An experimental study of the generalized second price auction.

Jinsoo Bae

The Ohio State University

John H. Kagel

The Ohio State University

6/18/2018

Abstract

We experimentally investigate the Generalized Second Price (GSP) auction used to sell advertising positions in online search engines. Two contrasting click through rates (CTRs) are studied, under both static complete and dynamic incomplete information settings. Subjects consistently bid above the Vikrey-Clarke-Grove's (VCG) like equilibrium favored in the theoretical literature. However, bidding, at least qualitatively, satisfies the contrasting outcomes predicted under the two contrasting CTRs. For both CTRs, outcomes under the static complete information environment are similar to those in later rounds of the dynamic incomplete information environment. This supports the theoretical literature that uses the static complete information model as an approximation to the dynamic incomplete information under which advertising positions are allocated in field settings.

We are grateful to Linxin Ye, Jim Peck, Paul J Healy, Kirby Nielsen, Ritesh Jain for valuable comments. This research has been partially supported by NSF grant SES Foundation, SES- 1630288 and grants from the Department of Economics at The Ohio state University. Previous versions of this paper were reported at the 2017 ESA North American Meetings in Richmond, VA. We alone are responsible for any errors or omissions in the research reported.

1. Introduction

Search engines such as Google, Yahoo and Microsoft sell advertisement slots on their search result pages through auctions, among which the *Generalized Second Price* (GSP) auction is the most prevalent format. Under GSP auctions, advertisers submit a single per-click bid. These bids are raked from highest to the lowest, with ad-slots assigned according to the ranking, with each advertiser paying a per-click price equal to the bid submitted by the next-highest bidder.

Edelman et al. (2007) and Varian (2007) were the first to characterize the Nash equilibrium of the GSP auction using a *static complete information* model about competitors' per-click values. They showed that truthful bidding is *not* a dominant strategy, that multiple Nash equilibria exist and that that these equilibria need *not* to be efficient. Edelman et al. (2007) proposed a refinement for the Nash equilibrium referred to as *locally envy-free* equilibria (LEFE), where no bidder would prefer another's slot to her own, given the ad-slot prices. [1] The LEFE predicts efficient allocations of ad-slots but still admits multiple equilibria.

Edelman et al. (2007) further proposed that the LEFE with the lowest possible bids would be the most likely equilibrium to emerge as the long-run outcome in GSP auctions, based on an *ascending clock* version of the GSP auction. Under this outcome, the allocation of ad-slots and the associated payments coincide with those of the dominant strategy equilibrium in the Vikrey-Clarke-Grove's (VCG) auction. In what follows, this refinement will be referred to as the *VCG-like* equilibrium.

Behavior in the GSP auctions is explored here in an experiment with three bidders and two ad-slots under two contrasting CTRs that result in distinctly different bidding behavior. Specifically in one treatment the CTRs are relatively far apart, where the first slots gets 11 clicks and the second slot 3 clicks. In this treatment the VCG-like equilibrium predicts mid-value bidders will employ modest reductions in bids relative to their valuations, while facing minimal competition from the high-value bidder.[2] In the second treatment, CTRs are very close to each other, 11 clicks for the first slot and 10 for the second. In this treatment the VCG-like equilibrium predicts that mid-value bidders engage in sharp price cutting due to the first and second positons

[1] Varian (2007) independently discovered locally envy-free equilibria, referring to them as *symmetric Nash equilibria*.

[2] VCG-like equilibrium predicts low-value holders always bid their values regardless of CTRs. On the other hand, the high-value holders' bids are not pinned down by VCG-like prediction since the highest bids do not determine any prices. Thus, the bidding behavior of the mid-value holders is crucial to examine the GSP auction outcomes under the two different CTRs in relation to the VCG-like equilibrium predictions.

being close to each other, so that the high-value bidder has an incentive to compete for the second position at a favorable price. This sets of Bertrand style competition between the two. The experiment is conducted in a static complete (SC) information environment, and in a dynamic incomplete information (DI) environment, closer in structure to how GSP auctions are conducted in practice.

An additional feature of the experimental design is that while per-click values are assigned randomly across auctions, a fixed ratio is maintained between values, which is required to insure that the contrasting predictions between the two CTRs are maintained.[3] The result is that for the mid-value bidder in the 11-3 treatment the upper bound for an LEFE in undominated strategies (LEFEU) coincides with value bidding, compared to sharp price reductions in the 11-10 treatment. There are also contrasting differences in mid-value bids needed to achieve the VCG-like equilibrium, with bids just above that of the low-value 11-10 treatment, compared to much more modest bid shaving under 11-3.

Behavior is broadly consistent with the predictions of the theory in that there is minimal bid shaving with 11-3 and substantial price shaving with 11-10. The contrast is particularly strong in bidding over rounds in the DI treatment: Under 11-3 bids hover around bidders' values, compared to sharp price cutting over time with 11-10. Efficiency, as traditionally defined in auction experiments, is consistently high averaging over 90% under 11-3 and around 75% under 11-10. Outcomes for mid-value bids under 11-3 lie within the range of LEFEU, but above the upper bound of the LEFEU for 11-10. This is a consequence of the fact that value bidding is at the upper bound of the LEFEU under 11-3, but well above it for 11-10. Bidders consistently bid higher than the VCG prediction under both CTRs, substantially less so under 11-3 than 11-10. Differences between mid-value bids relative to the VCG prediction are the same under SC compared to the last round under DI. This provides support for the idea that SC auctions can serve as a model for bidding in GSP auctions, as actually practiced. However, the predicted VCG-like revenue is not likely to be achieved. Part of this has to do with low-value bidders bidding above value, a common outcome in single unit second-price, private value auctions.[4] Median deviations of mid-value bids

[3] Subjects are unaware of this so that under DI there is no way they can deduce other bidders' values based on their own valuation.

[4] See Kagel et al., (1987), Kagel and Levin (1993), Andreoni et al., (2007), and Cooper and Fang, (2008). Garratt et.al, (2012) show that his results hold in a field setting using subjects who regularly participate in eBay auctions for Morgan silver dollars. In addition, Andreoni et al. (2007) show that with complete information about valuations bidders' exhibit rivalrous behavior, low value bidders bid above their values.

from the VCG-like equilibrium average around 25% and 150% under 11-3 and 11-10 respectively, under both SC and DI. Reasons for these marked differences are discussed below.

There have been several previous experimental studies of GPS auctions, the results of which are mixed.[5] Fukuda et al. (2013) and McLaughlin and Friedman (2016) compared revenues of GSP to sealed bid VCG auctions run as a control treatment, concluding that revenue is indistinguishable between the two. However, Noti et al. (2014), in comparing the two found higher revenue in the GSP compared to VCG auctions. In this experiment subjects were instructed to bidding their values in the VCG auctions was optimal.[6] All three studies employed a limited set of pre-determined valuations, which restricts the potential variation in outcomes that can be observed with random valuations such as those employed here.

The paper that is closest to ours is Che et al. (2017). They use new random valuations in each auction and two contrasting CTR treatments with characteristics similar to the 11-3 and 11-10 treatments employed here. However, unlike here, they use unrestricted random values which create different equilibrium outcomes between auctions, some of which may deviate from the general characteristics of the CTR treatments under study. In what follows, we employ the same fixed ratio for values that avoids this, while still employing random realization valuations. (More on this below.). Further, the fixed ratio employed creates very narrow bounds for LEFEU outcomes, and quite demanding VGC-like equilibrium outcomes for mid-value bidders. Nevertheless we view the two papers as complements, with each providing a similar structure for studying GSP auctions, but with some significant differences in experimental design.

The rest of the paper is organized as follows. The theoretical framework for analyzing GSP auctions is reviewed in section 2. Section 3 describes the experimental design and procedures. Experimental results are reported in section 4. Section 5 concludes by summarizing the findings.

---

[5] See Börgers et al. (2013) and Athey and Nekipelov (2012) for two empirical studies of GSP auctions using field data.

[6] Still bidders consistently bid above their values, with greater overbidding than in the GSP auctions. Fukuda et al (2013) also report bids above value in their VCG auctions for lower valued bidders, but with these bids typically below the value of the next highest valued bidder.

