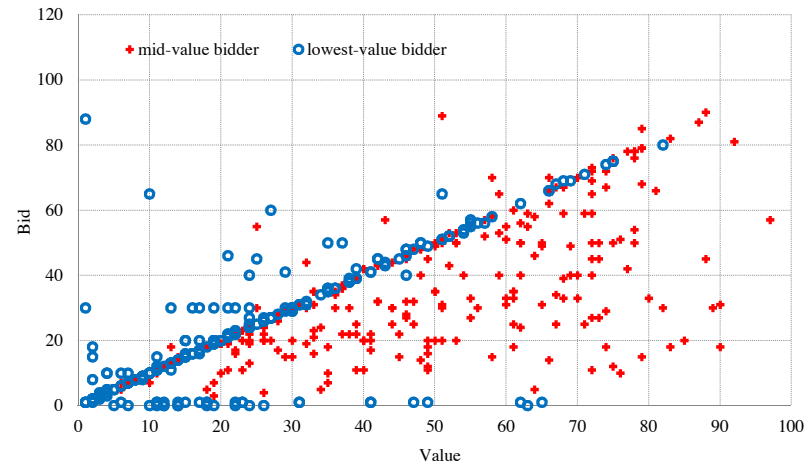
Table 10. Estimated model of strategic uncertainty: percentage deviations from the VCG

Treatment	Model						$\lambda$	Loglik
	Observed bids		Optimal bids		RUM-estimated bids			
	Mean	Median	Mean	Median	Mean	Median		
SC - 11	0.58	0.19	0.38	0.23	0.83	0.43	3.081	-1018.8
			(0.999)		(0.000)		(0.135)	
SC - 20	0.33	0.22	0.09	0.06	0.33	0.18	2.038	-1042.2
			(0.000)		(0.489)		(0.134)	

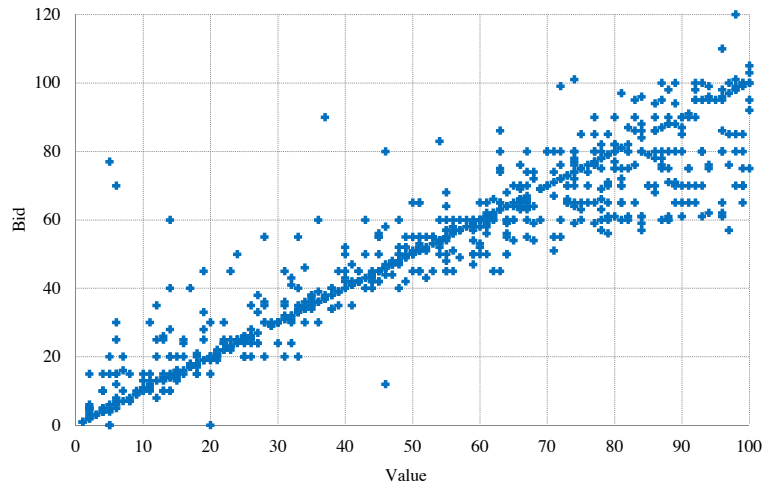
Note. The number in parentheses in columns of Model is the  $p$ -value of the Wilcoxon-Mann-Whitney nonparametric test of the equivalence of two distributions between observed bids and corresponding bids. Standard errors of  $\lambda$  estimates are also reported in parentheses.

