Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence

Börgers, Tilman; Cox, Ingemar; Pesendorfer, Martin and Petricek, Vaclav
University of Michigan (Ann Arbor)

September 2008

Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence*

Tilman Börgers[†] Ingemar Cox[‡] Martin Pesendorfer[§]

Vaclav Petricek[¶]

September 2008

*We are grateful to Sébastien Lahaie, David Pennock, Isabelle Perrigne and Quang Vuong for very helpful comments, and to Di Li for superb research assistance. We have also benefited from the insights of seminar audiences at Columbia, Penn State, Michigan, Michigan State, and Vienna. 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.

[†]Department of Economics, University of Michigan, tborgers@umich.edu.

[‡]Departments of Computer Science and Electrical Engineering, University College London, ingemar@ee.ucl.ac.uk.

[§]Department of Economics, London School of Economics, m.pesendorfer@lse.ac.uk.

[¶]Department of Computer Science, University College London, petricek@acm.org.

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 its search engine. We present results that indicate that this auction has a multiplicity of Nash equilibria. We also show that weak dominance arguments do not in general select a unique Nash equilibrium. We then analyze bid data assuming that advertisers choose Nash equilibrium bids. We offer some 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 such as Google and Yahoo provide a service where users enter search terms and receive in response lists of links to pages on the World Wide Web. Search engines use sophisticated algorithms to determine which pages will be of most interest to their users. But they also offer to advertisers against payment the opportunity to advertise their pages to all users who entered specific terms. 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.

Sponsored links are an important new marketing instrument. Sponsored links offer advertisers a more targeted method of advertising than traditional forms of advertising such as television or radio commercials, because sponsored links are only shown to users who have expressed an interest in a search term that is related to the product that the advertiser seeks to sell. For companies that run search engines advertising revenue constitutes a major component of their total revenue. Google reported for the first six months of 2006 a total revenue of $4.71 billion of which $4.65 billion originated in sponsored search incomes.[1] For the same period, Yahoo reported a total revenue of $3.14 billion, of which $2.77 billion were attributed to "marketing services."[2]

[1]These figures are taken from the quarterly report filed by Google Inc. to the United States Securities and Exchange Commission on August 9, 2006. The figures are not audited. The report was accessed by the authors at: `http://investor.google.com/pdf/20060630_10-Q.pdf` on August 13, 2006.

[2]These figures are taken from the quarterly report filed by Yahoo Inc. to the United States Securities and Exchange Commission on August 4, 2006.

The major search engines use auctions to sell spaces for sponsored links. A separate auctions 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*. At the time, in 2004, at which we observed 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*.

We examine a theoretical model, and bidding data, for Overture's auction. We seek to extract information about bidders' valuation of sponsored search advertisements, and we seek to understand how bidders responded 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 Search Marketing and by Google, 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 by Yahoo Search Marketing and by Google, is a "generalized second price auction:"[3] 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;

The figures are not audited. The report was accessed by the authors at: `http://www.shareholder.com/Common/Edgar/1011006/1104659-06-51598/06-00.pdf` on August 13, 2006.

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

etc.[4]

The generalized second price auction is a method for allocating heterogeneous objects, such as positions on a page of search results, to bidders. It is based on the assumption that bidders agree which object has the highest value, which one has the second highest value, etc. The generalized second price auction is a somewhat surprising choice of auction format in the light of the recent auction literature. An example of a modern auction format that is used to allocate multiple, heterogeneous goods to bidders each of whom acquires at most one unit the simultaneous ascending auction described in Milgrom (2000). In this auction, bidders can specify in each round which object that they are bidding for. 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 recently offered theoretical analyses of the generalized second price auction that suggest that the auction may yield an efficient allocation of positions to bidders. The first part of this paper reinvestigates the theory of the generalized second price auction. We come to somewhat different conclusions than Edelman et. al. and Varian. These authors' work relies on a relatively narrow specification of bidders' payoff functions: bid-

[4]Google also uses a generalized second price format, but, when ranking advertisers and determining their payments, Google incorporates the likelihood that a user will actually click on the advertisers' link.

ders' 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 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.

We propose a more flexible specification of bidders' preferences than is used by Edelman et. al. and Varian. We undertake a more exhaustive analysis of the set of Nash equilibria. We find that existence of pure strategy Nash equilibrium can be proved quite generally. In fact, the generalized second price auction typically has many Nash equilibria. Moreover, we suggest that there are no strong theoretical reasons to expect the equilibria of the generalized price auction to be efficient.