2. Theoretical framework[7]

There are three bidders and two advertising positions, with CTRs c_1 and c_2, where $c_1 > c_2$. Each bidder submits a single per-click bid, with these bids ranked from highest to lowest. The bidder with the kth highest bid wins the kth highest slot and pays the (k+1)th highest bid per click. The bidder with the lowest bid wins nothing and pays nothing. To streamline notation, renumber the bidders in the decreasing order of bids so that $b_1 > b_2 > b_3$ and let v_k be the per-click value of the bidder assigned ad-slot k. The payoff for getting the kth highest slot is $c_k \times (v_k - b_{k+1})$

Definition 1. A Nash equilibrium of the GSP auction is a bid profile $\mathbf{b} = (b_1, b_2\ b_3)$ that satisfies the following inequalities

$$c_1(v_1-b_2) \quad \geq c_2\,(v_1-b_3)$$
$$\geq 0$$
$$c_2(v_2-b_3) \quad \geq c_1\,(v_2-b_1)$$
$$\geq 0$$
$$0 \geq (v_3-b_2)\,c_2$$ [8]

The first two conditions mean the highest bidder has no incentive to deviate to win the bottom position or to lose all positions. The next two conditions mean the second-highest bidder has no incentive to deviate to win the top position or to lose all positions. The last condition means the lowest bidder has no incentive to bid high enough to get one of the two ad-slots. There will typically be a range of bids that satisfy these inequalities, and that these allocations need not be efficient (i.e., assortative). These two properties are not surprising since the Nash equilibrium for single unit second-price auctions (SPA) has the same properties under complete information.[9] However, the GSP auction has a clear disadvantage compared to an SPA as truthful bidding is not a dominant strategy.[10]

Definition 2. A bid profile $\mathbf{b} = (b_1, b_2\ b_3)$ is a locally envy-free equilibrium (LEFE) if it satisfies the following inequalities

$$c_1(v_1-b_2) \quad \geq c_2\,(v_1-b_3)$$

[7] The results here are based on Edelman et al. (2007). Also see Varian (2007).

[8] This implies $0 \geq (v_3-b_1)\,c_1$

[9] Under complete information, an SPA admits multiple Nash equilibrium, including inefficient equilibria. Imagine a case where a lower value holder bids more than the value of the highest value holder, all others bid 0. This constitutes an inefficient Nash equilibrium in which bidders fail to delete weakly dominated strategies.

[10] The GSP auction does not necessarily elicit truthful bids since a bidder may bid below his value to get a lower position at a cheaper price if the CTR of the lower position is high enough (Edelman and Ostrovsky, 2007). Bidding above value is a dominated strategy in the GSP auction.

$$\geq 0$$

$$c_2(v_2\text{-}b_3) \quad \geq c_1 (v_2\text{-}b_2)$$

$$\geq 0$$

$$0 \geq (v_3\text{-}b_3)\ c_2$$ [11]

In a NE bidders may envy a higher position and its associated price, but not so under an LEFE.

For example, suppose that

$$c_1 (v_2\text{-}b_2) \geq c_2 (v_2\text{-}b_3) \geq c_1 (v_2\text{-}b_1)$$

This is a NE but not an LEFE. The second-highest bidder envies the highest bidder who gets ad-slot c_1 and pays b_2 per click (left inequality), but has no incentive to increase her bid to get the top position, as she would earn less than her current earnings (right inequality).

LEFE have several desirable properties: (i) any LEFE is efficient[12] ($v_1>v_2>v_3$), and (ii) any LEFE has a stable assignment of ad positons in that no bidder wishes to exchange his ad-slot and payment with another bidder's ad-slot and their payment.

Definition 3. An LEFE with the lowest bids is obtained by reclusively choosing the lowest bid that satisfies the LEFE.

$$b_3 = v_3$$

$$b_2 = (1\text{-}c_2/c_1)\ v_2 + (c_2/c_1)\ v_3$$

$$b_1 > b_2$$ [13]

Among the set of LEFE, Edelman et al. (2007) suggest that the lowest LEFE is the most plausible outcome, where allocations and payments coincide with the dominant bidding strategies in a VCG auction. [14] Hence the name for this equilibrium – a *VCG-like equilibrium*. The VCG-like equilibrium corresponds to a socially efficient allocation of ad-slots despite the fact that truthful bidding is *not* a dominant strategy.

Note the VCG-like equilibrium predicts a certain amount of bid shaving for the second highest bidder (b2 < v2). However, the amount of bidding shaving depends critically on the ratio

[11] This implies $0 \geq (v_3\text{-}b_2)\ c_1$

[12] However, not all efficient equilibria belong to the LEFE. The LEFE is a small subset of the efficient equilibria.

[13] Note that b_1 does not appear in any conditions of LEFE. Thus, the only condition for b_1 is to be greater than b_2 as assumed.

[14] Edelman et al. (2007) showed that the Lowest LEFE is the unique perfect Bayesian equilibrium in an ascending clock version of GSP auction.

of c2/c1, with the amount of bid shaving increasing as c2/c1 increases. This comparative static prediction is used to determine the experimental treatments – the two CTR sets (c1, c2) = (11, 3) and (11, 10).

While in general there will be sharp price cutting under 11-10 compared to 11-3, there is one complicating factor in maintaining these contrasting predictions across random realizations of click values. For example, if v_2 is close to v_3, there would not be much bid shaving in the 11-10 treatment, contrary to the predicted outcome. On the other hand, if v_2 is much higher than v_3, the amount of bid shaving will be substantial even with 11-3, where minimal bid shaving is predicted. The relative closeness between v_2, v_3 needs to be restricted by $\frac{3}{11} < \frac{v^1-v^2}{v^1-v^3} < \frac{10}{11}$ to maintain the contrasting prediction between the two treatments (see the appendix).[15] Specifically, we use a fixed ratio $\frac{v^1-v^2}{v^1-v^3} = 0.58$ to maintain these contrasting predictions between the two CTRs throughout the experiment.[16] Keeping the same ratio in values also enables normalizing the auction outcomes across auctions, which enables comparing outcomes across auctions as if subjects played the same auction.

While theorists' focus on a VCG-like equilibrium being achieved under SC, a key question is the extent to which this can be achieved, over time, with DI. Appropriate levels of bid shaving on the part of the mid-value bidder are critical to achieving a VCG-like equilibrium. Assuming that bidders do not use dominated strategies (i. e., do not bid above their values), given the normalization in the ratio of values employed, mid-value bids satisfying an LEFE are bounded as follows for 11-3 and 11-10, respectively:

$$\frac{8}{11}\ v_2 + \frac{3}{11} v_3 \le b_2 \le v_2$$

$$\frac{1}{11}\ v_2 + \frac{10}{11} v_3 \le b_2 \le \frac{1}{11}\ v_1 + \frac{10}{11} v_3$$

Where in both cases the lower bound corresponds to the VCG-like equilibrium. In what follows we look at whether mid-value bids satisfy these inequalities for achieving an LEFE in undominated

[15] We use v^k to denote k[th] highest value. This coincides with v_k in an LEFE, which is efficient. Since the two, however, in general could be different, we use v^k whenever the order of values are need to be emphasized.

[16] Since the values in the experiment are integers, they cannot exactly satisfy $\frac{v^1-v^2}{v^1-v^3} = 0.58$. So that draws satisfied $0.57 < \frac{v^1-v^2}{v^1-v^3} < 0.59$.

strategies (which will be referred to as LEFEU) and the degree to which both SC and DI auctions correspond to the VCG-like outcome.

3. Experimental design and procedures

Each experimental session started with 10 static complete information (SC) auctions, with new values randomly drawn from the interval [1, 100] in each auction.[17] These were followed by 8 dynamic incomplete information (DI) auctions with 10 rounds per auction, with new values randomly drawn from the interval [1, 100]. CTRs remained the same in each experimental session – 11-3 or 11-10. Instructions for the SC and DI parts were read separately prior to the start of the treatment in question. Subjects were randomly reassigned to three-person groups prior to the start of each auction.

In the SC auctions information regarding all three bidders' valuations were posted on subjects' screens, with each bidder submitting a single bid for the two ad-slots (named bundles A and B), with the number of items (the CTR value for each bundle), displayed at their computer screens. Bids were ranked from highest to lowest, with the highest bidder getting the larger of the two bundles, with the second-highest bidder getting the smaller bundle. Payoffs in each bundle were equal to the number of items in the bundle multiplied the winners' unit value, at a cost equal to the number of items in the bundle multiplied by the next highest unit-value bid. The lowest value bidder got no items and paid nothing. This was explained to subjects using Figure 1.