Figure 1. Scatter plots of values and bids

1-1. Static complete-information game with  $\#(A) = 11$



1.2. Static incomplete-information game with  $\#(A) = 11$  (all bidders)



1.3. Dynamic incomplete-information game with  $\#(A) = 11$ : last five periods

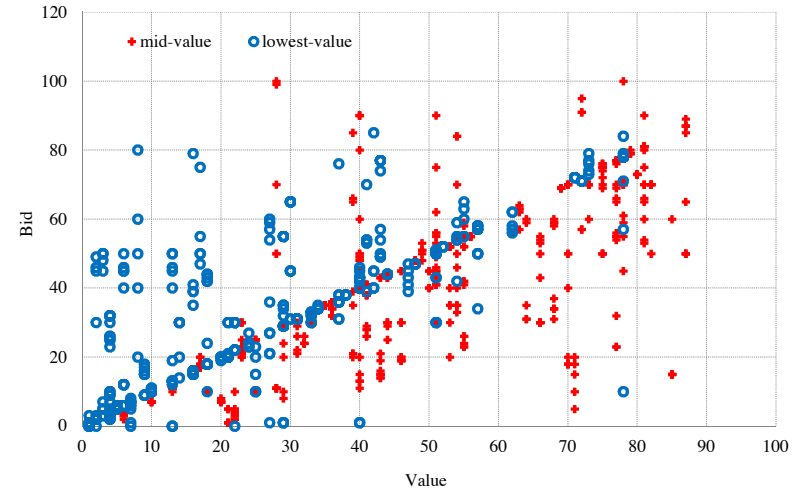
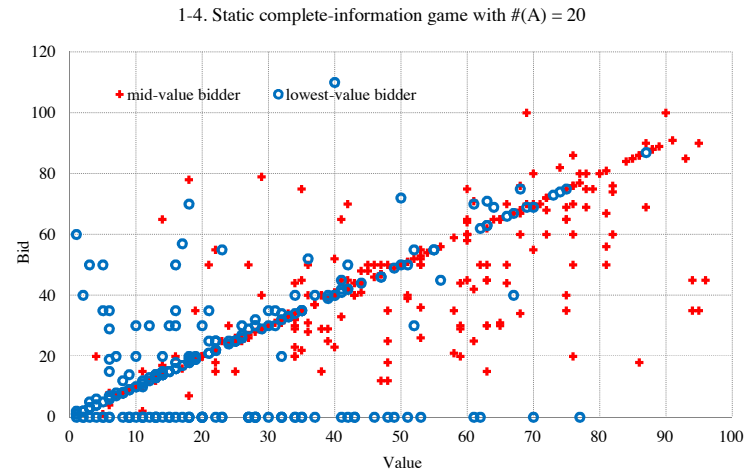
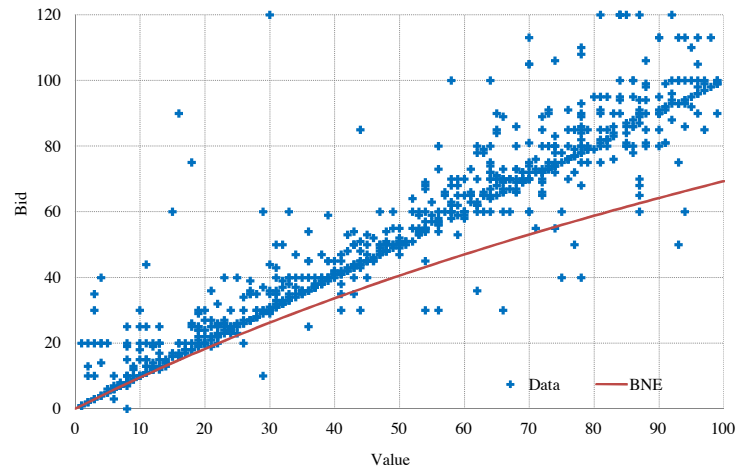


