Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence*

Tilman Börgers
University of Michigan

Ingemar Cox
University College London

Martin Pesendorfer
London School of Economics

Vaclav Petricek
eHarmony

November 25, 2012

*We are grateful to Sébastien Lahaie, David Pennock, Isabelle Perrigne, Quang Vuong, Rob Porter (the Coeditor) and two referees for very helpful comments, and to Yan-Min Choo and Di Li for superb research assistance. We have also benefited from the insights of a number of seminar audiences. Tilman Börgers thanks the ESRC for financial support under grant RES-000-23-0201 and Martin Pesendorfer thanks the NSF for financial support under grant SES 0214222.

Abstract

This paper presents a game theoretic analysis of the generalized second price auction that the company Overture operated in 2004 to sell sponsored search listings on search engines. We construct a model that embodies few prior assumptions about parameters, and we present results that indicate that this model has under quite general assumptions a multiplicity of Nash equilibria. We then analyze bid data assuming that advertisers choose Nash equilibrium bids. We offer preliminary conclusions about advertisers? true willingness to bid for sponsored search listings. We find that advertisers? true willingness to bid is multi-dimensional and decreasing in listing position.

1 Introduction

Internet search engines sell to advertisers the opportunity to advertise links to their pages on the search result page seen by users who entered a specific search term. These advertisements are called "sponsored links." Sponsored links are displayed on the same page as the links determined by the search engine's own algorithm, but separately from these. The major search engines use auctions to sell spaces for sponsored links. A separate auction is run for each search term. Advertisers' bids determine which advertisers' sponsored links are listed and in which order. The subject of this paper is an early version of an auction of sponsored link spaces that was operated until 2005 by a company called *Overture*.[1] We shall examine a theoretical model of Overture's auction and confront this model with bidding data that we have collected. We seek to extract from the data information about bidders' valuations of sponsored search advertisements, and we seek to understand how bidders respond to the incentives created by the auction rules.

Bidders in Overture's sponsored search auction, and also in the current sponsored search auctions run by Yahoo or Google, for example, bid a payment *per click*. Whenever a search engine user clicks on an advertiser's sponsored link that advertiser has to make a payment to the search engine. The auction format that Overture used, and that is also currently used, in a somewhat different format, by

[1]In 2004, when we collected data about Overture's auction, advertisers bid in Overture's auction for sponsored search listings on Yahoo's search pages. Indeed, Overture, which had started as an independent company, had been acquired at this point by Yahoo, and it was later to be renamed *Yahoo Search Marketing*.

Yahoo and Google, is a "generalized second price auction:"[2] The highest bidder is listed first and pays per click the second highest bid; the second highest bidder is listed second and pays per click the third highest bid; etc.[3]

The generalized second price auction is a method for allocating heterogeneous objects, the positions on a page of search results. It is based on the assumption that bidders agree which object has the highest value, which one has the second highest value, etc. Other auction formats for allocating heterogenous objects do not rely on this assumption. An example is the simultaneous ascending auction described in Milgrom (2000). In this auction, bidders can specify in each round for which object they want to bid. Bids are raised in multiple rounds. Within the limits of the auction rules, they can switch from bidding for one object to bidding for another object. The auction closes when no further bids are raised. By contrast, in the generalized second price auction, bidders submit a single-dimensional bid without specifying what they are bidding for. It seems worthwhile to investigate the properties of this new auction format.

Edelman et. al. (2007) and Varian (2007) have offered theoretical analyses of the generalized second price auction that suggest that the auction may yield an efficient allocation of positions to bidders. These authors' work relies on a relatively narrow specification of bidders' payoff functions: bidders' values per click do not depend on the position in which their advertisement is placed, and click rates are assumed to grow at the same rate for all advertisers as one moves

[2]This expression was introduced by Edelman et. al. (2007).

[3]Today, when ranking advertisers and determining their payments, Yahoo and Google incorporate the likelihood that a user will actually click on the advertisers' link.

up in sponsored link position. These authors' work also relies on a selection from the set of Nash equilibria of the generalized second price auction. The authors focus on equilibria that, although, of course, they are strategic equilibria, are very similar to Walrasian equilibria. A recent working paper, Athey and Nekipelov (2011), extends the auction set-up to incorporate random bidder-specific weights when determining the advertisers ranking. Such weights are a feature that Yahoo and Google use today.

There is very little empirical evidence to this point that justifies the assumptions that Edelman et. al. and Varian incorporated into their models. In our paper the purpose of the theoretical analysis is to lay the grounds for our own, and for future, empirical investigations of data from sponsored search auctions. We therefore propose a more flexible specification of bidders' preferences than is used by Edelman et. al. and Varian, having in mind that further restrictions will have to be motivated by data. We then use the same Nash equilibrium refinement, "symmetric" Nash equilibria, that also Edelman et. al. and Varian used. We find that some of the previous literature's results on symmetric Nash can easily be generalized to our model. However, unlike in previous papers, the symmetric Nash equilibria of the generalized price auction in our paper will be efficient only under restrictive assumptions, but not in general. We also point out that there is a very large set of Nash equilibria that are not symmetric. Our empirical analysis is not based on any selection from the set of Nash equilibria.

We then proceed to an analysis of bidding data for selected search terms. We have collected our data from Overture's website in the spring of 2004. We use a revealed preference approach to infer the structure of bidders' valuations. The more restrictive specifications of preferences used by previous authors are nested by our model, and therefore correspond to parameter restrictions within our model.

The evidence suggests that the properties of valuations that previous authors have postulated do not hold in practice. Our non-parametric revealed preference approach suggests that values per click decline in listing position. Moreover, even with our flexible specification of payoffs we find that we can rationalize most bidders' behavior only over relatively short time periods, after which we have to postulate an unexplained structural break in preferences. Thus we find that it is not easy to rationalize bidding behavior as equilibrium behavior. The most prominent model of equilibrium bidding has great difficulties accounting for real world bidding. Unlike the literature seems to suggest, more theoretical and empirical work is needed to find an empirically satisfactory model of bidding behavior.

Bidding in sponsored search auctions has previously been examined empirically by Edelman and Ostrovsky (2007) and by Varian (2007). Edelman and Ostrovsky's data are from an even earlier version of the Overture auction than we consider. That version was a generalized first price format rather than a generalized second price format. This differentiates their paper from ours. Moreover, unlike us, Edelman and Ostrovsky do not use a structural model of equilibrium bidding, and they do not present valuation estimates in any detail.

Varian (2007) uses bidding data for Google's sponsored search auction on one particular day. He finds evidence that supports a model of equilibrium bidding in which bidders' valuations are not rank dependent. By contrast, we use data that have been collected over a period of several months. To interpret observed bids as equilibrium bids over extended time periods, we need to allow valuations to depend on rank, and we need to allow for structural breaks. Varian's model is based on an equilibrium selection that implies efficiency of equilibria. Our analysis, using a data set that extends over time, and using a more general structural model, does not find evidence of efficiency of equilibria.

While it is a strength of our analysis in comparison to Varian's that our bidding data cover several months, a strength of Varian's analysis is that he has (proprietary) click rates available to him. When interpreting our results it must be kept in mind that our findings may be distorted by the lack of precise click rates.

The theory of sponsored search auctions is also related to the theory of contests and tournaments with multiple, ranked prizes (e.g. Moldovanu and Sela (2001), Moldovanu et. al. (2007)). One can interpret the "effort level" in these models as the bid in our model. However, the generalized second price rule seems specific to the sponsored search auction context.

This paper is organized as follows. Section 2 presents the model. Section 3 discusses Nash equilibria of the model. We reinvestigate "symmetric equilibria" (using Varian's 2007 terminology), as well as other Nash equilibria. Section 4 describes the data. Section 5 reports the results of revealed preference tests. Section 6 concludes.

2 Model

There are K positions $k = 1, 2, \ldots, K$ for sale, and there are N potential advertisers $i = 1, 2, \ldots, N$. We shall refer to the potential advertisers as "bidders." We assume $K \geq 2$ and $N \geq K$. Bidders $i = 1, 2, \ldots, N$ simultaneously submit one-dimensional non-negative bids $b_i \in \Re_+$. Bids are interpreted as payments *per click*. The highest bidder wins position 1, the second highest bidder wins position 2, etc. The bidder with the K-th highest bid wins position K. All remaining bidders win no position. The highest bidder pays per click the second highest bid, the second highest bidder pays per click the third highest bid, etc. The K-th highest bidder pays per click the $K + 1$-th highest bid if there is such a bid. Otherwise, if $N = K$, the K-th highest bidder pays nothing. We will explain later how we deal with identical bids, i.e. ties. We follow Edelman et. al. (2007) and refer to this auction as a "generalized second price auction."