Figure 1

Bidder	Ranked Bids	**Unit Valuations**	**Unit Prices**	Bundle Earned	Payoff
2	ub2	V2	ub1	A (11)	(V2 - ub1) x 11= xx
1	ub1	V1	ub3	B (3)	(V1 - ub3) x 3 =zz
3	ub3	V3	-	-	-

Feedback following each auction consisted of a table reporting values, bids, prices paid per unit, the bundle assignment and earnings of all bidders. Subjects were provided with an on screen calculator where they could enter what they believed others would bid and their own bid, which would calculate their expected rank and earnings. Subjects could use the calculator as many times as they wanted within a 90 second time interval for calculating.[18]

[17] These draws, done in advance, were repeated until three values satisfied the ratio on valuations discussed earlier.
[18] See http://econ.ohio-state.edu/kagel/Insts%20with%20screenshots.pdf for a complete set of instructions and screen shots.

The DI auctions followed similar procedures in each of the 10 rounds, except that bidders only knew their own valuations, and payoffs were computed based on one, randomly drawn round of the auction. Subjects had 60 seconds to work the calculator, composing as many hypothetical scenarios as time permitted. Feedback following each auction round consisted of a table reporting back bids, prices paid per unit, the bundle assignment and what own earnings would be if that was the payoff round. [19]

Earnings were in terms of experimental currency units (ECUs). Subjects were provided with starting capital balances of 500 ECUs, with earnings from each auction added to, or subtracted, from this. For the SC auctions earnings were converted into dollars at the rate of rate of 200ECUs = $1 in 11-3 treatment and 320ECUs in the 11-10 treatment, the latter on account of the higher CTRs in 11-10. Earnings for the randomly chosen round for payment in the DI auctions were converted into dollars at a higher rate to compensate for the fact that they were paid for one of the 10 rounds. Earnings averaged $24.19 per subject for sessions lasting 2 hours.

The experiment was run in the Ohio State University Experimental Economics Laboratory between March 2017 and April 2017. Subjects in the experiment were primarily undergraduate students drawn from all disciplines and recruited through ORSEE (Greiner, 2004). Each subject participated in single experimental session. The experiment was computerized, programmed using z-Tree (Fischbacher, 2007). There were three sessions with the 11-3 treatment and three with the 11-10 treatment. Sessions were run with between 12 and 24 subjects in each session.[20]

4. Predicted Outcomes and Propositions to be Investigated.

Based on the fixed ratio between click values $(v^1-v^2)/(v^1-v^3) = 58\%$, predicted outcomes can be discussed in terms of the following normalized valuations - $v^1 = 100$, $v^2 = 42$, and $v^3 = 0$. Normalized bids are used since otherwise results can be misleading.[21] Using these normalized valuations in place of the general restrictions on the bounds for mid-value bids constituting on LEFEU, are reduced to:

$$29.09 \leq b_2 \leq 42$$

[19] Not bidders' valuations in the last round of the DI auctions.

[20] 11-3 sessions had 24, 12, and 18 subjects respectively, and 11-10 sessions had 24, 15, and 15 subjects.

[21] For example, consider two realization of values (100, 42, 0) and (49, 42, 37). In the first case, the VCG-like prediction for the mid-value bidder under 11-3 treatment is 30.5, 40.6 in the second case. So if the mid-value bidder bids his own value 42, the percentage deviation is 31.7% in the first case and 3.4% in the second case. With normalization, the percentage deviation is the same, 31.7%.

for 11-3 and

$$3.64 \leq b_2 \leq 9.09$$

for 11-10.

Several observations are worth noting with respect to these bounds: First, for 11-3 the upper bound for mid-value bids corresponds to value bidding, which may also serve as a focal point. In contrast, for 11-10 the upper bound involves substantial price shaving, cutting bids just over 75% relative to v^2. Second, for 11-10, the lower bound on bids, the VCG-like bid, is just above the normalized value for v^3, whereas it is well above that for 11-3. Among other things, this means that mid-value bids will be much more sensitive to low-value bids above value, as would be anticipated based on single unit second-price auctions.[22] Third, for 11-3 the range of bids satisfying these bounds is twice that of 11-10. Based on the above observations we expect to see substantially greater price cutting under 11-10 compared to 11-3. And that deviations from the VCG-like equilibrium will be greater for 11-10 than 11-3, along with greater deviations from the NE and the LEFEU equilibria.

5. Experimental results

5.1 Efficiency and revenue

Table 1 reports average efficiency across treatments along with the frequency of fully efficient outcomes. For the DI auctions average efficiency is reported for the first, last and all rounds.[23] Several things stand out: First average efficiency is significantly higher under 11-3 compared to 11-10, under both SC and DI. Under 11-3, the average efficiency is remarkably high, averaging over 90% for both SC and DI treatments. In contrast, average efficiency under 11-10 is far less than fully efficient in both cases, ranging between 73.5%-74.7%. Similar differences hold with respect rank order efficiency – the frequency with which bidders with valuations 1, 2 and 3, obtained CTRs ranked 1, 2, and 3 respectively. As will be shown below, these differences are a result of the sharp price cutting over time in 11-10 compared to 11-3. With the downward

[22] See Kagel et al., (1987), Kagel and Levin (1993), Andreoni et al., (2007), and Cooper and Fang, (2008). In addition, Andreoni et al. (2007) show that when the values of other bidders are known, bidders are more likely to bid above their values as they can use that information to reduce the risk of losing money.

[23] Average efficiency is defined as (S_{actual} – S_{random})/ (S_{max}-S_{random}), where S_{actual} is the actual realized surplus from the auction, S_{random} is the mean surplus for all possible allocations, and S_{max} is the maximum possible surplus.

adjustment of prices in over time in 11-10, it is not uncommon for the v^2 bidder to displace the v^1 bidder, along with sporadic bidding above value on the part of the v^3 displacing v^1 and/or v^2.

Table 1
Efficiency
(standard errors of the mean in parentheses)

CTRs	Information	Average efficiency	Frequency of rank order efficiency
11-3	SC	91.7% (2.91)	86.8% (2.85)
	DI (Round 1)	84.6% (3.16)	78.4% (3.43)
	DI (All Rounds)	92.0% (2.19)	86.7% (2.82)
	DI (Round 10)	94.1% (1.63)	88.9% (2.62)
11-10	SC	73.5% (3.88)	56.7% (3.69)
	DI (Round 1)	71.4% (4.46)	62.5% (4.04)
	DI (All Rounds)	75.8% (4.17)	57.9% (4.13)
	DI (Round 10)	74.7% (5.67)	54.8% (4.17)

Not reported are tests for differences in efficiency between the SC and DI treatments. These show that for 11-3, both average and rank order efficiency are significantly lower in the first round of DI ($p < 0.01$) than under SC, as might be expected given the absence of information about bidders' valuations. But there are no significant differences between the two averaged over all DI rounds or DI in round 10. For 11-10, the only significant difference is that rank order efficiency is marginally higher ($p < 0.10$) in the first round of DI bidding compared to SC. However there are no significant differences in rank order efficiency or average efficiency between the two averaged over all rounds of DI, in DI round 10.

One reference point against which to judge these efficiency outcomes is to compare them to outcomes for budget constrained, zero intelligence (ZI) bidders (Gode and Sunder, 1993) who bid randomly between 0 and their valuations. Outcomes for ZI bidders have been shown to track efficiency outcomes quite well in continuous double auction experiments. ZI average efficiency values are 55.4% under 11-3 and 53.6% under 11-10, both well below the levels actually achieved.[24] ZI rank order efficiency averages 47.9% for both CTRs, well below the 11-3 levels, but only marginally lower than under 11-10.[25]

Conclusion 1: Efficiency levels are substantially higher under 11-3 compared to 11-10, which is expected given the sharp price cutting over time with 11-10 compared to 11-3. As reported on below, the sharp price cutting under 11-10 results in some reversals of winning bids relative to valuations. Minimal price cutting under 11-3 precludes this. There are no significant differences between efficiency in the last round with DI compared to the SC for both 11-3 and 11-10. Average efficiency is substantially higher than would have been achieved with zero intelligence bidders.