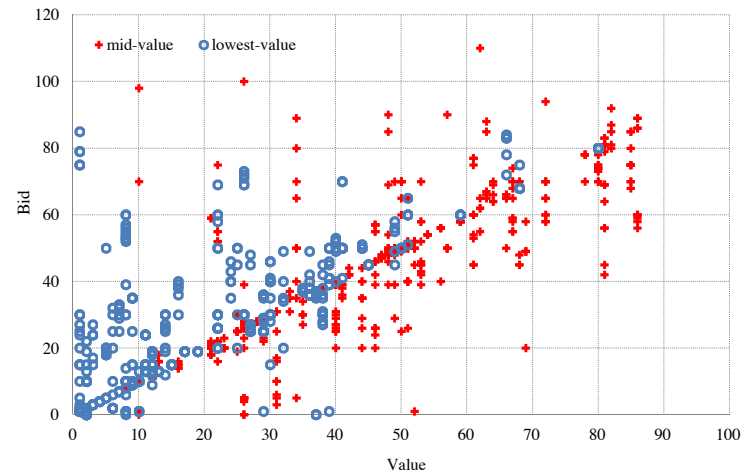
Figure 1 continued



1.5. Static incomplete-information game with  $\#(A) = 20$  (all bidders)



1.6. Dynamic incomplete-information game with  $\#(A) = 20$ : last five periods

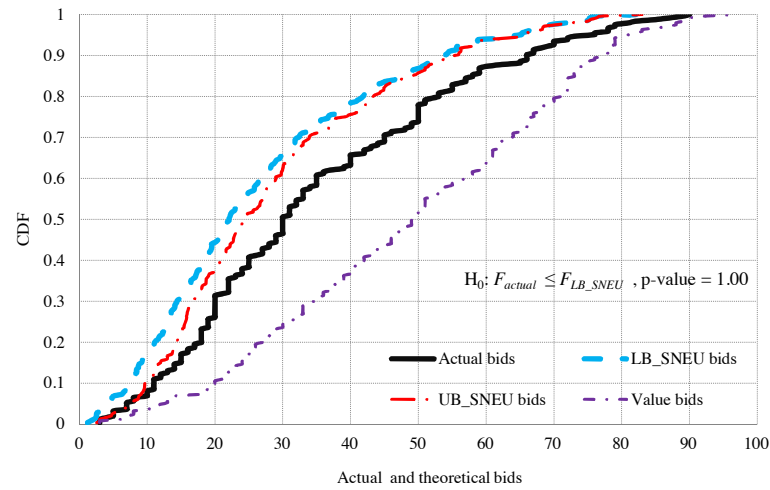


Note: For graphical illustration, we screen out bids above 120. This results in excluding 9 and 12 observations in SI with  $\#(A) = 11$  and 20 (Figure 1-2 and 1-5), respectively, 7 observations for mid-value bidders in DI with  $\#(A) = 11$  (Figure 1-3), and 1 observation for each of mid-value and lowest-value bidders in DI with  $\#(A) = 20$  (Figure 1-6).

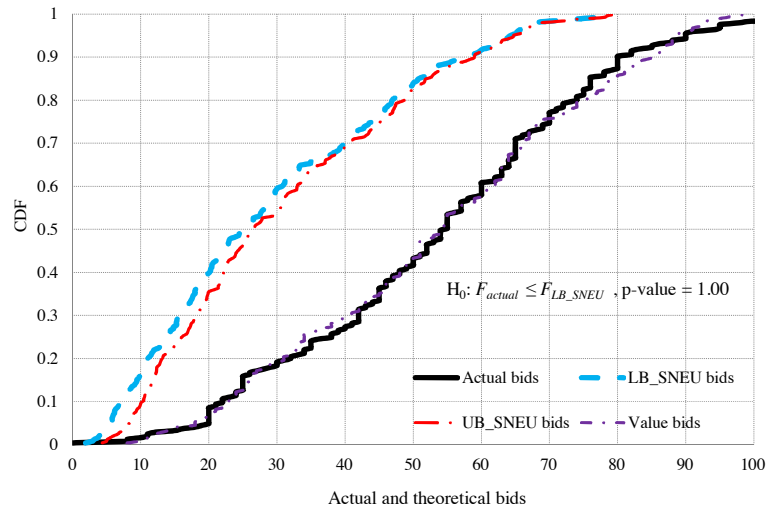


Figure 2. Cumulative distribution functions of observed and theoretical bids for mid-value bidders