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.

Bidding in sponsored search auctions has previously been examined empirically by Edelman and Ostrovsky (2007) and by Varian (2007). Edelman and Ostrovsky's data concern an even earlier version of the Overture auction than we consider. At the time for which Edelman and Ostrovsky have data, Overture used 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.

Other works that are related to ours include Athey and Ellison (2007), who model search behavior by consumers explicitly. The theory of sponsored search auctions also appears to be potentially 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 that is the subject of our paper seems specific to the sponsored search context.

This paper is organized as follows. Sections 2-6 describe our theoretical analysis. Section 2 presents the model. Section 3 discusses a type of Nash equilibria that Varian (2007) has called "symmetric." Section 4 analyzes "asymmetric" Nash equilibria Section 5 discusses refinements of Nash equilibria. Sections 6 and 7 constitute the empirical analysis. Section 6 describes the data. Section 7 reports the results of revealed preference tests. Section 8 concludes. One of the proofs of our theoretical results is in an Appendix.

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. The impression value is thus similar to the value that advertisers derive from other forms of advertising, such as television advertising, that are less targeted than sponsored search advertising. Impression values seem to be referred to frequently in marketing professionals' conversations.[5]

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 traditional economic considerations, but also by human physiology (where do people look first on a computer screen?) and by human psychology, and we do not know of good ways of capturing these factors in a formal model. Another reason is that our paper focuses on bidding data, and does not involve any data about users' behavior.

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 on 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, we find multiple equilibria even with the specification of payoffs given in (1). By leaving allocative externalities out of our model we thus identify an additional 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

[5]Note that we have not ruled out that the impression value is zero.

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. Given independence, an alternative interpretation of the expression in (1) and of each of the variables in (1) is thus to read them as expected values.

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 the notation in expression (3), but it will be useful to keep in mind that the model admits the alternative interpretation given 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 click rates for different positions is the same for 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 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 $1 \leq k \leq K$ and all alternative positions k' with $k < k' \leq K$:*

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

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

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

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

$$c^{k}_{\phi(k)} \left(v^{k}_{\phi(k)} - b_{\phi(k+1)}\right) \geq 0$$

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

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

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). Clearly, the static model is very stylized. Interactions in practice take place over time. Moreover, the common knowledge assumption, literally interpreted, is, of course, 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.

3 Symmetric Nash Equilibria

We shall initially focus on a particular type of Nash equilibrium, namely equilibria in which bidders don't even have an incentive to win a higher position k if they have to pay b_{k+1} rather than b_k. Varian (2007) has called such equilibria "symmetric Nash equilibria." We discuss asymmetric equilibria in the next section. In Section 5 we shall ask whether there are good reasons to focus on symmetric equilibria.

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:*

$$c_i^{k-1} v_i^{k-1} > c_i^k v_i^k \qquad \textit{and} \qquad v_i^{k-1} \geq v_i^k$$

The first inequality says that the expected value of a higher position for bidder i is at least as large as the expected value of a lower position. The second inequality says that the same monotonicity is true for bidder i's willingness to bid. 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.

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:

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

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

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. □

Two remarks are in order. First, as the second part of Assumption 1 is somewhat restrictive, one might wonder whether it can be relaxed. We have not pursued this question. Second, 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 who will win which 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*

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

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:

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

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.

Proof. Le ϕ be a ranking of bidders that is compatible with a symmetric Nash equilibrium, and let $\hat{\phi}$ be an alternative ranking. Without loss of generality assume that ϕ is the identity mapping. Let p^k (for $k = 1, 2, \ldots, K$) be the Walrasian prices associated with the symmetric equilibrium. By definition of the Walrasian equilibrium we have for all positions k that are won under $\hat{\phi}$ by bidders $\hat{\phi}(k)$ that would also win a position under ϕ, i.e. for whom $\hat{\phi}(k) \leq K$:

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

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

For all positions k that are won under $\hat{\phi}$ by bidders $\hat{\phi}(k)$ that would not win a position under ϕ, i.e. for whom $\hat{\phi}(k) > K$:

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

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

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

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

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. $\square$

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 position $k = 1, 2, 3, \ldots, K-1$*

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

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 bidders are labeled such that $v_1 > v_2 \ldots > v_N$.