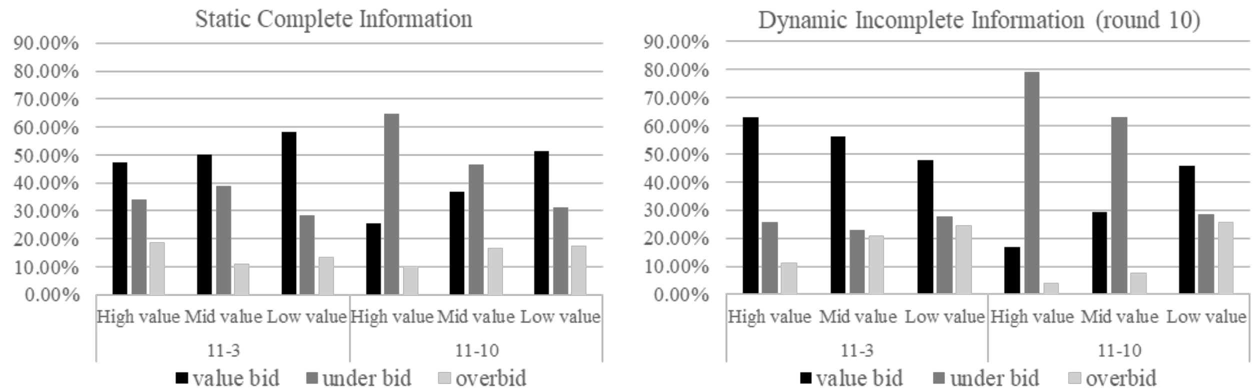
Figure 2 shows the frequency of truthful, under and over bidding relative to valuations under SC (left panel) and for the last round under DI (right panel). With SC, the frequency of truthful bidding for both high and mid-value bidders is substantially higher under 11-3 compared to 11-10. These differences are even greater in the last round for DI. As shown below, under 11-10 bid shaving increases over time under DI, as bidders adjust their bids based on information from previous rounds and competition based on previous rounds bids. The limited bid shaving under 11-3 results in efficiency close to the 100% level reported in Table 1. In contrast, bid shaving under 11-10 often results in an inefficient allocation as high-value bidders try to under-cut mid-value bidders, as a result of strategic uncertainty about their rival's action, as well as their rivals valuations. Although bidding above value is a dominated strategy, it is present, to some extent, for all valuations, consistent with the non-negligible frequency of bidding above value in single unit second-price auctions. Bidding above value for low-value bidders is more common under DI than SC, and tends to be more prevalent for 11-10.

Figure 3 reports the percent of each bundle's allocation relative to valuations. For 11-3 bidders valuations are highly correlated with bundles obtained for both SC and DI. For example, under SC, the frequency with which bundles A, B, and no-bundle were assigned to the high, mid, and low-value bidders was 94.45, 87.5%, and 88.1%, respectively. Under 11-10 the results are quite different. In this case, under SC, bundles A, B and no bundle were assigned to high, medium and low bidders 68.9%, 58.9% and 75.6%, respectively, with most of the missed

[24] Simulations employed 10,000 observations per auction.
[25] The CTRs do not affect the frequency of fully efficient outcomes as the latter only considers the order of bids.

allocations occurring between the high and mid-value bidders. Similar patterns are reported under the DI treatments.

Figure 2. Frequency of Truthful, Under and Over Bidding Relative to Valuations.*



*Truthful bidding includes bids within ± 2 ECUs

Figure 3. Percent of each bundle's allocation in relationship to valuations

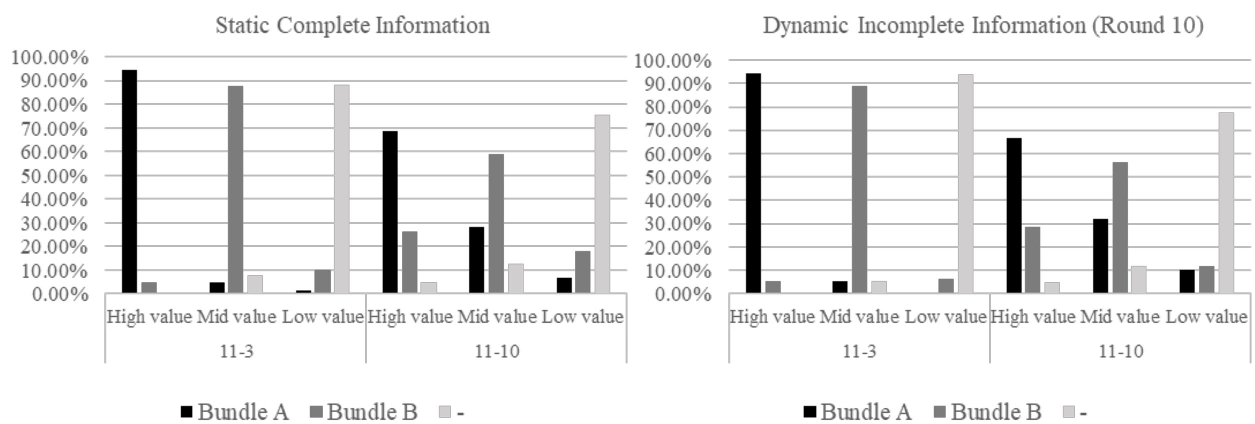


Table 2
Percentage Differences between Observed and Predicted VCG-like Revenue[a]
(standard errors of the mean in parentheses)

CTR	Information	Lower Bound (VCG-like revenue)
11-3	SC	1.15 (2.62)
	DI Round 1	11.7*** (2.6)
	DI All Rounds	12.4*** (3.4)
	DI Last Round	12.2*** (4.0)
11-10	SC	61.0*** (12.2)
	DI Round 1	50.3*** (10.2)
	DI All Rounds	42.7*** (10.9)
	DI Last Round	42.6*** (13.5)

[a] Mean differences. *** Significantly different from zero at $p = 0.01$

Previous studies have focused on the extent to which revenue deviates from a VCG-like equilibrium, particularly with DI information. Table 2 reports these deviations in percentage terms. For the 11-3 SC auctions, average revenue is slightly higher than the VCG-like equilibrium. However, this does not mean that a VCG-like equilibrium has been achieved. Rather, as the data in figure 2 indicates, the VCG-like revenue achieved in this case is a fortuitous combination of value bidding, in conjunction with bidding below and above value. Bidding above value should not be observed in a VCG-like equilibrium. For the 11-3 DI auctions, revenue averages around 12% higher than the VCG-like equilibrium ($p < 0.01$), as this fortuitous combination breaks down a bit.

For 11-10 revenue is substantially higher than the VCG like equilibrium under SC, indicating that this should not be taken for granted, as much of theory does. The interesting thing here is that although revenue is still much higher than the VCG like outcome in the last round of DI bidding, it is *lower* compared to SC, at just under 43% in the last round of DI bidding ($p < 0.05$ under a two-tailed Mann-Whitney test). While part of this is a result of the close to value bidding in one 11-10 SC sessions (see below), this difference remains statistically significant even after dropping this session ($p < 0.05$).

Conclusion 2: Bidding reflects the contrasting predictions with respect to equilibrium outcomes under the two CTRs: Significant price cutting under 11-10 compared to minimal price shaving under 11-3. In both cases, aside from the notable exception of the 11-3 SC sessions, revenues are significantly higher than the VCG-like equilibrium. In 11-3 SC this is a result of fortuitous combination of value bidding, bid shaving, and bidding above value. The latter should not be observed in a true VCG-like equilibrium.

5.2 Bidding over Time

Figure 4 plots median bids separately for high, mid and low-valued bidders under SC in terms of normalized values, along with the VCG-like prediction for mid-value bids. For 11-3, with experience, high-value bids converge to their normalized value in all three sessions. In two of the three session's mid-value bids converge close to their values, *above* the VCG-like outcome. The exception is session 3 where median mid-value bids converge close to the VCG-like equilibrium. Median low-value bids tend to be at or below the zero normalized value. However, a closer look at the data shows that 28.9% of these low value bids were above value, with an average absolute deviation of 24.6 in terms of normalized bids. Obviously, this tends to discourage mid-value bids converging to the VCG-like prediction.