2-1. Static complete-information game with  $\#(A) = 11$



2-2. Static incomplete-information game with  $\#(A) = 11$



2-3. Dynamic incomplete-information game with  $\#(A) = 11$ : last five periods

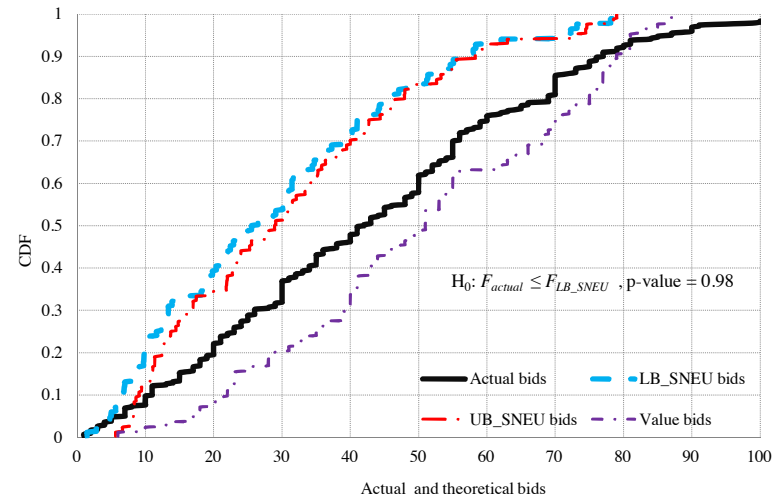
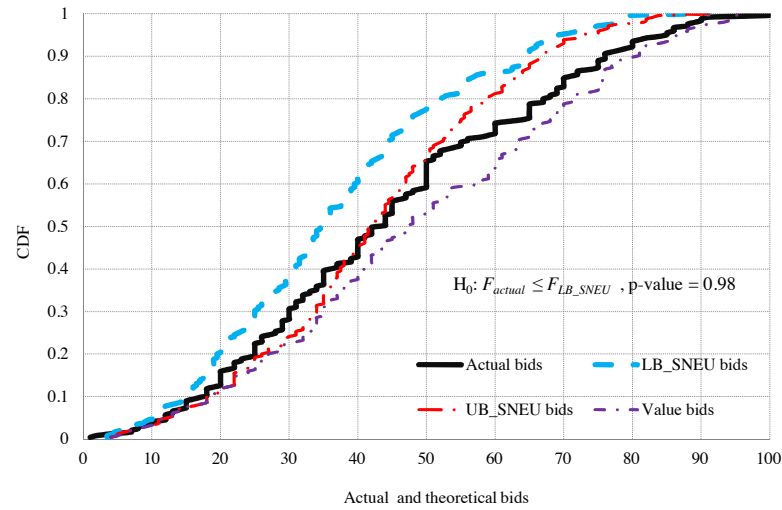
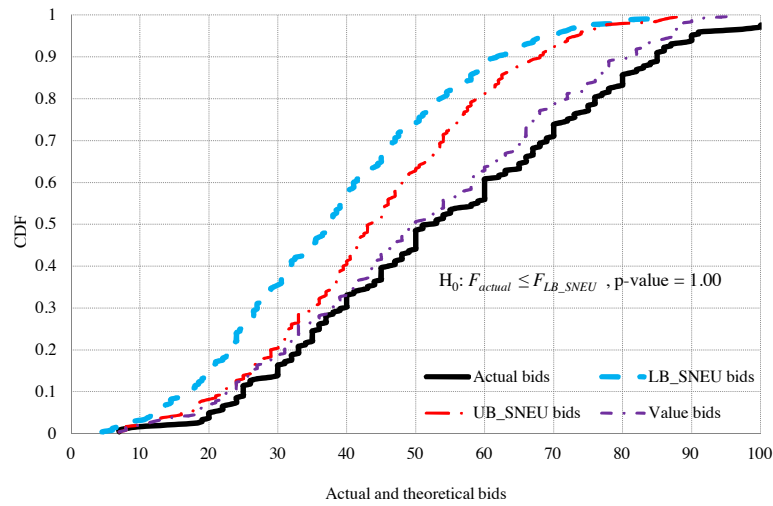