The payoff to bidder i of being in position k if he has to pay b per click is:

$$c_i^k(\gamma_i^k - b) + \omega_i^k \tag{1}$$

Here, $c_i^k > 0$ is the *click rate* that bidder i anticipates if he is in position k, that is, the total number of clicks that bidder i will receive in the time period for which the positioning resulting from the auction is valid. Next, $\gamma_i^k > 0$ is the *value per click* for bidder i if he is in position k. This is the profit that bidder i will make from each click on his advertisement. Finally, $\omega_i^k \geq 0$ is the *impression value* of being in position k for bidder i. The impression value describes the value that bidder i derives from merely being seen in position k, independent of whether a search engine user clicks on bidder i's link. We have in mind that companies derive value from the fact that a sponsored search link reminds customers of the existence of their company, and that it makes users more likely to buy in the future, even if those users do not click on the link and make a purchase at the time of their search.[4]

Our representation of bidders' payoffs is "reduced form," that is, we do not describe explicitly the behavior of users of search engines that generates bidders' payoffs. One reason for not modeling users' behavior explicitly is that this behavior is presumably driven not only by economic considerations, but also by human physiology (where do people look first on a computer screen?) and psychology, and we do not know of good ways of capturing these factors in a model. Another reason is that we only have bidding data, not data about users' behavior.[5]

A restrictive assumption implicit in equation (1) is that click rate, value per click, and impression value for bidder i in position k do not depend on the identity of the bidders that win other positions. In practice, this identity might matter. Bidder i might attract a larger click rate in second place if the bidder in the top position is a large, widely known company than if the bidder in the top position is small and not well-known. In auction theory, this is known as an "allocative externality." It is well-known that such externalities may create multiple equilibria in single unit auctions (Jehiel and Moldovanu, 2006). In our multi-unit auction,

[4]Note that we do not rule out that the impression value is zero.

[5]Athey and Ellison (2011) offer a model in which search engine users are described as expected utility maximizing economic agents.

we find multiple equilibria even with the specification of payoffs given in (1). By leaving allocative externalities out of our model we thus identify a different source of multiplicity of equilibria. Our modeling choice also reflects that we do not attempt to identify and measure allocative externalities. Measuring allocative externalities would require sufficient data variation in the allocation realization which we cannot guarantee as our data set is too small.

Equation (1) seems to assume that bidders know click rates, values per click, and impression values. We can, however, allow the possibility that bidders are uncertain about these variables, and maximize the expected value of the expression in (1). The expected value will have the same form as (1), with all three variables replaced by their expected value, if all three variables involved are stochastically independent.

We shall refer to the value of b which makes the payoff in expression (1) zero as *bidder i's willingness to bid for position k*. We denote it by v_i^k:

$$v_i^k = \gamma_i^k + \frac{1}{c_i^k}\omega_i^k \tag{2}$$

We can now equivalently write bidder i's payoff as:

$$c_i^k(v_i^k - b) \tag{3}$$

This expression makes clear that our model is equivalent to one in which there is no impression value, and the value per click is v_i^k rather than γ_i^k. We shall conduct our analysis using expression (3), but it will be useful to keep in mind that the model admits the alternative interpretation in expression (1).

Our model nests as special cases those of Lahaie (2006), Edelman et. al (2007), and Varian (2007). These authors assume that the values per click are independent of the position, that is, for every $i = 1, 2, \ldots, N$ there is some constant v_i such that:

$$v_i^k = v_i \text{ for all } k = 1, 2, \ldots, K \tag{4}$$

and that the ratio of the click rate of one position to that of another is the same across all bidders, that is, for every bidder $i = 1, 2, \ldots, N$ and every position $k = 1, 2, \ldots, K$ there are numbers a_i and c^k such that:

$$c_i^k = a_i c^k \tag{5}$$

Our analysis is more general than the analysis in the papers cited above, although in Propositions 2 and 3 below we shall focus on the specification in equation (5). We start with a more general specification than previous papers for two reasons. First, a priori neither of the two generalizations seems unreasonable. The value of a click may depend on the position from which the click is received, if, for example, search engine users work their way from the top of the page to the bottom, and pick the first advertiser whose offerings match their needs. Then clicks that arrive at an advertisers' page from a bottom position represent buyers with specialized interests for whom the profit per click may be different than for other bidders. The click rate ratio for different positions may depend on the identity of the advertiser if, for example, one bidder is a prominent internet book trader, whereas the other advertiser is not so prominent. For a prominent advertiser dropping one position may lead to a lower percentage loss of clicks than for a less prominent advertiser. Our second argument for working with a more general specification than previous literature is that we wish to allow the data to inform us about model parameters, and therefore we want to embed as few assumptions as possible in the model. Unfortunately, our data will not allow us to draw inferences about click rates, but the data will suggest that the value of a click to an advertiser depends on the position of the advertiser.

We shall study pure strategy Nash equilibria of the auction game. A pure strategy Nash equilibrium is a vector of bids $(b_1, b_2, \ldots, b_N)$ such that each bid maximizes the bidder's payoffs when the bids of the other bidders are taken as given. To give a formal definition, we need to deal with ties. A *ranking* of bidders is a bijection $\phi : \{1, 2, \ldots, N\} \to \{1, 2, \ldots N\}$ that assigns to each rank ℓ the bidder $\phi(\ell)$ who is in that rank. A ranking of bidders is compatible with a given bid vector $(b_1, b_2, \ldots, b_N)$ if $\ell \leq \ell' \Rightarrow b_{\phi(\ell)} \geq b_{\phi(\ell')}$, that is, higher ranks are assigned to bidders with higher bids, where ties can be resolved arbitrarily. A ranking of bidders that is compatible with a given bid vector thus represents one admissible way of resolving ties in this bid vector. We now define a Nash equilibrium to be a bid vector for which there is some compatible ranking of bidders so that no bidder has an incentive to unilaterally change their bid.

Definition 1. *A vector of bids* $(b_1, b_2, \ldots, b_N)$ *is a* Nash equilibrium *if there is a compatible ranking* ϕ *of bidders such that:*

- *For all positions k with $1 \le k \le K$ and all alternative positions k' with $k < k' \le K$:*

$$c_{\phi(k)}^{k}\left(v_{\phi(k)}^{k} - b_{\phi(k+1)}\right) \ge c_{\phi(k)}^{k'}\left(v_{\phi(k)}^{k'} - b_{\phi(k'+1)}\right)$$

- *For all positions k with $1 \le k \le K$ and all alternative positions k' with $1 \le k' < k$:*

$$c_{\phi(k)}^{k}\left(v_{\phi(k)}^{k} - b_{\phi(k+1)}\right) \ge c_{\phi(k)}^{k'}\left(v_{\phi(k)}^{k'} - b_{\phi(k')}\right)$$

- *For all positions k with $k \le K$:*

$$c_{\phi(k)}^{k}\left(v_{\phi(k)}^{k} - b_{\phi(k+1)}\right) \ge 0$$

- *For all ranks ℓ with $\ell \ge K+1$ and all positions k with $1 \le k \le K$:*

$$c_{\phi(\ell)}^{k}\left(v_{\phi(\ell)}^{k} - b_{\phi(k)}\right) \le 0$$

Here, if $K = N$, we define $b_{\phi(N+1)} = 0$.

The first two conditions say that no bidder who wins a position has an incentive to deviate and bid for a lower or for a higher position. Note the following asymmetry. A bidder who bids for a lower position k has to pay $b_{\phi(k+1)}$ to win that position, but a bidder who bids for a higher position k has to pay $b_{\phi(k)}$ to win that position. The last two conditions say that no bidder who wins a position has an incentive to deviate so that he wins no position, and no bidder who wins no position has an incentive to deviate so that he wins some position.

Our approach of modeling the auction as a static game of complete information and focusing on Nash equilibria of this game follows previous papers: Lahaie (2006), Edelman et. al. (2007), and Varian (2007). The static model is very stylized. Interactions in practice take place over time. Moreover, the common knowledge assumption, literally interpreted, is unrealistic. However, the idea of our approach is that the repeated nature of the interaction with almost continuous opportunities for bid adjustment allows bidders to converge fast to a Nash equilibrium of the auction. We do not model this adjustment process explicitly. However, we have in mind that bidders behave naively in this process. Therefore, the adjustment process itself need not be in equilibrium. But after a short while, taking others' bids as given, each bidder behaves optimally. In particular, we shall assume that static equilibrium has been reached at every instance in our data set.[6]

3 Nash Equilibria

In the first part of this section we focus on a particular type of Nash equilibria, namely equilibria in which bidders do not have an incentive to win a higher position k even if they have to pay only the lower price b_{k+1}, rather than b_k. Var-

[6] Che et. al. (2011) offer some experimental evidence that an adaptive learning process under incomplete information does indeed guide bidders towards similar behavior as they exhibit in the static, complete information game.

ian (2007) has called such equilibria “symmetric Nash equilibria.” There may be multiple such equilibria, but under additional assumptions the the allocation of positions to bidders is uniquely determined. Symmetric equilibria have been studied in detail in the earlier literature. Our more general framework allows us to uncover step by step the assumptions needed for various conclusions about symmetric equilibria.