Results are different for 11-10. In sessions 2 and 4, over time high-value bidders cut their bids substantially, relative to their values, with mid-value bids below their normalized value in session 2, and occasionally so in session 4. The price cutting by high-value bidders results in some reversals of winning positions, with the high-value bidder getting bundle B and the mid-value getting A. Session 6 has a different pattern, with high-value bidders reducing their bids substantially in the first several auctions, only to revert to close to value bidding after that. The apparent reason for this is that in auctions 3 and 4 of session 6, 60% of low value bids were above value, with an average absolute deviation of 32.6 in normalized bids.[26] This seems to have spooked high value bidders to cut their bid shaving, reverting to value bidding, taking the pressure off of

[26] For example, in session 6 subject 12 was assigned the high-value (100) in auctions 2 and 3, submitting bids of 55.3 and 27.9, respectively. However, in auction 3, the bidders assigned value 0, wildly overbid at 202.3, with the mid-value holder bidding 41.9, so the high value bidder was not assigned an ad-slot even though he had the high value. From that auction on, subject 12 reverted to value bidding, rarely deviating from it.

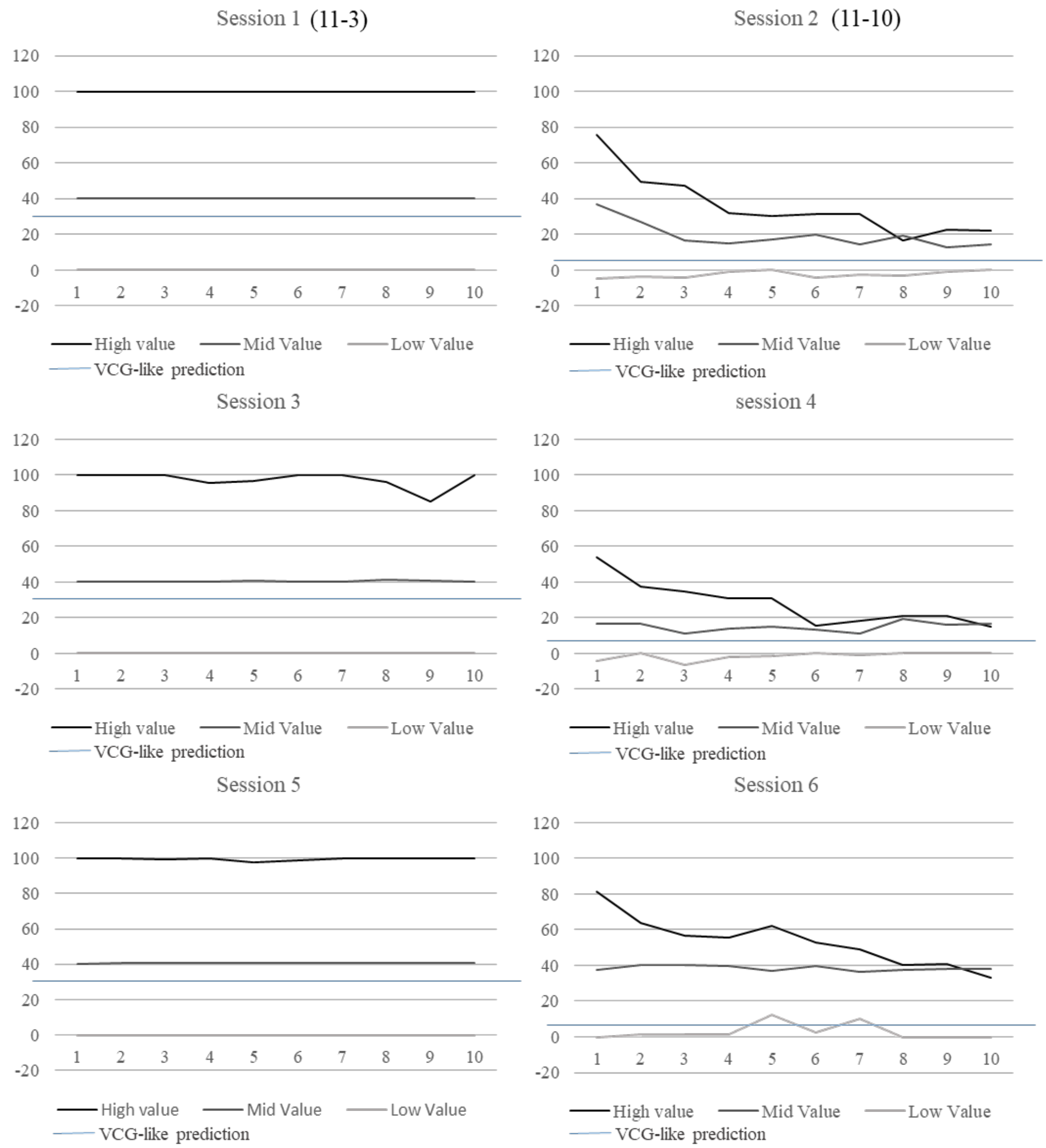
mid-value bidders to bid much below their values. As shown in Table 3 below, low-value bids above value have a strong, statistically significant effect on both high and mid-value bidders for 11-10, but not so for 11-3.

Figure 4. Median bids with over time: Static complete information

Figure 5 plots median bids across arounds under DI. Unlike with SC, under 11-3 there are only minor deviations from value bidding. One explanation for this is that under SC, complete information about the other bidders' values leaves room to bid below value without risking adverse consequences. However, under DI, given the incomplete information, there is little room to bid below one's value, without risking an adverse outcome. This is in line with Gomes and Sweneey (2014) who show that under incomplete information, as the CTR of the second ad-slot gets smaller relative to the first, the Bayesian Nash equilibrium results in bids closer to value.

Under 11-10, there is sharp price cutting on the part of high-value bidders over rounds,with some criss-crossing with mid-value bids over the last 3 rounds. In sessions 2 and 4, there is also sharp price cutting on the part of mid-value bidders, but still above the VCG-like equilibirum, with minimal price cutting in session 6. What is interesting here is that unlike with SC in session 6, under DI high-value bidders engage in sharp price cutting over rounds. This suggests that as information is revealed across rounds, high-value bidders see the opportunity for higher earnings, and are more comfortable going for it. Here too bids above value on the part of low-value bidders result in sharp increases in mid and high-value bids (Table 4 below).

Figure 5. Median bids across rounds: Dynamic incomplete information



Tables 3 and 4 report regressions for bidding across auctions under SC and across rounds within auctions under DI. As with the figures, the regressions are based on normalized values and the corresponding normalized bids. All regressions employ session dummies that have been normalized to sum to zero. These have been suppressed in the Tables to save space.[27] Standard

[27] For 11-3 there are no significant session level effects for SC, and a marginally significant effect ($p = 0.08$) for DI mid-value bids. For 11-10 there are significant differences ($p < 0.05$) between intercepts for both high and mid-value bids, for both SC and DI.

errors are clustered at the subject level and outliers (bids over 300 in normalized bid) were removed. Bidding above value on the part of low-value bidders will no doubt impact both mid and high-value bidders, particularly for 11-10 auctions. Two dummy variables are employed to account for this: Modest bids above value are defined as bids above value but less than or equal to the median overbid on the part of low-value bidders (wOB_{t-1}). Strong bids above value (sOB_t) are defined as bids above value that are greater than the median overbid. For high (mid-value) bidders right hand side variables include the lagged mid-value (high-value) bids. A time trend is added to the DI auctions to capture the obvious time trends reported in the Figure 5. Bid estimation specification is as follow:

$$\mathrm{H_t} = Constant + wOB_{t-1} + sOB_{t-1} + M_{t-1} + \mathrm{t}\ (\text{only in DI}) + \epsilon_\mathrm{t}$$

$$\mathrm{M_t} = Constant + wOB_{t-1} + sOB_{t-1} + H_{t-1} + \mathrm{t}\ (\text{only in DI}) + \epsilon_\mathrm{t}$$