Figure 2 continued

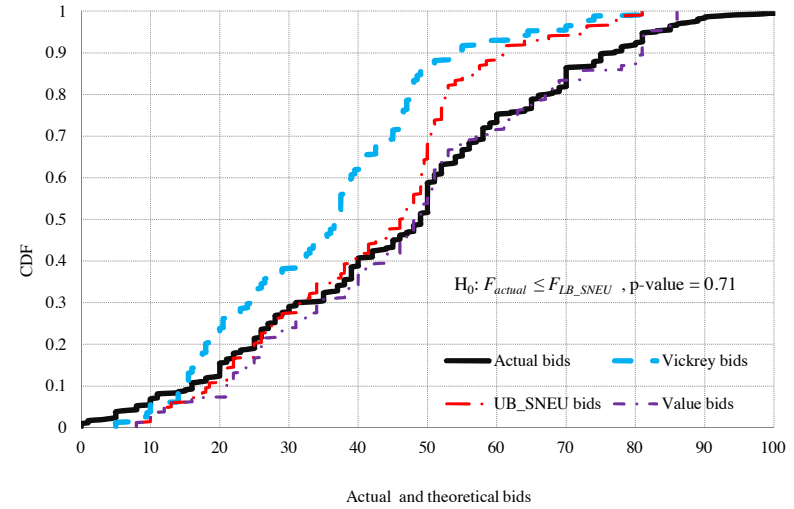
2-4. Static complete-information game with  $\#(A) = 20$