In the second part of the section we also allow asymmetric equilibria. It is difficult to characterize the set of all Nash equilibria. We present an example that indicates that the set of all Nash equilibria may be surprisingly large. We then examine critically possible arguments for focusing on symmetric Nash equilibria. We are skeptical about these arguments. In our empirical analysis we shall be agnostic in the sense that we do not make any assumption about the type of Nash equilibrium that bidders play.

Definition 2. *A vector of bids* $(b_1, b_2, \ldots, b_N)$ *is a* symmetric Nash equilibrium *if there is a compatible ranking* ϕ *of bidders so that the bid vector satisfies the conditions of Definition 1, and:*

- *For all positions* k *with* $1 \leq k \leq K$ *and all alternative positions* k' *with* $1 \leq k' < k$*:*

$$c^k_{\phi(k)}\left(v^k_{\phi(k)} - b_{\phi(k+1)}\right) \geq c^{k'}_{\phi(k)}\left(v^{k'}_{\phi(k)} - b_{\phi(k'+1)}\right)$$

- *For all ranks ℓ with $\ell \geq K+1$ and all positions k with $1 \leq k \leq K$:*

$$c^k_{\phi(\ell)} \left(v^k_{\phi(\ell)} - b_{\phi(k+1)} \right) \leq 0$$

The sense in which Nash equilibria that satisfy the conditions of Definition 2 are "symmetric" is that all bidders, when contemplating to bid for position k, expect to pay the same price for this position, namely $b_{\phi(k+1)}$. Thus, the vector $\left(b_{\phi(2)}, b_{\phi(3)}, \ldots, b_{\phi(K+1)}\right)$ can be interpreted as a vector of Walrasian equilibrium prices. If each bidder takes these prices as given and fixed, and picks the position that generates for him the largest surplus at these prices, then for each position there will be exactly one bidder who wants to acquire that position, provided that indifferences are resolved correctly. Thus the market for each position "clears": demand and supply are both equal to 1.

We now introduce an assumption that guarantees the existence of a symmetric Nash equilibrium.

Assumption 1. *For every bidder $i = 1, 2, \ldots, N$ and for every position $k = 2, 3, \ldots K$ the following two inequalities hold:*

$$c^{k-1}_i v^{k-1}_i > c^k_i v^k_i \qquad \textit{and} \qquad v^{k-1}_i \geq v^k_i$$

The first inequality says that the value of a higher position for bidder i is larger than value of a lower position. The second inequality says that bidder i 's willingness to bid of a higher position is at least as large as i's willingness to bid for a lower position. Even if the value per click and the impression value are larger for larger positions, the second inequality in Assumption 1 may be violated if the click rates increases too fast in comparison to the impression value. This can be seen from equation (2). Thus, the second part of Assumption 1 is somewhat restrictive. If one assumes, however, as the previous literature has done, that values per click are independent of the position, then the second inequality in Assumption 1 is automatically satisfied, and the first inequality reduces to the requirement that higher positions have higher click rates than lower positions.

Proposition 1. *Under Assumption 1 the game has at least one symmetric Nash equilibrium in pure strategies.*

Proof. STEP 1: We show the existence of Walrasian equilibrium prices for the K positions. This is essentially an implication of Theorem 3 in Milgrom (2000). Milgrom proves existence of competitive equilibrium indirectly. He postulates that K objects are sold through a *simultaneous ascending auction*, and that bidders bid *straightforwardly*. He then proves that the auction will end after a finite number of rounds, and that the final prices paid for the K objects converge to Walrasian equilibrium prices as the increment in the simultaneous ascending auction tends to zero. This implies that Walrasian equilibrium prices exist. To apply Milgrom's argument to our context, we need to modify his construction, and assume that bids in the simultaneous ascending auction are payments per click, rather than total payments. With this modification, Milgrom's argument goes through without change. Milgrom's result assumes that objects are substitutes: each bidder's demand for an object does not decrease as the prices of the other objects increase. This assumption is obviously satisfied in our setting with single unit demand.

Denote by ϕ a ranking of the bidders that is compatible with the Walrasian equilibrium, that is, in the Walrasian equilibrium position k is obtained by agent $\phi(k)$. Denote by $(p_1, p_2, \ldots, p_K)$ some vector of Walrasian equilibrium prices that has been constructed by Milgrom's method. Observe that, as one can easily show, $N = K$ implies $p_K = 0$.

STEP 2: We show that $p_1 \geq p_2 \geq \ldots \geq p_K$. Indeed, suppose that for some k we had $p_{k-1} < p_k$, and consider the bidder i who acquires position k. Because position k is the optimal choice for bidder i at the given prices,

$$c_i^k \left(v_i^k - p_k\right) \geq c_i^{k-1} \left(v_i^{k-1} - p_{k-1}\right) \tag{6}$$

Because $p_{k-1} < p_k$ this implies:

$$c_i^k \left(v_i^k - p_k\right) > c_i^{k-1} \left(v_i^{k-1} - p_k\right) \tag{7}$$

which is equivalent to:

$$\left(c_i^{k-1} - c_i^k\right) p_k > c_i^{k-1} v_i^{k-1} - c_i^k v_i^k \tag{8}$$

The expression on the right hand side of (8) is by Assumption 1 positive. The expression on the left hand side is linear in p_k. For $p_k = 0$ it equals zero and is thus smaller than the right hand side. The largest possible value of p_k is v_i^k. We now show that even for this largest value of p_k the expression on the left hand side is smaller than the expression on the right hand side:

$$\left(c_i^{k-1} - c_i^k\right) v_i^k \leq c_i^{k-1} v_i^{k-1} - c_i^k v_i^k \Leftrightarrow \tag{9}$$

$$v_i^k \leq v_i^{k-1} \tag{10}$$

which holds by Assumption 1. Thus, there is no value of p_k for which (8) could be true, and the assumption $p_{k-1} < p_k$ leads to a contradiction.

STEP 3: We now construct a symmetric Nash equilibrium. For each k with $2 \leq k \leq K$ we set the bid of the bidder who wins position k in the Walrasian equilibrium equal to the price that position $k-1$ has in that equilibrium:

$$b_{\phi(k)} = p_{k-1} \tag{11}$$

For bidder $\phi(1)$ who wins position 1 we can choose any bid $b_{\phi(1)}$ that is larger than p_1. Finally, if there are bidders i who don't obtain a position in the Walrasian equilibrium, we set their bids equal to p_K. Because the Walrasian prices are ordered as described in STEP 2 these bids imply that every bidder who wins a position in the Walrasian equilibrium wins the same position in the auction, and pays in the auction the price that he pays in the Walrasian equilibrium. Moreover, because we have implemented a Walrasian equilibrium, no bidder prefers to acquire some other position at the price that the winner of that position pays over the outcome that he obtains in the proposed bid vector, and hence we have a symmetric Nash equilibrium. □

Three remarks are in order. First, note that Proposition 1 does not assert that there is only one symmetric Nash equilibrium. In many examples, there are indeed many symmetric Nash equilibria. Second, as the second part of Assumption 1 is somewhat restrictive, one might wonder whether it can be relaxed. We have not pursued this question. Third, the simultaneous ascending auction to which we refer in Step 1 of the above proof may be regarded as a an alternative to the generalized second price auction used by Overture. We have not attempted to evaluate the relative merits of this alternative auction format for sponsored search positions.

If we knew bidders' valuations v_i^k, could we predict the winner of each position in a symmetric Nash equilibrium? We shall consider this question under the following simplifying assumption.

Assumption 2. *For every bidder $i = 1, 2, \ldots, N$ and for every position $k = 1, 2, \ldots, K$ there are numbers $a_i > 0$ and $c^k > 0$ such that*

$$c_i^k = a_i c^k$$

for all i and all k.

Proposition 2. *Under Assumption 2 a ranking ϕ of bidders that is compatible* *with a symmetric Nash equilibrium maximizes*

$$\sum_{k=1}^{K} c^k v_{\phi(k)}^k$$

among all possible rankings ϕ*.*

For generic parameters, there will be a unique allocation of positions to bidders that maximizes the sum in Proposition 2. In this sense, Proposition 2 provides conditions under which we can unambiguously predict which bidder will win which position in a symmetric equilibrium.

The function that according to Proposition 2 symmetric Nash equilibrium rankings maximize is similar to a utilitarian welfare function. However, a utilitarian welfare function would assign to each ranking the sum of all bidders' valuations of positions, that is:

$$\sum_{k=1}^{K} a_i c^k v_{\phi(k)}^k$$

In the expression in Proposition 2 the bidder specific factors a_i are omitted. It is intuitively plausible that the Overture auction cannot lead to an allocation which takes these factors into account: these factors only affect the absolute level of click rates, but not their ratio; incentives in the auction only depend on the ratio of click rates. If, however, there is no heterogeneity in click rates, i.e. the factors a_i are the same for all bidders, then Proposition 2 implies that symmetric Nash equilibria maximize utilitarian welfare.