4 Asymmetric Nash Equilibria

The game defined in Section 2 has further Nash equilibria if we allow for equilibria that are not symmetric, that is, asymmetric equilibria. It is hard to give a complete description of *all* Nash equilibria. We provide two partial results. The first result concerns the case discussed at the end of the previous section.

Proposition 3. *Under Assumptions 1 and 3 there is an asymmetric Nash equilibrium in which bidder 1 wins position 2, bidder 2 wins position 1, and bidder i wins position i for all $i = 3, 4, ..., K$.*

Proof. Suppose that $b_1, b_2, \ldots, b_N$ is a symmetric Nash equilibrium in which bidder i wins position i for $i = 1, 2, \ldots, K$. By Proposition 1 and Corollary 1 such an equilibrium exists. Define a new vector of bids, $\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_N$, as follows: $\tilde{b}_i = b_i$ for $i = 3, \ldots, N$, $\tilde{b}_1 = b_3 + \varepsilon$ where $\varepsilon > 0$ is very close to zero, and $\tilde{b}_2$ is arbitrary but very large, and, in particular, larger than $\tilde{b}_1$. We now show that we can choose ε so small that no bidder has an incentive to deviate and bid for a different position. We ignore the possibility of deviating and bidding for position 1, because by choosing $\tilde{b}_2$ sufficiently large we can eliminate all incentives to bid for position 1.

We first consider the incentives of bidder 2. Bidder 2 wins position 1 at a price that is ε larger than the price that he paid in the original equilibrium for position 2. If bidder 2 were to deviate and bid for position 2, the change in his payoff would be:

$$\begin{aligned} & a_2 c^2 (v_2^2 - b_3) - a_2 c^1 (v_2^1 - b_3 - \varepsilon) \\ = \; & a_2 \left[c^2 v_2^2 - c^1 v_2^1 + (c^1 - c^2) b_3 + c^1 \varepsilon \right] \end{aligned}$$

We want to show that for sufficiently small but positive ε this expression is negative. For this, it is obviously sufficient to show that $c^2 v_2^2 - c^1 v_2^1 + (c^1 - c^2) b_3$ is strictly negative. This term is linear in b_3. Note that $0 \leq b_3 < v_2^2$. The reason why b_3 has to be strictly less than v_2^2 is that otherwise bidder 2 would make non-positive profits in the original equilibrium. Bidder 2 could then deviate and make positive profits by bidding for the lowest position. It thus suffices to show that the expression in question is strictly negative for $b_3 = 0$, and that it is non-positive for $b_3 = v_2^2$. The first claim follows from Assumption 1, and the second claim follows if we substitute $b_3 = v_2^2$ to obtain $c^1(v_2^2 - v_2^1)$ which is non-negative by Assumption 1. We infer that bidder 2 has no incentive to bid for position 2. Bidder 2 does not have an incentive to deviate and bid for an even lower position because such a deviation was not profitable in the original equilibrium, and in the new equilibrium bidder 2 obtains a higher profit than in the original equilibrium.

We next consider the incentives of bidder 1. Bidder 1 obtains position 2 at the same price at which originally bidder 2 obtained position 2. We show that bidder 1 does not have an incentive to bid for a lower position because in the original equilibrium bidder 2 did not have an incentive to bid for a lower position. Bidder 2 does not have an incentive to bid for a lower position if and only if the following inequality is true for all $k \geq 3$:

$$\begin{aligned} a_2 c^2 (v_2^2 - b_3) &\geq a_2 c^k (v_2^k - b_{k+1}) \Leftrightarrow \\ c^2 v_2^2 - c^k v_2^k &\geq c^2 b_3 - c^k b_{k+1} \end{aligned}$$

By Assumption 3 this implies:

$$\begin{aligned} c^2 v_1^2 - c^k v_1^k &\geq c^2 b_3 - c^k b_{k+1} \Leftrightarrow \\ a_1 c^2 (v_1^2 - b_3) &\geq a_1 c^k (v_1^k - b_{k+1}) \end{aligned}$$

which says that bidder 1 does not have an incentive to bid for position k.