2-5. Static incomplete-information game with  $\#(A) = 20$

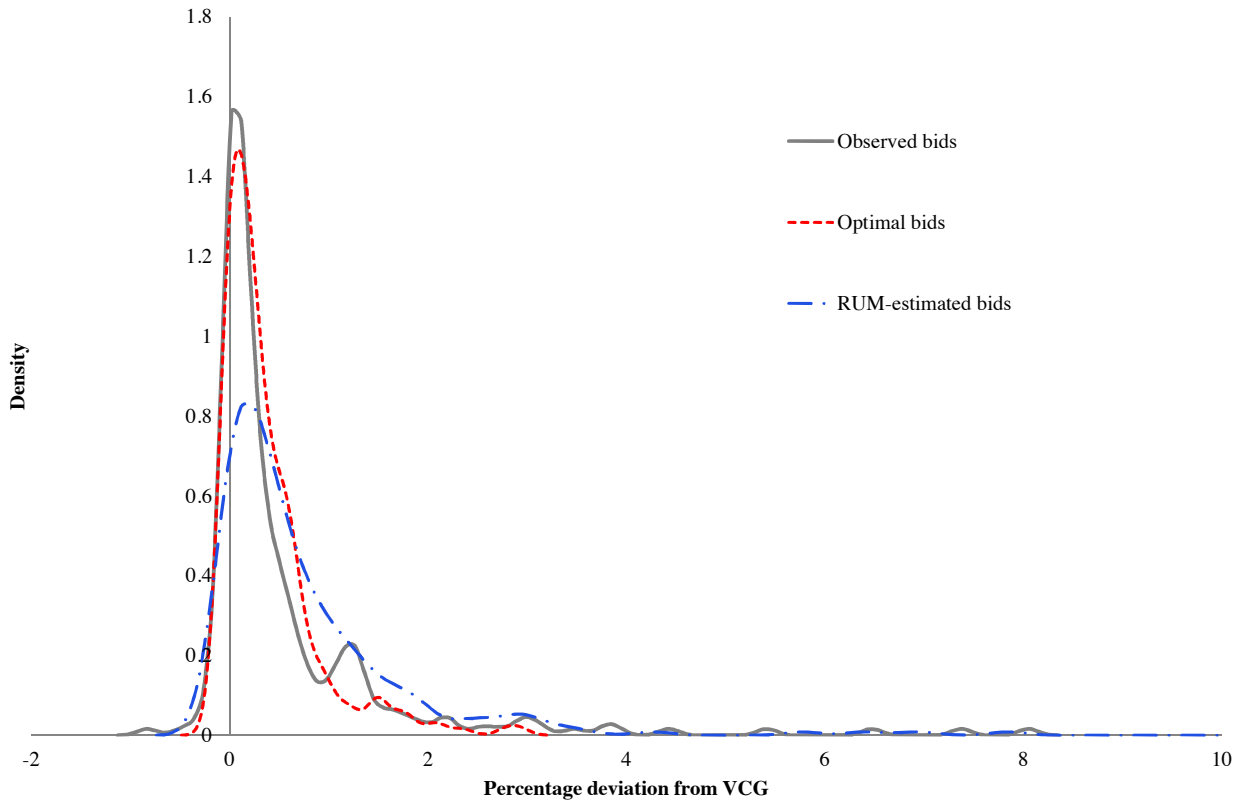


2-6. Dynamic incomplete-information game with  $\#(A) = 20$ : last five periods

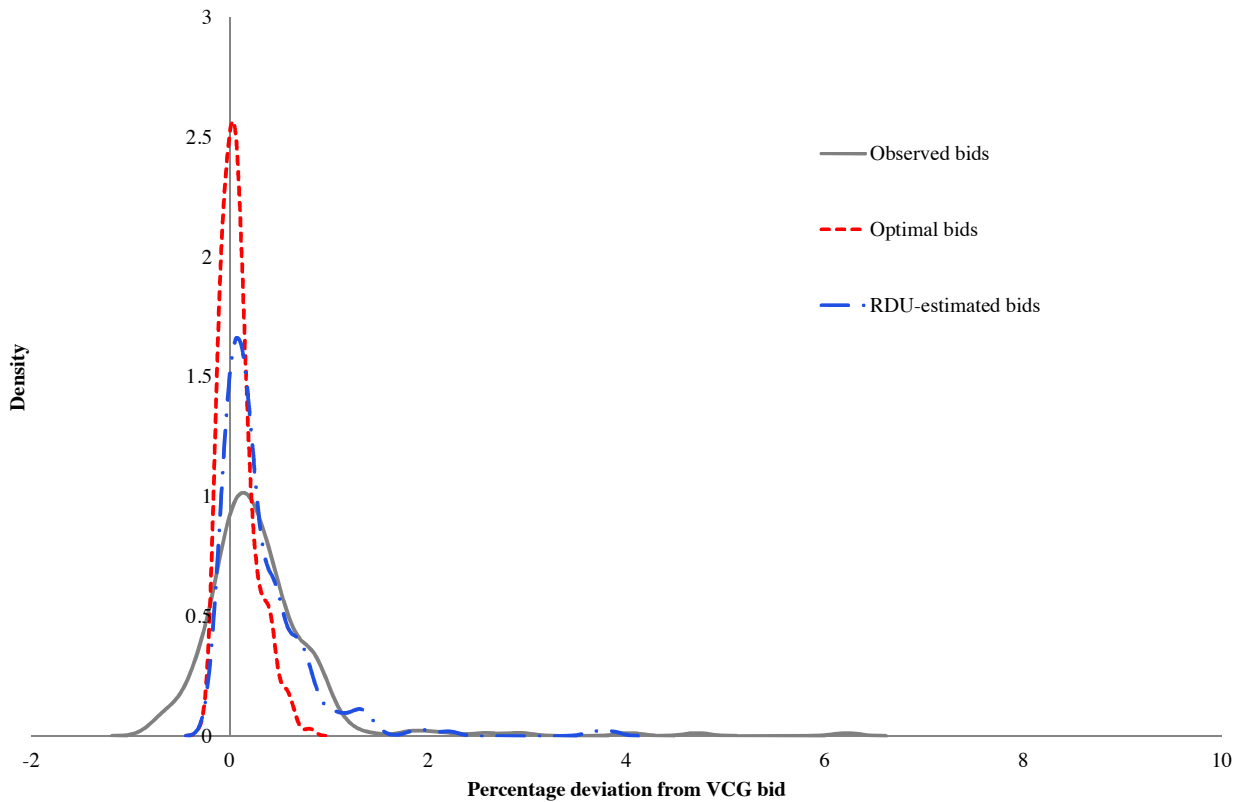


**Figure 3. Goodness of fit of the model of strategic uncertainty**  
(Kernel density estimates of percentage deviation from VCG bids)