Note that Proposition 2 is about the equilibrium allocation of positions to bidders. It does not imply the uniqueness of equilibrium bids. There may be multiple bids supporting the same equilibrium allocation. In the second part of this section we shall discuss the multiplicity of Nash equilibria in our model in more detail.

Proof. Le ϕ be a ranking of bidders that is compatible with a symmetric Nash equilibrium. Without loss of generality assume that ϕ is the identity mapping. Let $\hat{\phi}$ be an alternative ranking. Consider some position k, and consider the incentives of the bidder who wins position k under the ranking $\hat{\phi}$, that is bidder $\hat{\phi}(k)$. If we denote by p^k (for $k = 1, 2, \ldots, K$) the Walrasian prices associated with the symmetric equilibrium, then, by the definition of Walrasian equilibrium:

$$a_{\hat{\phi}(k)} c^k \left(v^k_{\hat{\phi}(k)} - p^k \right) \leq a_{\hat{\phi}(k)} c^{\hat{\phi}(k)} \left(v^{\hat{\phi}(k)}_{\hat{\phi}(k)} - p^{\hat{\phi}(k)} \right) \Leftrightarrow \quad (12)$$

$$c^k \left(v^k_{\hat{\phi}(k)} - p^k \right) \leq c^{\hat{\phi}(k)} \left(v^{\hat{\phi}(k)}_{\hat{\phi}(k)} - p^{\hat{\phi}(k)} \right) \quad (13)$$

If $\hat{\phi}(k) > K$:

$$a_{\hat{\phi}(k)} c^k \left(v^k_{\hat{\phi}(k)} - p^k \right) \leq 0 \Leftrightarrow \quad (14)$$

$$c^k \left(v^k_{\hat{\phi}(k)} - p^k \right) \leq 0 \quad (15)$$

Summing (13) and (15) over all $k = 1, 2, \ldots, K$ we obtain:

$$\sum_{k=1}^{K} c^k \left(v^k_{\hat{\phi}(k)} - p^k \right) \leq \sum_{k \in \{1,\ldots,K | \hat{\phi}(k) \leq K\}} c^{\hat{\phi}(k)} \left(v^{\hat{\phi}(k)}_{\hat{\phi}(k)} - p^{\hat{\phi}(k)} \right) \quad (16)$$

which implies:

$$\sum_{k=1}^{K} c^k \left(v^k_{\hat{\phi}(k)} - p^k \right) \leq \sum_{k=1}^{K} c^k \left(v^k_k - p^k \right) \Leftrightarrow \tag{17}$$

$$\sum_{k=1}^{K} c^k v^k_{\hat{\phi}(k)} \leq \sum_{k=1}^{K} c^k v^k_k \tag{18}$$

Thus, the value of the function in Proposition 2 under $\hat{\phi}$ is not larger than it is under ϕ. □

To illustrate how Proposition 2 allows one to predict symmetric equilibrium allocations we consider the case in which bidders are ranked according to a single crossing condition: the marginal value of higher positions decreases as a player's index goes up.

Assumption 3. *Assumption 2 holds, and for all bidders* $i = 1, 2, \ldots, N-1$ *and all positions* $k = 1, 2, 3, \ldots, K-1$

$$c^k v_i^k - c^{k+1} v_i^{k+1} > c^k v_{i+1}^k - c^{k+1} v_{i+1}^{k+1}$$

The following is an immediate implication of Proposition 2.

Corollary 1. *Under Assumption 3 in every symmetric Nash equilibrium bidder* i *wins position* i *for* $i = 1, 2, \ldots, K$.

If Assumptions 1 and 3 hold simultaneously we can infer the existence of a symmetric equilibrium in which bidder i wins position i. Existence results that have been obtained constructively by Edelman et. al. (2007, Theorem 1) and Varian (2007, Section 2) are implications of this observation. These authors study models in which Assumption 2 holds, values v_i^k are independent of position k, and $c^k > c^{k+1}$ for $k = 1, 2, \ldots, K-1$. This implies Assumption 1. Assumption 3 is then satisfied if, in addition, bidders are labeled such that $v_1 > v_2 \ldots > v_N$.

Varian (2007, Fact 5) notes that in his model a sufficient condition for a bid vector to be a symmetric Nash equilibrium is that no bidder has an incentive to bid for an adjacent position instead of the position he obtains, that is, if the inequalities in Definition 2 hold for adjacent positions, they hold for all positions. The same can be shown in our model under Assumption 3.

We now turn to an analysis that also allows asymmetric Nash equilibria. It is hard to give a complete description of *all* Nash equilibria. We give instead an example. In this example, every allocation of positions to bidders can be an equilibrium allocation. We describe corresponding bid vectors. Note that the example satisfies Assumption 3, and thus in any symmetric equilibrium bidders are allocated positions as in Corollary 1.

Example 1. *There are 3 bidders and 3 positions. Click rates are bidder independent:* $c_i^1 = 3, c_i^2 = 2, c_i^3 = 1$ *for all bidders* $i = 1, 2, 3$. *The willingness to bid per click is independent of a bidder's position:* $v_1^k = 16, v_2^k = 15, v_3^k = 14$ *for all positions* $k = 1, 2, 3$. *Whenever one bidder bids 11, another bids 9, and another bidder bids 7, then this will be a Nash equilibrium. Thus, all allocations of positions to bidders are possible equilibrium allocations.*

Although we do not offer in this paper a complete characterization of asymmetric Nash equilibria, our empirical analysis will not rule such equilibria out by assumption. Varian (2007) offers no game theoretic motivation for focusing on symmetric Nash equilibria. But Edelman et. al. (2007, p. 249) have argued that the selection of symmetric equilibria can be justified by the assumption that bidders raise their bids to induce a higher payment for the next highest bidder, but that they do so only up to the point $\bar{b}$ at which they would not regret having raised their bid if the next highest bidder were to lower his bid slightly below $\bar{b}$. Edelman et. al. refer to the selected equilibria as " locally envy-free." We regard this argument as not entirely compelling because it is not clear that the relevant case for bidders to consider is the case that other bidders lower their bids just below $\bar{b}$.

Edelman et. al. (2007) offer two further justifications for their selection. The first (their footnote 17) is that there is an analogy between symmetric Nash equilibria and the requirement in single unit, second price auctions that bidders bid at least their true value. We argue that in single unit, second price auctions this requirement is not attractive per se. Edelman et. al. (2007, Section IV) also introduce an ascending price auction with incomplete information, and show that the unique perfect Bayesian equilibrium of this auction results in rankings and payments identical to those in symmetric Nash equilibria of the static, complete information model. They interpret the ascending price auction as a description of the process by which bidders arrive at equilibrium. One can conceive of other models of this process, and we prefer to remain agnostic on this point.[7]

[7] An earlier version of this paper also discussed refining the set of Nash equilibria by ruling out weakly dominated strategies. The strategies in Example 1 are not weakly dominated.

4 Data

We have collected bid data for five search terms over a period from February 3rd 2004 to May 31, 2004. The search terms are *Broadband, Flower, Loan, Outsourcing* and *Refinance*.[8] For each search term, the data describe the current bid levels every 15 minutes[9] yielding 96 bid observations per bidder per day.[10] We include a bid observation (and time period) for bidder i in the final data only when the bidder places a new bid or alters the bid level of an existing bid. The data selection avoids a set of issues related to delays in bidders' response times.[11]

We augmented the bid data with weekly click-through data for 46 weeks in 2004.[12] Based on the click through data we calculate that the ratio c^{k-1}/c^k equals about 1.5 for top positions on average across our search terms. We use this number in the subsequent analysis. The assumption of a common click through ratio is restrictive as it does not permit the possibility of bidder heterogeneity in click through ratios. We make the assumption as our data do not contain information

[8] Initially, search words were chosen at random by using an English dictionary, and we collected one sample of bid prices for each search word. We then selected the search words that achieve high bid prices. The motivation for our selection was that bidders may be more likely to behave optimally when more money is at stake.

[9] The data were collected using the publicly accessible bid tool on the webpage `http://uv.bidtool.overture.com/d/search/tools/bidtool`. The data retrieval time interval ranges between 10 and 20 minutes.

[10] Bidders revise their bids frequently and the 15 minute sampling frequency was chosen to capture bid changes accurately. On average a new bid is chosen, or an existing bid is revised every 43 minutes, yielding an average of 33 changes per day. There is variation across search terms with the average number of bid revisions ranging from five per day for *Outsourcing* to 63 per day for *Flower*.

[11] In particular, the data selection avoids the concern that an initially payoff maximizing bid may no longer be an optimal bid choice when an opponent's bid level changes.

[12] The data were kindly provided to us by Yahoo.

on bidder specific click throughs. The empirical findings have to be interpreted subject to this caveat.