Finally, we argue that bidders $i = 3, 4, \ldots, N$ have no incentive to bid for a different position. Recall that we started with a symmetric equilibrium. Thus, these bidders have no incentive to deviate in the original equilibrium if thy assume that all positions are available to them at the prices which the current winners of those positions pay. Because the price of none of the positions $2, 3, \ldots, K$ have changed, these bidders continue to have no incentives to bid for any of those positions. As noted, an incentive to bid for position 1 can be ruled out by making bidder 2's bid $\tilde{b}_2$ arbitrarily high.

Finally, observe that the equilibrium that we have described is not symmetric. If bidder 1 could obtain position 1 at the same price as bidder 2 obtains it, then for sufficiently small ε he would want to deviate. $\square$

For a further illustration of the multiplicity of Nash equilibria in our model we now take a closer look at the case that $N = K = 3$. For this case explicit calculations that we provide in the Appendix prove the following result.[6] Note that this result relies on none of the assumptions used earlier.

Proposition 4. *Suppose $K = N = 3$. An equilibrium in which bidder i wins position i for $i = 1, 2, 3$ exists if and only if either $c_3^2 v_3^2 \leq c_3^3 v_3^3$ and*

$$c_1^1 v_1^1 \geq c_1^3 v_1^3$$
$$c_2^2 v_2^2 \geq c_2^3 v_2^3$$

or, alternatively, $c_3^2 v_3^2 > c_3^3 v_3^3$ and in addition to the two conditions above also the following two conditions hold:

$$\left(c_1^1 v_1^1 - c_1^3 v_1^3\right) \geq \frac{c_1^1}{c_3^2}\left(c_3^2 v_3^2 - c_3^3 v_3^3\right)$$
$$\left(c_1^1 v_1^1 - c_1^2 v_1^2\right) + \frac{c_1^2}{c_2^2}\left(c_2^2 v_2^2 - c_2^3 v_2^3\right) \geq \frac{c_1^1}{c_3^2}\left(c_3^2 v_3^2 - c_3^3 v_3^3\right)$$

The conditions in Proposition 4 are very weak. The second to last inequality in Proposition 4, for example, requires that the marginal value to bidder 1 of being in position 1 rather than position 3 (the left hand side of the inequality) is at least as large as a variable that is proportional to the marginal value to bidder 3 of being in position 2 rather than position 3, where the proportionality factor is some ratio of

[6] Lahaie (2006, Lemma 3) provides a necessary condition for the existence of a Nash equilibrium that assigns position i to bidder i for all $i = 1, 2, ..., K$. He asserts that this condition is also sufficient, but after publication of his paper he found this part of the claim to be incorrect. We are grateful to Sebastién Lahaie for helpful discussions regarding his result. Our Proposition 3 corrects Lahaie's work for the special case that $N = K = 3$.

click rates. The last inequality is a similarly weak inequality relating the marginal value that bidder 1 derives from being in position 1 rather than position 2, the marginal value that bidder 2 derives from being in position 2 rather than position 3, and the marginal value that bidder 3 would have if he were in position 2 rather than position 3.

We now give an example in which Proposition 4 implies that every allocation of positions to bidders can be an equilibrium allocation. We describe corresponding bid vectors.

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.*

5 Refinements of Nash Equilibrium

In this section we ask whether there are good reasons to expect only some of the Nash equilibria described in the previous sections to be played and not others. In other words, we ask whether there are plausible ways of refining the set of Nash equilibria in the auction game that we are studying. Edelman et. al. (2007) and Varian (2007) focus on symmetric Nash equilibria. We comment on these authors' approaches towards the end of this section. The purpose of this section is to examine the equilibrium selection issue from a different angle than these authors.

The classic way of selecting among equilibria in second price auctions is to rule out Nash equilibria in weakly dominated strategies. Weak dominance arguments are powerful in our model only if the number of bidders is $N = 2$. In that case each bidder knows that even a bid of zero guarantees at least the second position. Bidders only bid for the marginal benefit of being in the first rather than the second position. The auction is strategically equivalent to a single unit, second price auction. It is well-known that the single unit Vickrey auction has multiple Nash equilibria,[7] but that the only strategy that is not weakly dominated is to bid one's true value. This observation extends to our setting. Although the multiplicity of equilibria described in the previous section also prevails in the case of $N = 2$, it is easily seen that each bidder i has a weakly dominant strategy, namely to place the bid b_i that makes bidder i indifferent between obtaining the first position with bid b_i, and obtaining the second position for free. This bid thus solves the following equation:

$$c_i^1(v_i^1 - b_i) = c_i^2 v_i^2 \qquad \Leftrightarrow \tag{19}$$

$$b_i = v_i^1 - \frac{c_i^2}{c_i^1} v_i^2 \tag{20}$$

Unfortunately, the situation changes quite dramatically when $N \geq 3$. This is shown in the following result that provides a range of not weakly dominated bids.