Table 3. Regressions: High and mid-value bids under SC

	11-3 treatments		11-10 treatments	
VARIABLES	High Bid	Mid Bid	High Bid	Mid Bid
wOB_{t-1}	0.602 (5.33)	0.740 (1.76)	12.37 (8.39)	3.852 (4.05)
sOB_{t-1}	0.524 (6.93)	-1.051 (8.60)	33.85** (12.77)	26.48*** (7.44)
M_{t-1}	0.152 (0.153)		0.047 (0.101)	
H_{t-1}		-0.000 (0.001)		-0.011 (0.017)
Constant	93.1*** (6.94)	37.1*** (1.76)	58.6*** (12.3)	32.4*** (3.96)
Observations	151	151	154	154

Standard errors, in parentheses, clustered at the subject level. *** p<0.01, ** p<0.05, * p<0.1

For the SC 11-3 sessions (left hand side of Table 3) none of the right hand side variables, other than the intercept, are statistically significant, nor are these variables significant as a whole.[28] The constant for the mid-value bidder (37.1) is within the range for LEFE bidding in undominated strategies (29.09-42). In contrast, in the 11-10 SC sessions (right hand side of Table 3), strong bids above value in the previous auction prompt higher bids for both high and mid-value bidders. The constant for the high-value bidder (58.6) is well below the value reported for 11-3, consistent with the sharp price discounting compared to 11-3. The constant for mid-value bidders is well above the range for LEFE bidding in undominated strategies – 32.4 compared to a normalized high 9.09.

Table 4. Regressions: High and mid-value bids under DI

	11-3 treatments		11-10 treatments	
VARIABLES	High Bid	Mid Bid	High Bid	Mid Bid
wOB_{t-1}	-4.56 (4.79)	0.300 (0.434)	6.416 (3.96)	6.954*** (1.70)
sOB_{t-1}	8.52 (12.4)	1.041* (0.558)	20.68*** (4.10)	13.85*** (1.75)
M_{t-1}	0.019 (0.014)		0.455*** (0.034)	
H_{t-1}		0.008 (0.006)		0.193*** (0.007)
T	-0.408 (0.334)	0.025 (0.065)	-2.085*** (0.553)	0.154 (0.237)
Constant	96.2*** (4.34)	40.0*** (0.75)	29.8*** (4.00)	7.66*** (1.76)
Observations	1,280	1,280	1,294	1,294

Standard errors, in parentheses, clustered at the subject level. *** p<0.01, ** p<0.05, * p<0.1

Table 4 shows the regression results for the DI treatment. For 11-3, the constants are just below value bidding for both high and mid-value bidders. Strong bids above value have a modest

[28] F values of .88 and .45 for high and mid-value bidders respectively.

impact on mid-value bids, and both time trends are small and not significant.

11-10 is markedly different. Bids of both high and mid-value bidders are responsive to each other's past bids, as the two undercut each other over time. They both increase significantly in response to strong bids above value on the part of low-value bidders. Mid-value bids are also increasing in response to weak bids above value on the part of low-value bidders, as a consequence of the sharp bid shaving on their part.[29] The negative time trend for high-value bidders reflects the sharp price cutting over time in efforts to get the second ad-slot at a favorable price. Constants under 11-10 are sharply lower compared to 11-3, as well as to the constants under 11-10 with SC, with the constant for the mid-value bid within the range for an LEFEU.

Note that low-value bidders typically do *not* bid above v_2, doing so 6.1% on average under SC for both treatments, and typically less than this under DI. Bids of this sort are suggestive of rivalrous behavior, but commonly do not result in losses (8.3% and 12.3% of the time under SC for 11-3 and 11-10) with losses, conditional on winning averaging of \$0.65 and \$0.45 per auction. Losses were less common under DI (4.7% and 8.3% of the time under 11-3 and 11-10) with losses conditional on winning higher - \$1.47 and \$1.04 – reflective of the higher conversion rate to dollars under DI. Rivalrous bidding of this sort has been documented in experiments (Andreoni et al., 2007), as well as in field settings (Cramton, 1997). Nevertheless, these bids clearly discourage sharp discounting of bids for mid-value bidders, as required for an LEFE or VCG-like equilibrium.

Conclusion 3: Bidding over time is qualitatively consistent with expectations regarding differences between the 11-3 and 11-10 treatments. For SC there is minimal bid shaving under 11-3 compared to substantial bid shaving for 11-10 sessions (with the notable exception of session 6 in which low-value bidders were bidding substantially above their values). Further, for 11-3 mid-value bid constants for both SC and DI regressions lie within the range of an LEFEU, as value bidding *is* the upper bound of this interval. For 11-10 under DI there is crisscrossing between mid and high-value bids in later rounds as bidders compete for the second advertising slot at favorable prices. Further, the constant for the mid-value bids lies within the range of an LEFEU, but above the VCG-like equilibrium. The sharp price cutting under DI with 11-10, is a nice example of Bertrand competition with incomplete information.[30]

---

[29] Of course, high and mid-value bidders do not know how far low bids are above value under DI. But they do know how close these bids were to their own value.

[30] There are no experimental studies of Bertrand pricing with incomplete information that we are aware of. The closest approximation to Bertrand competition over time, with incomplete information, are posted-price double auctions markets, where buyers are usually simulated (see Chapter 4 in Davis and Holt, 1993, for a summary of this literature). These to converge slowly, over time, to the competitive equilibrium.

5.3 Nash Equilibrium Bidding and Best Responding

Although bidding is not consistent with a VCG-like equilibrium, the question addressed here is how often was a Nash equilibria (NE) achieved, along with the costs associated with failing to achieve one. Table 5 shows the frequency of bid profiles in the NE in each treatment.[31]

The frequency with which bid profiles are a NE is substantially higher under 11-3 compared to 11-10, with the absolute frequency quite high (low) for 11-3 (11-10). These results are not surprising given the contrasting equilibrium properties of the two treatments: Under 11-3, value bidding constitutes a Nash equilibrium and serves as a focal point. In contrast, value bidding is not a NE under 11-10 and, since both ad-slots have almost the same value, there is a coordination issue in achieving a NE. The net result is that the set of NE is much smaller under 11-10.

Table 5. Frequency of bid profile in Nash equilibria

CTRs	Information	Frequency of NE	# of observation
11-3	SC	76.1%	180
	DI (All periods)	76.3%	1440
	DI (last period)	73.6%	144
11-10	SC	15.0%	180
	DI (All periods)	13.8%	1440
	DI (last period)	18.0%	144

A more interesting question is whether bid profiles evolve to a NE with repetition, particularly in the DI treatment. Figures 6 reports this for SC and DI auctions. Since under 11-10 the NE set is quite narrow, in both cases the frequency with which bid profiles are close to the NE is also reported (close as measured by within a 5 ECU or a 10 ECU radius of the NE).

For both SC and DI there is not much of an increase in NE over time for 11-3 as NE are relatively high to begin with. For 11-10 the frequency of NE increases from the first couple of SC auctions to later auctions, with some dips along the way, it being hard for inexperienced bidders to coordinate to achieve a NE equilibrium, even in a one shot SC game. However, with experience

[31] NEs with dominated strategies are included here. They range between 30-39% for both SC and DI auctions, with low-value bidders responsible for the largest share of these above value bids, with the exception of 11-3 SC where high-value bidders bid above value most often.

bidders learn the equilibrium structure of the 11-10 auctions, resulting in a noticeable increases when accounting for either 5 or 10ECU miscalculations. For 11-10 DI auctions the frequency of NE increases monotonically over rounds within an auction, as bidders utilize information based on prior rounds bids. This increase is particularly striking for both the 5 ECU and 10 ECU bands, to the point that in the last round of bidding, the 10 ECU band achieves a NE around 60% of the time.