A. Static complete-information game with  $\#(A) = 11$



B. Static complete-information game with  $\#(A) = 20$



# Supplementary Material: An Experimental Study of Sponsored-Search Auctions

(Not for Publication)

## Appendix I-A: Long Run Average Perspective

Edelman and Schwarz (2010) use a general specification of a dynamic model with incomplete information, and use the revenue implication of such a model to bound the revenue. While their bound is the equilibrium revenue of the ascending bid auction of EOS with optimal reserve price, this bound can be further tightened to be VCG level in the setup where there is no reserve price, and all positions are sold always. Suppose there are  $n$  bidders, each of which draws a per-click value of  $v \in [0, 1]$ , according to distribution  $F$ , which admits density  $f$  and satisfies the regularity condition:  $J(v) := v - \frac{1-F(v)}{f(v)}$  is strictly increasing in  $v$ . There are  $1 \leq k < n$  positions with position  $a$  generating  $c^a$  units of clicks. There are no reserve prices. There are  $T \geq 2$  periods with possible discount factor  $\delta \in [0, 1)$ . Let  $H_i^t$  be the set of possible histories (or states) that player  $i$  has observed until time  $t$  (prior to his bidding). A Bayesian strategy for player  $i$  is a mapping  $\beta_i$  from his type  $v_i$  to a sequence of functions, one for each period, such that the  $t$  period function maps from  $H_i^t$  to a non-negative real number, his bid at period  $t$ . We assume that each bidder always bids at least zero.

**Proposition 2.** *In any Bayesian Nash equilibrium, the per-period average revenue for the seller cannot exceed her VCG payoff:*

$$\mathbb{E} \left[ \sum_{a=1}^k c^a J(v^{(a)}) \right],$$

where  $v^{(j)}$  is the  $j$ -th highest order statistic. The upper bound for the expected revenue is attained if the positions are assigned efficiently, namely position  $j$  is assigned to the  $j$ -th highest value bidder, but the expected revenue falls strictly below the upper bound if the allocation is inefficient with positive probability in any period  $t = 1, \dots, T$ .

*Proof.* Let  $\beta := (\beta_1, \dots, \beta_n)$  be a Bayesian Nash equilibrium strategy profile of this dynamic game. Let  $x_{it}^a(\beta_1(v_1), \dots, \beta_n(v_n))$  be the probability of bidder  $i$  winning position  $a = 1, \dots, k$  at period  $t$  in that equilibrium, evaluated given period 1 information. Let  $p_i(\beta_1(v_1), \dots, \beta_n(v_n))$  be bidder  $i$ 's expected discounted intertemporal payment at time 1.

Define

$$X_{it}^a(v_i) := \mathbb{E}_{v_{-i}}[x_{it}^a(\beta_1(v_1), \dots, \beta_n(v_n))]$$

and

$$P_i(v_i) := \mathbb{E}_{v_{-i}}[p_i(\beta_1(v_1), \dots, \beta_n(v_n))].$$

Let  $i$ 's interim payoff be

$$U_i(v'|v) := \sum_{t=1}^T \sum_{a=1}^k \delta^{t-1} [v c^a X_{it}^a(v')] - P_i(v'),$$

when he adopts  $\beta(v')$  even though his true type is  $v_i$ , and all others adopt their equilibrium strategies. Since  $\beta$  is Bayesian Nash equilibrium,

$$U_i(v|v) \geq U_i(v'|v), \forall v, v'.$$

By the Envelope theorem,

$$U_i(v|v) = U_i(0|0) + \int_0^v \left[ \sum_{t=1}^T \sum_{a=1}^k \delta^{t-1} c^a X_{it}^a(s) \right] ds = \int_0^v \left[ \sum_{t=1}^T \sum_{a=1}^k \delta^{t-1} c^a X_{it}^a(s) \right] ds.$$

By the usual method, the seller's total expected revenue is

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^n P_i(v_i) \right] &= \mathbb{E} \left[ \sum_{i=1}^n \left( J(v_i) \sum_{t=1}^T \sum_{a=1}^k \delta^{t-1} c^a X_{it}^a(v_i) \right) \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n \sum_{a=1}^k c^a \left( J(v_i) \sum_{t=1}^T \delta^{t-1} X_{it}^a(v_i) \right) \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n \sum_{a=1}^k c^a J(v_i) Y_i^a(v_i) \right] \end{aligned} \quad (16)$$

where  $J(v) := v - \frac{1-F(v)}{f(v)}$  and  $Y_i^a(v_i) := \sum_{t=1}^T \delta^{t-1} X_{it}^a(v_i)$ .