[7] See, for example, Blume and Heidhues (2004) for the case of incomplete information.

Intuitively, the reason why in the case $N \geq 3$ we obtain a range of not weakly dominated bids rather than a single such bids is that for $N \geq 3$ the marginal gain of a bidder who raises his bid is no longer clear unambiguously defined. Raising one's bid may, in the best case, move a bidder up from no listing to top position, but it may also, for example, move a bidder up by only one position, from position k to $k+1$. The range of bids is the range of marginal utilities derived from any such marginal improvement in a bidder's position, modified by a correction factor that takes into account how positions affect click rates.

Proposition 5. *Suppose $N \geq 3$, and Assumption 1 holds. Consider any bidder i. If $N = K$, a bid b_i is not weakly dominated by any other bid $\hat{b}_i$ if and only if:*

$$\min\{v_i^k - \frac{c_i^{k'}}{c_i^k} v_i^{k'} \mid k' > k\} \leq b_i \leq v_i^1$$

If $N > K$, a bid b_i is not weakly dominated by any other bid $\hat{b}_i$ if and only if:

$$\min(\{v_i^k - \frac{c_i^{k'}}{c_i^k} v_i^{k'} \mid k' > k\} \cup \{v_i^K\}) \leq b_i \leq v_i^1$$

Proof. We give the proof in the case $N = K$. The proof in the case $N > K$ is analogous. We first show that any bid outside the range described in Proposition 5 is weakly dominated. First, obviously any bid $b_i > v_i^1$ is weakly dominated by bid v_i^1. It remains to show that any bid below the boundary described in Proposition 5 is weakly dominated. Let b_i be any such bid, and let $\hat{b}_i > b_i$ be another such bid that is also lower than the lower boundary in Proposition 5. We shall show that $\hat{b}_i$

weakly dominates b_i. For some bid vectors of the other bidders it will not make a difference whether bidder i bids $\hat{b}_i$ or whether he bids b_i. Suppose it does make a difference, and that bidder i, by bidding $\hat{b}_i$ acquires position k whereas bidding b_i yields position $k' > k$. We shall show that it is better to bid $\hat{b}_i$ than b_i. The worst case is that bidding b_i acquires position k' at price 0, whereas bidding $\hat{b}_i$ acquires position k at price $\hat{b}_i$. We shall show that in even this case it is better to bid $\hat{b}_i$ rather than b_i:

$$c_i^k(v_i^k - \hat{b}_i) \;>\; c_i^{k'} v_i^{k'} \Leftrightarrow \tag{21}$$

$$\hat{b}_i \;<\; v_i^k - \frac{c_i^{k'}}{c_i^k} v_i^{k'} \tag{22}$$

This holds by construction.

We now show that no bid that satisfies the inequality in Proposition 5 is weakly dominated. Consider any bid $b_i \leq v_i$, and consider any other bid $\hat{b}_i \neq b_i$. We shall construct a vector of bids of the other bidders such that b_i achieves a higher payoff than $\hat{b}_i$. Suppose first $\hat{b}_i < b_i$. Consider a vector of bids of all other bidders such that no two bids are equal to each other, the highest bid of the other bidders is $\hat{b}_i + \varepsilon$ and the second highest of the other bidders bid is $\hat{b}_i - \varepsilon$. Here, ε is a positive number. Suppose that it is sufficiently small so that bidder i, if he bids b_i, wins position 1 and has to pay for it $\hat{b}_i + \varepsilon$, but if he bids $\hat{b}_i$ he wins position 2 and has to pay $\hat{b}_i - \varepsilon$. By Assumption 1 bidder 1 strictly prefers position 1 to position 2 if he has to pay the same price for both positions. Therefore, for ε sufficiently close to zero, he also prefers bidding b_i to bidding $\hat{b}_i$.