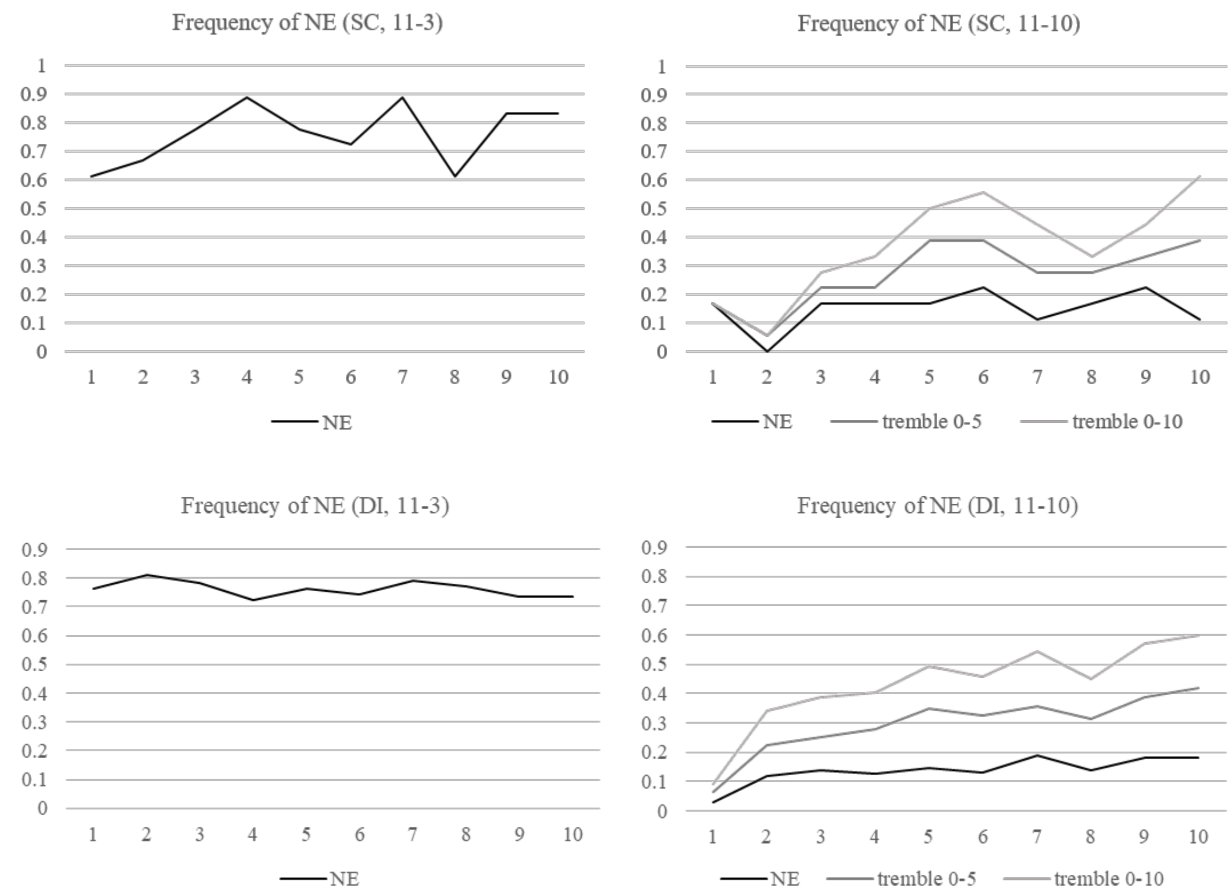


Figure 6. Frequency of Nash Equilibria over Time: SC (top panel) and DI (bottom panel): 11-10 auctions also report outcome within a 5 ECU or a 10 ECU radius of the NE.

Table 6 shows profits of bidders relative to best responding. Under 11-3, bid profiles mostly constitute a NE, so that losses relative to best responding are quite small - $0.05 for SC and $0.13 in the last round with DI. Losses relative to best responding are higher with 11-10 as a consequence of the coordination issues inherent in that treatment, averaging $0.20 with SC and $0.50 in the last round with DI.[32] However, given the large standard errors, the difference in percentage losses between SC and the last round of DI are not significant for both CTRs.

Table 6. Average Losses Relative to Best Responding
(standard errors in parentheses)

	CTR 11-3				CTR 11-10		
	Round	Losses: dollar amount	Profit if BR	Percent loss	Losses: dollar amount	Profit if BR	Percent loss
DI	1	$0.17 (0.12)	$1.88 (0.97)	9.0%	$0.70 (0.35)	$2.57 (1.42)	27.3%
	2	$0.11 (0.11)	$1.82 (1.00)	6.2%	$0.54 (0.25)	$2.26 (1.32)	24.0%
	3	$0.10 (0.03)	$1.80 (1.05)	5.3%	$0.45 (0.21)	$2.34 (1.39)	19.2%
	4	$0.19 (0.08)	$1.93 (1.17)	9.6%	$0.53 (0.32)	$2.36 (1.49)	22.4%
	5	$0.13 (0.05)	$1.90 (1.16)	7.0%	$0.52 (0.30)	$2.34 (1.45)	22.4%
	6	$0.17 (0.07)	$1.98 (1.26)	8.5%	$0.52 (0.30)	$2.37 (1.51)	21.9%
	7	$0.10 (0.03)	$2.05 (1.26)	4.8%	$0.45 (0.29)	$2.41 (1.56)	18.6%
	8	$0.13 (0.06)	$1.94 (1.15)	6.8%	$0.49 (0.29)	$2.42 (1.57)	20.1%
	9	$0.13 (0.05)	$2.02 (1.20)	6.4%	$0.41 (0.21)	$2.35 (1.51)	17.6%
	10	$0.13 (0.04)	$1.99 (1.15)	6.5%	$0.50 (0.35)	$2.33 (1.48)	21.4%
SC		$0.05 (0.03)	$0.73 (0.37)	7.3%	$0.20 (0.08)	$0.79 (0.40)	24.8%

Much of the literature on GSP auctions focuses on the VCG-like equilibrium for SC auctions although in practice GSP auctions involve dynamic incomplete information. The idea being interactions between competitors will converge to a VCG-like equilibrium. Table 2, reported earlier, showed that there were significant deviations from VCG-like with respect to revenue under 11-10 for both SC and DI auctions, as well as for 11-3 auctions with DI. What follows extends this investigation of VCG-like outcomes, in terms of the behavior of mid-value bidders.

Table 7 shows the extent to which median bids of mid-value bidders' deviate from a VCG-like equilibrium along with Kolmogorov-Smirnov tests for whether the distribution of

[32] Note that absolute costs and profits are higher for DI compared to SC since the conversion rate from ECUs to dollars was tripled under DI to compensate that payment was only for 1 out of 10 rounds of bidding.

outcomes differs between SC outcomes and each round of DI bids. For 11-3, median deviations from the VCG-like outcome are not significantly different under SC compared to any round of DI. This is consistent with the argument for focusing on SC auctions, but outcomes under SC are significantly higher (25.9%) than the VCG-like equilibrium. No doubt the similarity between DI rounds compared to SC happens as value bidding serves as a focal point, and under the normalization employed here, is the same for DI and SC.

In contrast, for 11-10 the VCG-like equilibrium entails bidding close to v^3 along with serious price competition between high and mid-value bidders. So in this case the large percentage deviations from value bidding under the VCG-like equilibrium (3.64 compared to 42) are no doubt inhibited by the persistent bidding above value on the part of low-value bidders. However, here too the median percentage deviation between the last round of DI bidding and SC auctions, is consistent with using SC auctions as a model for DI outcomes.[33]

Table 7. Median percentage deviation of mid-value bidders from VCG prediction

CTR		11-3			11-10		
	Rounds	Median %	Std. err.	K-S test[a] (p-value)	Median %	Std. err.	K-S test[a] (p-value)
Dynamic complete information	1	27.7	0.01	0.71	165.4	0.01	0.16
	2	26.8	0.01	0.62	163.8	0.02	0.32
	3	28.3	0.01	0.47	160.9	0.03	0.35
	4	28.1	0.01	0.52	155.0	0.04	0.04**
	5	27.3	0.01	0.68	159.3	0.04	0.13
	6	27.2	0.01	0.57	155.9	0.07	0.03**
	7	26.8	0.01	0.81	156.5	0.04	0.01**
	8	28.6	0.01	0.11	157.3	0.03	0.05**
	9	28.5	0.01	0.52	160.1	0.04	0.05**
	10	28.5	0.01	0.22	152.9	0.05	0.04**
Static complete information		25.9	0.04	-	163.6	0.02	-

[a] Kolmogorov-Smirnov test for equal distribution of outcomes between each DI round and SC. ** Significantly different at the 5% level. The bootstrap standard errors for median are reported.

[33] For 11-10 the difference is significant at better than the 5% level, but not terribly different in terms of the economic significance.