Note that  $X_{it}^a(v_i) \geq 0$  for all  $i, a, v_i, t$ , and  $\sum_{a=1}^k X_{it}^a(v_i) \leq 1$ , which implies that

$$Y_i^a(v_i) \geq 0, \forall a, \text{ and } \sum_{a=1}^k Y_i^a(v_i) \leq \frac{1 - \delta^T}{1 - \delta}, \forall i. \quad (17)$$

Since all positions are assigned in each period,  $\sum_{i=1}^n X_{it}^a(v_i) = 1$ , for each  $t$  and  $a$ , which in turn implies that

$$\sum_{i=1}^n Y_i^a(v_i) = \frac{1 - \delta^T}{1 - \delta}, \forall a. \quad (18)$$

For the vector of random variables  $\{v_1, \dots, v_n\}$ , let  $v^{(j)}$  be the  $j$ -th highest order statistic. If  $J(\cdot)$  is nondecreasing, the maximum of (16) subject to (17) and (18) is achieved point-wise by assortative matching (the bidder with the  $j$ -th highest value is assigned to position  $j$ ) and equals

$$\mathbb{E} \left[ \sum_{a=1}^k c^a J(v^{(a)}) \left( \frac{1 - \delta^T}{1 - \delta} \right) \right].$$

It is straightforward to show that this revenue is precisely what the seller will get from the VCG auction.

If  $J(\cdot)$  is strictly increasing (as is assumed), then the maximum is uniquely attained by the assortative matching. In other words, the VCG revenue is never attained if the matching fails to be assortative in any period with positive probability. ■

## Appendix I-B: Equilibria of Static Incomplete Information Model

Suppose that three bidders independently draw their per-unit values from the interval  $[0, 100]$  following the uniform distribution. Consider the GSP auction in which two positions,  $A$  and  $B$ , with  $c_A$  and  $c_B < c_A$  clicks are offered. Assuming that there exists a symmetric, strictly increasing equilibrium bidding strategy  $b : [0, 100] \rightarrow \mathbb{R}_+$ , we must have for all  $v \in [0, 100]$

$$v = \arg \max_{v' \in [0, 100]} \pi(v'; v) := v [c_A F(v')^2 + c_B 2(1 - F(v'))F(v')] - \int_0^{v'} c_A b(s) dF(s)^2 - 2(1 - F(v')) \int_0^{v'} c_B b(s) dF(s),$$

which yields the first-order condition

$$(c_A F(v) + c_B(1 - F(v)))(v - b(v)) - \int_0^v c_B(v - b(s)) dF(s) = 0.$$

Differentiating both sides with  $v$  and substituting  $F(v) = v/100$  result in the following differential equation

$$(c_A - 2c_B)(v - b(v)) + (c_A v + c_B(100 - v))(1 - b'(v)) - c_B v = 0 \quad (19)$$

along with the boundary condition  $b(0) = 0$ .

With  $c_A = 20$  and  $c_B = 10$ , the differential equation (19) yields a strictly increasing solution,

$$b(v) = 100 \ln[100 + v] - 100 \ln[100].$$

So  $b(\cdot)$  constitutes a Bayesian Nash equilibrium.<sup>38</sup> With  $c_A = 11$  and  $c_B = 10$ , however, (19) does not yield any strictly increasing solution, which implies there is no efficient equilibrium. According to the Proposition 6 of Gomes and Sweeney (2014), no symmetric equilibrium exists.

---

<sup>38</sup>One can verify that the function  $\pi(v'; v)$  is supermodular in  $v'$  and  $v$  so the second-order condition is also satisfied.