The imposed ratio of 1.5 is based on the expected click through rates. The weekly number of clicks per position exhibits a lot of variation. For this reason we use the median number of clicks per position to measure the expected click through rates. We denote the median number of clicks by $\overline{c}^k(\ell)$ for search term ℓ and position k. The expected click ratios are calculated as the ratio of median number of clicks across search terms, that is $\frac{1}{5}\sum_{\ell=1}^{5}(\overline{c}^{k-1}(\ell)/\overline{c}^k(\ell))$, and equal $1.44, 1.55, 1.50$ and 1.28 for positions $k = 2, 3, 4, 5$ respectively. The standard deviations of the click ratios, calculated across search terms, equal $0.20, 0.12, 0.29$ and 0.20 for positions $k = 2, 3, 4, 5$, respectively. We cannot reject the null hypothesis that the ratio equals 1.5 for top positions.

The price paid reflects a lower bound on an advertiser's willingness to pay per click. The lower bound varies substantially across categories. The price for the top *Broadband* position equals \$2.05 on average. The average top position price equals \$2.44, \$4.62, \$2.54, \$6.92 for the search terms *Flower*, *Loan*, *Outsourcing*, and *Refinance* respectively.

The price difference between two adjacent positions is 20 cents on average across search terms for the top ten positions. The price difference between two adjacent positions varies across search terms and ranges from 14 cents for *Outsourcing* to 31 cents for *Refinance*.

There is substantial dispersion in bids over time suggesting that revealed preference arguments may achieve tight bounds on advertisers' willingness to pay.

The bid dispersion varies in magnitude across categories. The low standard deviation occurs for *Outsourcing* with a standard deviation of the top position price equalling \$0.27. On the other extreme is the category *Broadband* with a standard deviation of \$0.81. The empirical distribution reveals that ninety percent of high *Outsourcing* position price observations fall into the interval \$2.00 to \$3.00. Ninety percent of *Broadband* price observations fall into the interval \$1.32 to \$3.25.

The dispersion in bids over time does not reflect a clear pattern attributable to days of the week or hours of the day. We illustrate the patterns of bids over time in two ways: First, we qualitatively examine position bids and illustrate that there is no systematic pattern of variation in position bids across days of week and hour of the day. Second, we conduct statistical tests to confirm that there is no significant variation across days of the week or hours of the day for the majority of search terms and positions.

Table 1 reports the average submitted bid for the top three positions for each day of the week for the search term *Outsourcing*. The evidence for other position bids and other search terms is qualitatively similar. The observations in the Tables include all submitted bids b_k for position k for a given day. The number in parenthesis denotes the sample standard deviation.

The top position bid b_1 is surprisingly constant in Table 1. It varies by 15 cents across days of the week ranging between 2.58 (Thursday) and 2.73 (Friday). There is no observation for b_1 on Sundays. A statistical test of equality of average top position bid b_1 across days of the week cannot be rejected at the one percent

Table 1. Outsourcing Position Bids.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
b_1	2.62	2.70	2.60	2.58	2.73	2.71	.
	(0.25)	(0.24)	(0.22)	(0.23)	(0.25)	(0.28)	
b_2	2.51	2.50	2.54	2.33	2.42	2.14	2.59
	(0.21)	(0.15)	(0.18)	(0.36)	(0.31)	(0.35)	(0.09)
b_3	2.08	2.06	2.25	1.98	2.37	2.38	2.21
	(0.46)	(0.41)	(0.41)	(0.37)	(0.27)	(0.00)	(0.13)

significance level. For comparison on average the position bid b_2 is 25 cents lower than the position bid b_1 and the difference in the top two bid prices, $b_1 - b_2$, is of larger magnitude than the within day of the week variation in b_1. The average outsourcing bid b_2 ranges between 2.14 (Saturday) and 2.59 (Sunday). The Saturday low bid price b_2 is a substantial 19 cents away from the second lowest bid price (Thursday). A statistical test of equality of average top position bid b_2 across days of the week is rejected. The position bid b_3 ranges between 1.98 (Thursday) and 2.38 (Saturday). A statistical test of equality of average top position bid b_3 across days of the week cannot be rejected.

Table 1 illustrates that there is variation in position bid prices across the days of the week, but there is no clear day of the week pattern evident. For example, on Saturday the low price for position bid b_2 occurs while at the same day the high position bid b_3 arises. The rejection of the null of equal coefficients for position bids b_2 across days of the week appears attributable to Saturday outliers.

Next, we examine the evidence on hourly position prices. Table 2 illustrates the within day hourly price variation for *Refinance*.

Table 2. Refinance, Hourly Position Bid Prices.

Hour	b_1		b_2		b_3	
0:00	7.02	(0.57)	6.95	(0.48)	6.96	(0.53)
1:00	7.03	(0.75)	7.03	(0.51)	6.63	(0.39)
2:00	7.13	(1.07)	6.94	(0.49)	6.82	(0.47)
3:00	6.86	(0.61)	6.85	(0.61)	6.73	(0.63)
4:00	7.30	(0.95)	7.11	(1.04)	6.85	(0.79)
5:00	6.87	(0.61)	6.90	(0.76)	6.77	(0.65)
6:00	7.05	(0.90)	7.05	(0.76)	7.06	(0.62)
7:00	6.83	(0.55)	6.72	(0.54)	6.70	(0.64)
8:00	6.98	(0.57)	6.94	(0.72)	6.83	(0.38)
9:00	6.85	(0.61)	6.91	(0.57)	7.16	(0.53)
10:00	6.86	(0.62)	6.83	(0.59)	6.86	(0.58)
11:00	6.92	(0.71)	6.84	(0.52)	6.87	(0.45)
12:00	6.90	(0.57)	6.86	(0.53)	6.93	(0.52)
13:00	6.88	(0.57)	6.79	(0.50)	6.66	(0.57)
14:00	6.94	(0.61)	6.98	(0.60)	6.84	(0.60)
15:00	6.96	(0.54)	6.76	(0.58)	6.55	(0.60)
16:00	6.86	(0.67)	6.73	(0.66)	6.56	(0.58)
17:00	7.01	(0.62)	6.89	(0.57)	6.86	(0.67)
18:00	7.04	(0.49)	7.00	(0.43)	6.94	(0.57)
19:00	7.82	(5.44)	6.99	(0.65)	6.76	(0.60)
20:00	7.02	(0.47)	6.93	(0.49)	6.71	(0.55)
21:00	7.04	(0.46)	6.83	(0.52)	6.69	(0.52)
22:00	7.02	(0.51)	7.05	(0.56)	6.93	(0.55)
23:00	7.04	(0.64)	7.01	(0.58)	7.09	(0.54)

There is again surprisingly little variation in bid prices within a day. The top position bid price b_1 ranges between 6.83 (at 7am) and 7.82 (at 7pm). There is no discernible trend during the day in position bid levels. A statistical test of equal position bids b_1 within the day cannot be reject. The position bid b_2 ranges between 6.72 (7am) and 7.11 (4am). The position bid b_3 ranges between 6.55 (3pm) and 7.16 (9am). For both, b_2 and b_3, a statistical test of equal average bid prices across hours of the day cannot be rejecteded. Overall, we do not see a systematic discernible pattern in variations of position prices during the day.

To examine systematically whether position prices vary significantly across days of the week or hours of the day we consider a statistical test of equality of day of the week (or hour) effects. We examine the null hypothesis of constant position prices for individual search words and position bids. The null hypothesis for search term ℓ and position k can be written as $b_k^j(\ell) = b_k(\ell)$ for all days of the week j (or all hourly intervals j). We conduct separate tests for all top five position bids and all search terms. In each case we consider a test for days of the week and another test for hours within the day. We construct an F-statistic by regressing bid position prices on two sets of regressors. First, in the unrestricted case we include a full set of day of the week dummies (or hourly interval dummies). Second, in the restricted case we include a constant only. We use the 1% significance level as the test criterion.

In total we conduct 50 tests. On 28 of 50 tests, that is on 56% of all tests, we cannot reject the null of equal coefficients. For the majority of tests we find no significant variation over time. Excluding the search term loan, the number of cases in which the null is not rejected increases to 70%. We may interpret the statistical test evidence as an indication that there are no systematic variations in bid prices across days of the week or hours of the day.

Summarizing, we can conclude that a qualitative examination of position bids reveals no systematic patterns of variation in position bids across days of week and hours of the day. There is some variation in prices over time possibly caused by individual outliers, but there is no clear evidence of systematic variation. A statistical examination of equal position bids over time reveals no significant variation across days of the week and hours of the day for the majority of search terms and positions.