Conclusion 4: The frequency of Nash equilibrium outcomes with 11-3 is relatively high much lower under 11-10 (around 75% versus15%). In addition, average losses relative to best responding are substantially higher under 11-10 (around 25% versus 7% under 11-3). Both of these differences can largely be accounted for by the fact that value bidding constitutes an NE under 11-3, but entails sharp price cutting, along with coordination issues, under 11-10. Median percentage deviations from the VCG-like equilibrium are substantially lower with 11-3. This can be accounted for by the sharp price discounting under 11-10 in conjunction low-value bids commonly exceeding valuations. However, for both CTRs, these percentage deviations from the VCG-like equilibrium are essentially the same between SC and DI auctions, consistent with the theorists focus on SC auctions as a stand in for DI auctions, which are closer in structure the GSP auctions in field settings. But these results also cast doubt on the relevance of the VCG-like equilibrium as a reference point for auction outcomes.

Summary and Conclusions

This paper experimentally investigates outcomes under GSP auctions for selling on line ad-slots. There are two main innovations in the paper: (i) using a cross-over design to compare static complete information (SC) auctions to dynamic incomplete information (DI) auctions so that outcomes are observed for the same set of subjects, and (ii) restricting the relationship between random valuations across auctions designed to maintain the contrasting predictions between the two contrasting click through rates (CTRs) studied. Consistently high efficiency levels are observed for both CTRs under both SC, and in later rounds of DI auctions. The primary comparative static predictions of the theory are shown to be correct in that there is sharp, Bertrand like, competition between bidders when the value of the CTRs are close in value, but minimal competition when they are not. In addition, in comparing SC and DI auctions, SC outcomes are shown to be quite similar to DI outcomes under both sets of CTRs. This provides empirical support for the common theoretical practice of using behavior in SC auctions to model outcomes in DI auctions, where the latter are closer in structure to GSP auctions outside the lab.

GSP auctions have a large number of Nash equilibria (NE). Around 75% of auction outcomes correspond to a NE with the 11-3 set of CTRs, although there is a fairly high frequency of bidding above value on the part of the high value bidder. Strict NE are observed only about 15% of the time under the 11-10 CTRs. Strict NE are inherently much less likely to be observed under 11-10 due to the strong competition for the second ad-slot at a favorable price. However, this rate is substantially higher and increasing with experience after accounting for modest trembles around the NE. One equilibrium that theorists focus on is the VCG-like equilibrium under an auction mechanism that does not support truthful bidding. However, the experiment shows that VCG-like outcomes are not consistently observed. Part of the reason for this is the tendency for low value bidders to bid above their values, a common outcome in single unit second-price auctions conducted the laboratory, and in the one field experiment using experienced bidders.[34] The experimental results also call for new theoretical approaches that incorporate behavioral elements studying GSP auctions.

[34] See footnote 4 for references.

References

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

Athey, S. and D. Nekipelov (2012). A structural model of sponsored search advertising auctions. Unpublished manuscript.

Borgers, T., I. Cox, M. Pesendorfer, and V. Petricek (2013). Equilibrium bids in sponsored search auctions: Theory and evidence. American economic Journal: microeconomics 5 (4), 163- 187.

Cary, M., A. Das, B. Edelman, I. Giotis, K. Heimerl, A. R. Karlin, S. D. Kominers, C. Mathieu, and M. Schwarz (2014). Convergence of position auctions under myopic best-response dynamics. ACM Transactions on Economics and Computation 2 (3), 9.

Che, Y.-K., S. Choi, and J. Kim (2017). An experimental study of sponsored-search auctions. Games and Economic Behavior 102, 20-43.

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

Cramton, P. C. (1997). The FCC spectrum auctions: An early assessment. Journal of Economics and Management Strategy, 6, 431-495.

Davis, D. D. and Holt, C. A. (1993). Experimental Economics. Princeton University Press.

Dufwenberg, M. and U. Gneezy (2000). Price competition and market concentration: an experimental study. International Journal of industrial Organization 18 (1), 7-22.

Edelman, B. and M. Ostrovsky (2007). Strategic bidder behavior in sponsored search auctions. Decision Support Systems 43 (1), 192-198.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. The American Economic Review 97 (1), 242-259.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. Experimental economics 10 (2), 171-178.

Fukuda, E., Y. Kamijo, A. Takeuchi, M. Masui, and Y. Funaki (2013). Theoretical and experimental investigations of the performance of keyword auction mechanisms. The RAND Journal of Economics 44 (3), 438-461.

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

Gode, D. K. and Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. Journal of Political Economy, 101 (1), 119-37.

Gomes, R., & Sweeney, K. (2014). Bayes–Nash equilibria of the generalized second-price auction. Games and Economic Behavior, 86, 421-437.

Greiner, B. et al. (2004). The online recruitment system orsee 2.0-a guide for the organization of experiments in economics. University of Cologne, Working paper series in economics 10 (23), 63-104.

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

Kagel, J. H. and D. Levin. 2009. Implementing efficient multi-object auction institutions: An experimental study of the performance of boundedly rational agents. Games and Economic Behavior 66, 221-237.

McLaughlin, K. and D. Friedman (2016). Online ad auctions: An experiment. Technical report, WZB Discussion Paper.

Noti, G., N. Nisan, and I. Yaniv (2014). An experimental evaluation of bidders' behavior in ad auctions. In Proceedings of the 23rd international conference on World Wide Web, 619-630. ACM.

Varian, H. R. (2007). Position auctions. International Journal of industrial Organization 25 (6), 1163-1178.

Appendix A. Ratio between values in maintaining contrasting predictions.

While VCG-like equilibrium predicts bid shaving regardless of realization of per-click values, true bidding behavior may be governed by another equilibrium solution such as NE which equilibrium property depends on the realization of per click values. If v^2 v^3 are happened to be realized so close that $\frac{10}{11} \leq \frac{v^1-v^2}{v^1-v^3}$, bidding values constitutes a NE under 11-10 treatment (see below proposition) where strong bid shaving is predicted by VCG-like equilibrium. On the contrary, if v^2 is relatively higher than v^3 that $\frac{3}{11} \geq \frac{v^1-v^2}{v^1-v^3}$, value bidding do not constitute a NE under 11-3 treatment as the high-value bidder has incentive to get the second ad-slot at a cheaper price. This incentive to get the lower ad-slot triggers price cutting between high and mid-value bidder, most likely resulting in considerable bid shaving of mid-value bidders in the auction outcome. Thus, although VCG-like equilibrium predicts more bid shaving in 11-10 than 11-3, it is possible to observe opposite outcomes if the values are realized $\frac{10}{11} \leq \frac{v^1-v^2}{v^1-v^3}$ in 11-10 treatment or $\frac{3}{11} \geq \frac{v^1-v^2}{v^1-v^3}$ in 11-3 treatment. On this account we restrict values to be $\frac{3}{11} \leq \frac{v^1-v^2}{v^1-v^3} \leq \frac{10}{11}$.

Proposition 1. Value bidding as a Nash Equilibrium

Given valuations (v^1,v^2,v^3), value bidding constitutes a Nash equilibrium in GSP auctions iff $\frac{c_2}{c_1} \leq \frac{v^1-v^2}{v^1-v^3}$,

Proof. (→) Suppose $\frac{c_2}{c_1} \leq \frac{v^1-v^2}{v^1-v^3}$. We claim that value bidding is a NE. When all bidders submit their own values, it is clear that no bidder wants to outbid a value above his value and get a higher ad-slot. Thus, if no player has an incentive to underbid to get a lower position, value bidding is a NE. Since all winners earn positive payoff, it is obvious that no winners wants to deviate to be ranked 3rd and gets no ad-slot. Thus, the only potentially profitable deviation is the bidder with the top ad-slot deviates to get the second ad-slot. However this deviation is not profitable by assumption.

$$(v^1 - v^2)c_1 \geq (v^2 - v^3)c_2$$

Thus, when $\frac{c_2}{c_1} \leq \frac{v^1-v^2}{v^1-v^3}$, value bidding is a NE.

(←) Now suppose $\frac{c_2}{c_1} > \frac{v^1-v^2}{v^1-v^3}$. Then, value bidding is obviously not a NE. If bidders submit their values, the bidder with the highest value has an incentive to deviate to get the second ad-slot.