In the data we see that some bidders are regular bidders for top positions while other bidders achieve a top position on occasions only, or vanish after a short time. These two types of bidders may exhibit distinct valuation processes and we wish to distinguish them in the subsequent analysis. To illustrate the difference we determine the average position in the bid ranking during our sample period. There are 167 bidders with average ranking of one to ten and there are $1,227$ bidders with average ranking of ten or higher. The bidders with average ranking of one to ten win 85 percent of the top five positions. We focus on these regular bidders in the subsequent analysis.

5 Revealed Preferences

This section explores a non-parametric revealed-preference approach to infer bounds on advertisers' willingness to pay. We assume that the submitted bid maximizes the bidder's payoff. We use the bid data in conjunction with the optimality condition to deduce bounds on the willingness to pay. We illustrate when the bounds imply a non-empty set of valuations and examine the non-emptiness hypothesis empirically. We discuss the shape of the valuation profiles consistent with the bounds. Section 5.1 illustrates when a set of bid observations yields a non-empty set of valuations. Section 5.2 describes our empirical test results for the revealed preference hypothesis. Section 5.3 illustrates the shape of the valuation profile consistent with the bounds and describes caveats of the test procedure.

5.1 Test of the Revealed Preference Hypothesis

It is instructive to distinguish two types of bid submissions depending on whether the submitted bid wins an item or not. First, suppose the chosen bid of bidder i does not win a position which we call a *type one* bid submission. If we denote by $b_{\phi(k)}$ the kth highest bid, then, it must be that the bid prices exceed the valuation of the position:

$$v_i^k \leq b_{\phi(k)} \quad \text{for all } k \leq K \tag{19}$$

Thus, we obtain an upper bound on the valuation vector.

Second, suppose the bid by bidder i wins position $k \leq K$. We call this a *type two* submission. Optimality of the bid choice implies the following three inequalities:

$$-v_i^k \leq -b_{\phi(k+1)} \tag{20}$$

$$v_i^{k'} \leq \frac{c_i^k}{c_i^{k'}} v_i^k + \left[b_{\phi(k')} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)} \right] \quad \text{for } k' < k \tag{21}$$

$$v_i^{k'} \leq \frac{c_i^k}{c_i^{k'}} v_i^k + \left[b_{\phi(k'+1)} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)} \right] \quad \text{for } K \geq k' > k \tag{22}$$

The first inequality says that the valuation of position k is at least as large as the winning price which places a lower bound on the valuation v_i^k. The second and third inequalities say that the valuation for a position that is not won, $v_i^{k'}$ with $k' \neq k$, is bounded from above by a line with slope $\frac{c_i^k}{c_i^{k'}}$ and an intercept equal to $b_{\phi(k')} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)}$ for $k' < k$ and an intercept equal to $b_{\phi(k'+1)} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)}$ for $k' > k$, respectively.

The above inequalities are the same as in Definition 1. We can write these inequalities compactly in matrix notation as

$$\mathbf{A}_t \mathbf{v_i} \leq \alpha_t \tag{23}$$

where $\mathbf{v_i} = \left(v_i^1, v_i^2, \ldots, v_i^K\right)$ is a $K \times 1$ dimensional valuation vector; $\mathbf{A}_t$ is a $K \times K$ dimensional matrix and α_t is a $K \times 1$ dimensional vector. In type one submissions $\mathbf{A}_t$ equals the identity matrix and α_t is equal to $\left(b_{\phi(1)}, b_{\phi(2)}, \ldots, b_{(K)}\right)$. In type two submissions, when position k is won, $\mathbf{A}_t$ is equal to a matrix with entry (k, k) equal to -1, entry (k, k') equal to 0, entry (k', k') for $k' \neq k$ equal to 1, entry (k', k) equal to $-\left(c_i^k/c_i^{k'}\right)$ and all other entries equal to zero;[13] and vector α_t has entry k equal to $-b_{\phi(k+1)}$, entries k' where $k' < k$ equal to $b_{\phi(k')} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)}$, and entries $k' > k$ equal to $b_{\phi(k'+1)} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)}$.

Given a set of observations T, we denote the set of valuations that satisfy restriction (23) as $\mathbf{V_i^T}$,

$$\mathbf{V_i^T} = \left\{\mathbf{v_i} \in \Re_+^K | \mathbf{A_t v_i} \le \alpha_\mathbf{t} \text{ for all } t \in T\right\}$$

Revealed preference predicts that the set $\mathbf{V_i^T}$ is non-empty. The revealed preference hypothesis can be tested empirically. Observe though that the computational complexity of the empirical test can be high even for moderately sized K due to the curse of dimensionality.

Figure 1 illustrates the set $\mathbf{V_i^T}$ graphically in the case of two positions, $K = 2$. In Figure 1, we write " b_2^1" for $b_{\phi(2)}^1$ etc. We consider three hypothetical bid vector observations b^1, b^2, b^3 where the superscript in the bid vector indicates that bidder i wins position 1, position 2, or no position, respectively. The dark shaded area[14] describes the set of valuations consistent with inequalities (19) - (22) for the hypothetical bid vectors. The solid line segments b_2^1, a_5 and a_7 arise from inequalities (20) and (22) and bid vector b^1. The inequalities imply that valuations in the area south-east of the solid line segments are consistent with item 1 being won.[15] The dashed line segments b_3^3, a_3 and a_8 arise from inequalities (20) and

[13] Here, $k' \neq k$.

[14] The dark shaded area has boundary points a_1, a_2, a_3, a_4, a_5, and a_6.

[15] Valuations to the east of the solid line segment b_2^1, a_5 satisfy the property that the valuation v_1 exceeds the price paid for item 1. Valuations to the south-east of the line segment a_7, a_9 satisfy

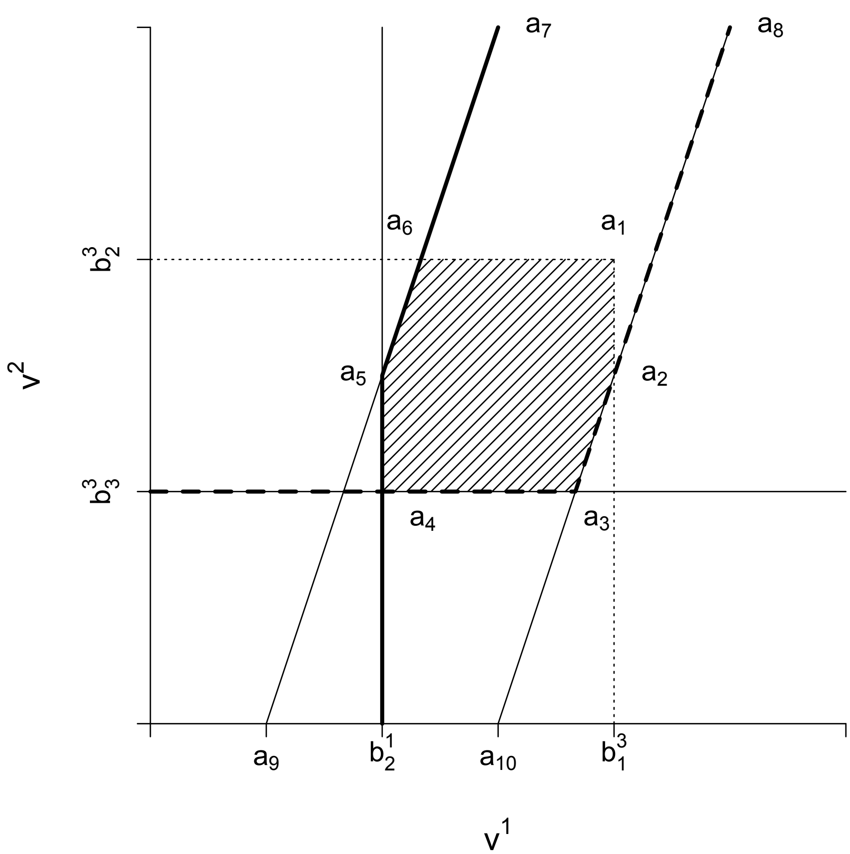

Figure 1: Valuations consistent with hypothetical bids

(21) and bid vector b^2. It says that valuations in the area north-west of the solid line segments are consistent with item 2 being won.[16] The dotted line segments b_2^3, a_1 and b_1^3 arise from inequality (19) and bid vector b^3. It says that valuations in the south-west of the dotted line segments are consistent with no item being won.[17]

Figure 1 can be easily extended to an arbitrary set of bids. To see that, partition the set of observations T into three sets T^1, T^2, T^3, so that T^1, T^2 denote the sets of bids in which position $1, 2$ is won and T^3 denotes the set of bids in which no position is won. The dotted line is defined by the minimum bids for positions 1 and 2, $b^3_{\phi(2)} = \min_{t \in T^3}(b^t_{\phi(2)})$, and $b^3_{\phi(1)} = \min_{t \in T^3}(b^t_{\phi(1)})$, the dashed line segments are defined by $b^2_{\phi(3)} = \max_{t \in T^2}(b^t_{\phi(3)})$ and $a_{10} = \min_{t \in T^2}(b^t_{\phi(1)} - (c_i^2/c_i^1)b^t_{\phi(3)})$, and the solid line segments are defined by $b^1_{\phi(1)} = \max_{t \in T^1}(b^t_{\phi(2)})$ and $a_9 = \max_{t \in T^1}(b^t_{\phi(2)} - (c_i^2/c_i^1)b^t_{\phi(3)})$. Hence, the bid vectors b^1, b^2, b^3 in Figure 1 denote the corresponding minima and maxima. If some set T^i is empty, then the corresponding boundary will not bind and the shaded area in the figure will be enlarged.[18]

With multiple positions, $K > 2$, the set $\mathbf{V_i^T}$ is contained in $\Re^K$. The boundary

the property that the bidder who buys item 1 would not have been better off by buying item 2.

[16]Valuations to the north of the solid line segment b_3^3, a_3 satisfy the property that the valuation v_2 exceeds the price paid for the item b_3^3. Valuations to the north-west of the line segment a_{10}, a_8 satisfy the property that the bidder who buys item 2 would not have been better off by buying item 1.

[17]Valuations to the west of the dotted line segment b_1^3, a_1 satisfy the property that the valuation v_1 is less than the price paid for item 1. Valuations to the south of the line segment b_2^3, a_1 satisfy the property that the valuation v_2 is less than the price paid for item 2.

[18]If T^1 is empty, then the left boundary of the shaded area will equal the vertical line $\left(0, v_i^2\right)$ as by assumption $v_i^1 > 0$. If T^2 is empty, then the bottom boundary of the shaded area will equal the horizontal line $\left(v_i^1, 0\right)$. If T^3 is empty, then the shaded area is unbounded to the north-east.

of the set $\mathbf{V}_i^T$ along dimension $(v_i^k, v_i^{k'})$ shares the features as in Figure 1 for any pair $(v_i^k, v_i^{k'})$.

Next, we state that a pairwise non-empty boundary is a necessary condition for the revealed preference hypothesis. We denote the set of bid observations in which the submitted bid wins position k by $T^k \subset \Re^N$, and the set of bid observations in which the submitted bid does not win any position by $T^{K+1} \subset \Re^N$. We adopt the convention that the maximum and minimum over an empty set equals $-\infty$ and $+\infty$, respectively.

Condition 1 (Non-empty Pairwise Boundaries)**.** *Given a set of observations T, a necessary condition for the valuation range $\mathbf{V_i^T}$ to be non-empty is that*

$$\max_{t \in T^k} \left(b^t_{\phi(k+1)}\right) \leq \min_{t \in T^{K+1}} (b^t_{\phi(k)}) \textit{ for all } k \leq K;$$

$$\max_{t \in T^k} \left(b^t_{\phi(k+1)} - \frac{c_i^{k'}}{c_i^k} b^t_{\phi(k'+1)} \right) \leq \min_{t \in T^{k'}} \left(b^t_{\phi(k)} - \frac{c_i^{k'}}{c_i^k} b^t_{\phi(k'+1)} \right)$$
$$\textit{for all } k, k' \leq K \textit{ with } k < k'.$$

The non-empty pairwise boundary condition is a necessary condition for a non-emptiness of the set $\mathbf{V_i^T}$. The first necessary condition states that the position price paid during some period cannot exceed the price of the same position during another period when the bidder doesn't win a position. The second necessary condition says that when position k is won the valuation difference, $v_i^k - \left(c_i^{k'}/c_i^k\right) v_i^{k'}$, is bounded from below by the price differences $b^t_{\phi(k+1)} - \left(c_i^{k'}/c_i^k\right) b^t_{\phi(k'+1)}$, and, when position k' is won it is bounded from above by the price differences $b^t_{\phi(k)} - \left(c_i^{k'}/c_i^k\right) b^t_{\phi(k'+1)}$, respectively. Observe that Condition 1 is not a sufficient condition as two two-dimensional areas that share one dimension need not overlap in the common dimension.

Examining empirically whether the set $\mathbf{V_i^T} \subset \Re^K$ is non-empty can be computationally complex for moderately sized K. Yet, Condition 1 can be examined at relatively small computational costs for all K. For computational reasons we proceed with a two step test approach of the revealed preference hypothesis: In the first step, we examine whether there is a violation of Condition 1. In the second step, we examine whether there is a non-empty set for those observations with a non-empty pairwise boundary.

A violation of the revealed preference hypothesis may be indicative of behavior inconsistent with rationality. Alternatively, it may suggest taste changes across subsets of the observations. For instance, preferences may be different during day-time from during night-time. The revealed preference hypothesis may be satisfied during day-time periods and during night-time periods, but not for both periods jointly.

5.2 Revealed Preference Test Results

This section reports the revealed preference hypothesis test results in two steps. First, we examine the non-empty pairwise boundary hypothesis. Second, we consider the full revealed preference test.

The *non-empty pairwise boundary hypothesis* is examined for a subset of our data consisting of bidders that submit a bid for a top five position on average.[19] In total there are 71 such bidders. We find no violation of the non-empty pairwise boundary condition for 21 of 71 bidders, or 30 percent. Violations arise for bidders submitting numerous bids. On average, a bidder with a violation submits 154 bids. In contrast, a bidder without a violation submits about 3 bids.

A violation may be attributable to a discrete change in an observable characteristic, such as a change from day-time to night-time. Alternatively, a violation may be attributable to a gradual change in observable characteristics, for instance when there is a time trend. Violations may also arise, if bidders are inexperienced and make periodic mistakes in assessing their willingness to pay or in submitting erroneous bids.

To examine whether violations arise suddenly or gradually, we select all bidders with a violation for the entire sample period. We determine the (maximal) length of sub-periods on which the non-empty boundary hypothesis holds. The algorithm is simple. For each bidder, we start with the first observation and then add on additional consecutive observations as long as no violation of the non-empty boundary hypothesis occurs. When a violation arises, we start a new set of observations. The algorithm partitions the set of observations into consecutive sub-period $T_{i1}, \ldots, T_{it_i}$ with the property that the non-empty boundary hypothesis is satisfied on each sub-period. Notice that period T_{i1} starts at the point of time

[19] An examination of all bidders shows that a violation of the non-empty boundary condition occurs for 14 percent of bidders only. The low violation rate may appear surprising initially. However, the bidders without a violation win position 70 or higher on average. For these bidders, the upper valuation bound is binding most of the time, and there are hardly any observations that provide a lower bound on the valuation range.

when bidder i places the first bid, or revises the existing bid for the first time. Typically period T_{i1} starts well inside of our sample period.

The length of the sub-periods without a violation amounts to 1.34 days on average. During the 1.34 days the bidder submits a total of 4.7 bids on average. The frequent violations suggest that valuations may vary over time, or that bidders may make mistakes periodically.

Next, we describe our test results of the revealed preference hypothesis. We examine whether the hypothesis holds for observations without a violation of the non-empty pairwise boundary condition.

The non-empty $\mathbf{V_i^T}$ *hypothesis.* In total we include 1618 observations. These include all observations of bidders with a non-empty pairwise boundary during the entire period and all observations with a non-empty pairwise boundary for sub-periods. To limit the computational complexity of the exercise, we examine the non-emptiness hypothesis for a five dimensional valuation profile consisting of the top five valuations $(v_i^1, v_i^2, \ldots, v_i^5)$. We do not examine the restrictions placed by the hypothesis for higher position valuations, $(v_i^6, v_i^7, \ldots, v_i^{10})$. For each test candidate, we take one million independently and identically distributed multivariate random draws from a uniform distribution.[20]

The results are the following: For 51 percent of observations the set $\mathbf{V_i^T}$ is non-empty. We can conclude that for about half the observations the revealed

[20]The support of the uniform distribution is defined by the position price when no item is won, and the price paid when the item is won. Specifically, we take as the upper bound for valuation v_i^k the low bid observation that does not win a top ten position, $\min_{t \in T^{11}} b_{(i)}^t$, and we take as the lower bound the price paid when position k is won, $\max_{t \in T^k} b_{\phi(k+1)}^t$. When the upper bound does not exist, we replace it with 15. When the lower bound does not exist, we set it to 0.

preference hypothesis is satisfied.

Summarizing, we find that bidders' behavior can only be rationalized over relatively short time periods, after which we have to postulate an unexplained structural break in preferences. Next, we examine the short time periods in more detail and test further restrictions on the valuation profiles. We examine the shape of the valuation profiles that are consistent with revealed preference.

5.3 Shape of the Valuation Profile

This section examines the shape of the valuation profile and describes caveats of our test procedure.

For the shape of the valuation profile we consider two alternative hypothesis: (i) constant valuations, $v_i^1 = v_i^2 = \ldots = v_i^5$; and (ii) monotone decreasing valuations, $v_i^1 > v_i^2 > \ldots > v_i^5$. The data include all observations that pass the revealed preference test.

The hypothesis of a constant valuation profile is tested in the following way. We fix a grid with 0.5 cent increment and determine whether there exists a constant valuation profile $\widetilde{v}_i \in \{0.005, 0.01, \ldots, 15\}$ such that $\widetilde{v}_i \in \mathbf{V_i^T}$.

The hypothesis of monotone decreasing valuations is tested by using a sample of randomly drawn monotone valuation profiles. We select one hundred thousand draws from a multi-variate uniform distribution and we check whether $\widetilde{v}_i \in \mathbf{V_i^T}$.

We find that 16 percent of observations pass the constant valuation test. We interpret the test result as a rejection of the null hypothesis of constant valuations.

We find that 98 percent of observations pass the monotone decreasing valuation test. We cannot reject the monotonicity of valuation profiles.

To examine whether the decrease amounts to at least five percent for all consecutive pairs of valuations we consider the hypothesis that $v_i^k > 1.05 \cdot v_i^{k+1}$ for $k = 1, ..., 4$. We cannot reject the null hypothesis of a five percent decline for all consecutive pairs for 98 percent of observations. The test results indicate that the willingness to pay decreases with the position.

We conclude this section with two caveats of the revealed preference approach. First, the chosen data partition may influence the interpretation of the test results. For example, it may be of interest to partition the data into day-time and night-time observations, and to examine whether the revealed preference hypothesis holds for the respective sub-samples. Yet, it is difficult to determine whether the newly created partition improves the fit simply due to the increased fineness of the partition, or indeed reflects a structural break.

A second caveat concerns the zero-one nature of the revealed preference test. A test may be rejected although the data almost satisfy optimality. The test does not take into account if a violation was a near miss or far off. Both count as a violation. An alternative approach which avoids this zero-one nature is to measure the magnitude of the departure from the revealed preference.

The alternative approach may be based on the assumption that there is a random component entering the willingness to pay of individual bidders. The valuations consist then of a parametric component plus an error. The error can be interpreted as optimization error, as in McKelvey and Palfrey (1995), or may reflect random components in valuations which are known to bidders but not observed by the econometrician. Maximum likelihood may then be used to estimate the parameters of the parametric component.

Assuming a linear specification, the valuation of bidder i for position k in period t is then given by,

$$\begin{aligned} v_i^k &= \mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \varepsilon_i^{kt} \\ &= \alpha_i^0 + \left(\alpha^1 \cdot \alpha_i^0\right) \cdot k + \varepsilon_i^{kt} \end{aligned}$$

where the coefficients α_i^0 measure bidder fixed effects. The coefficient α^1 enters the multiplicative term $(\alpha^1 \cdot \alpha_i^0)$ and measures the valuation decrease relative to the bidder specific intercept α_i^0. The top position has index k equal to one. The term ε_i^{kt} denotes the error term. The parametric assumption can be combined with the bounds on the valuations to obtain a set of inequalities for any bid observation b^t.[21] Assuming that the error ε_i^{kt} is iid standard normally distributed, allows us to

[21] For a type one submissions, when the bid does not win a top position and $t \in T_{K+1}$, the inequality is,

$$\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \varepsilon_i^k \le b_{\phi(k)}^t \quad \text{for all } k \le K \quad \text{for } k \le K.$$

For type two submissions, when the submitted bid wins position $k \le K$ and $t \in T^k$, the inequalities are

$$\begin{aligned} &\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \varepsilon_i^k \ge b_{\phi(k+1)}^t \\ &\mathbf{X_i^{k',t}}\alpha_\mathbf{i} + \varepsilon_i^{k'} \le \frac{c_i^k}{c_i^{k'}}\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \frac{c_i^k}{c_i^{k'}}\varepsilon_i^k + b_{\phi(k')}^t - \frac{c_i^k}{c_i^{k'}}b_{\phi(k+1)}^t \\ &\qquad \text{for } k' < k \\ &\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \varepsilon_i^{k'} \le \frac{c_i^k}{c_i^{k'}}\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \frac{c_i^k}{c_i^{k'}}\varepsilon_i^k + b_{\phi(k'+1)}^t - \frac{c_i^k}{c_i^{k'}}b_{\phi(k+1)}^t \\ &\qquad \text{for } K \ge k' > k \end{aligned}$$

derive the likelihood.[22]

We wish to explore this likelihood approach in future work. As an initial attempt we considered data for top three premium bidders who submit at least ten bids and who occupy a top ten position for more than two weeks during the sample period. The estimates yielded slope coefficients α^1 which are negative for all search terms and significantly different from zero. The slope coefficients range between -0.02 and -0.24 across search terms. These estimates indicate that values per click depend on the position in which their advertisement is placed and accord well with the evidence from the revealed preference tests. Further research will explore the likelihood approach in more detail allowing for richer data and more realistic specifications of the slope coefficients accounting for bidder heterogeneity and additional explanatory variables.

[22]The log-likelihood is given by

$$\ell = \sum_{t \in T^{K+1}} \sum_{k=1}^{K} \ln\left(\Phi\left(b^t_{\phi(k)} - \mathbf{X_i^{k,t}}\alpha_\mathbf{i}\right)\right) +$$

$$\sum_{k=1}^{K} \sum_{t \in T^k} \ln\left(\int_{b^t_{\phi(k+1)} - \mathbf{X_i^{k,t}}\alpha_\mathbf{i}}^{\infty} \left[\prod_{k'<k} \Phi\left(\frac{c_i^k}{c_i^{k'}}\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \frac{c_i^k}{c_i^{k'}}\varepsilon_i^k + b^t_{\phi(k')} - \frac{c_i^k}{c_i^{k'}} b^t_{\phi(k+1)} - \mathbf{X_i^{k',t}}\alpha_\mathbf{i}\right)\right.\right.$$

$$\left.\left.\cdot \prod_{k'>k} \Phi\left(\frac{c_i^k}{c_i^{k'}}\mathbf{X_i^{k,t}}\alpha_\mathbf{i} + \frac{c_i^k}{c_i^{k'}}\varepsilon_i^k + b^t_{\phi(k'+1)} - \frac{c_i^k}{c_i^{k'}} b^t_{\phi(k+1)} - \mathbf{X_i^{k,t}}\alpha_\mathbf{i}\right)\phi(\varepsilon_i^k) d\varepsilon_i^k\right]\right)$$

where the first line describes the contribution to the likelihood of type one bid submissions, and lines two and three describe the likelihood contribution of type two bid submissions.

6 Conclusion

We have presented a game theoretic analysis of the Yahoo sponsored search auction, and we have interpreted bidding data assuming that this theory is a correct model of bidders' behavior. Our analysis suggests that it might be interesting to consider a dynamic model of bidding behavior in the auction in which bidders pursue repeated game strategies. Another missing element in our model might be bidders' budget constraints. It seems common that bidders in sponsored search auctions have to respect budget constraints. The rich data that high frequency sponsored search auctions provide allows the examination of a variety of further issues.

References

Athey, S. and G. Ellison (2011), Position Auctions With Consumer Search, *Quarterly Journal of Economics* 126, 1213-1270.

Athey, S. and D. Nekipelov (2011), A Structural Model of Sponsored Search Advertising Auctions, technical report, UC Berkeley.

Blume, A. and P. Heidhues (2004), All Equilibria of the Vickrey Auction, *Journal of Economic Theory* 114, 170-177.

Che, Y.-K., S. Choi and J. Kim (2011), An Experimental Study of Sponsored-Search Auctions, discussion paper, Columbia University.

Edelman, B. and M. Ostrovsky (2007), Strategic Bidder Behavior in Sponsored

Search Auctions, *Decision Support Systems* 43, 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, *American Economic Review* 97, 242-259.

Jehiel, P., and Moldovanu, B. (2006), Allocative and Informational Externalities in Auctions and Related Mechanisms, in: R. Blundell, W. Newey, and T. Persson *The Proceedings of the 9th World Congress of the Econometric Society*, Cambridge: Cambridge University Press, 2006.

Lahaie, S. (2006), An Analysis of Alternative Slot Auction Designs for Sponsored Search, in: J. Feigenbaum, J. Chuang and D. Pennock (editors), *Proceedings of the 7th ACM Conference on Electronic Commerce, June 11-15, 2006, Ann Arbor, Michigan, USA*.

McKelvey, R. and T. R. Palfrey (1995), Quantal Response Equilibria For Normal Form Games, *Games and Economic Behavior* 10, 6-38.

Milgrom, P. (2000), Putting Auction Theory to Work, *Journal of Political Economy* 108, 245-272.

Moldovanu, B. and A. Sela, The Optimal Allocation of Prizes in Contests, *American Economic Review* 91, 542-558.

Moldovanu, B., A. Sela and X. Shi (2007), Contests for Status, *Journal of Political Economy* 115, 338-363.

Varian, H. (2007), Position Auctions, *International Journal of Industrial Organization*, 25, 1163-1178.