Now consider the case that $\hat{b}_i > b_i$. Assume that k, k' are the indices for which the minimum in Proposition 5 is attained. Let b_i be a bid that is equal or greater than this minimum. Let $\hat{b}_i > b_i$ be an alternative bid. Suppose that $N - k'$ bidders bid 0. Suppose that $k' - k$ of the remaining bidders bid $\hat{b}_i - \varepsilon > b_i$, and that all other bidders bid above $\hat{b}_i$. Here, $\varepsilon > 0$. Then bidding b_i wins position k' at price 0, whereas bidding $\hat{b}_i$ wins position k at price $\hat{b}_i - \varepsilon$. It is better to bid b_i if:

$$c_i^{k'} v_i^{k'} \quad > \quad c_i^k (v_i^k - \hat{b}_i - \varepsilon) \Leftrightarrow \tag{23}$$

$$\hat{b}_i - \varepsilon \quad > \quad v_i^k - \frac{c_i^{k'}}{c_i^k} v_i^{k'} \tag{24}$$

By construction

$$\hat{b}_i > v_i^k - \frac{c_i^{k'}}{c_i^k} v_i^{k'} \tag{25}$$

and hence for sufficiently small ε also (24) will be true. □

Observe that Proposition 5 examines only weak dominance when the dominating strategy is a pure strategy. In principle, it may be that more strategies can be ruled out when mixed strategies are considered. We conjecture that this is not the case. A formal examination of this issue would require us to specify bidders' risk attitudes. We have not pursued this issue.

Proposition 5 indicates that there is little chance of obtaining a substantial refinement of the set of Nash equilibria by appealing to weak dominance. In Example 1 neither of the equilibria displayed is ruled out by weak dominance, as the intervals of undominated bids are in that example $[\frac{16}{3}, 16]$, $[\frac{15}{3}, 15]$ and $\frac{14}{3}, 14]$

for bidders 1, 2, and 3 respectively. One can also verify that the equilibria constructed in Edelman et. al. (2007) and Varian (2007) are not in weakly dominated strategies.

Edelman et. al. (2007) and Varian (2007) in more special models than ours select among all Nash equilibria the symmetric Nash equilibria. Varian offers no game theoretic motivation for this. Edelman et. al. (2007, p. 249) argue that the selection can be derived from 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." This construction appears ad hoc. It is not clear why 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, but only in as far as it is implied by weak dominance. Our analysis shows that weak dominance does not always select symmetric Nash equilibria. 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. In any case, in this paper we interpret the data without committing to any particular equilibrium selection.

6 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 for every 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 issue related to delays in bidders' response times.[11]

We augmented the bid data with weekly click-through data for 46 weeks in

[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 bidtool 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 across search terms a new bid is chosen, or an existing bid is revised every 43 minutes across search terms, yielding an average of 63 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.

2004.[12] Based on the click through data we calculate that the ratio c_i^{k-1}/c_i^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 on bidder specific click throughs. The empirical findings have to be interpreted subject to this caveat.

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.

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 price difference between two adjacent positions is 20 cents on average

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

across search terms for 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*.

In the data we see that some bidders are regular bidders for premium positions while other bidders achieve a premium 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 the regular bidders in the subsequent analysis.

7 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-emptyness hypothesis empirically. We discuss the shape of the valuation profiles consistent with the bounds. Section 7.1 illustrates when a set of bid observations yields a non-empty set of valuations. Section 7.2 describes our empirical test results.

7.1 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 $k-th$ highest bid, then, it must be that the bid prices exceed the valuation of the position:

$$v_i^k \leq b_{\phi(k)} \quad \text{for all } k \leq K \tag{26}$$

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:

$$-v_i^k \leq -b_{\phi(k+1)} \tag{27}$$

$$v_i^{k'} \leq \frac{c_i^k}{c_i^{k'}} v_i^k + \left[b_{\phi(k')} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)} \right] \quad \text{for } k' < k \tag{28}$$

$$v_i^{k'} \leq \frac{c_i^k}{c_i^{k'}} v_i^k + \left[b_{\phi(k'+1)} - \frac{c_i^k}{c_i^{k'}} b_{\phi(k+1)} \right] \quad \text{for } K \geq k' > k \tag{29}$$

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.

We can write the above inequalities compactly in matrix notation as

$$\mathbf{A}_t \mathbf{v_i} \leq \alpha_t \tag{30}$$

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 (30) as $\mathbf{V_i^T}$,

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

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, $K =$

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

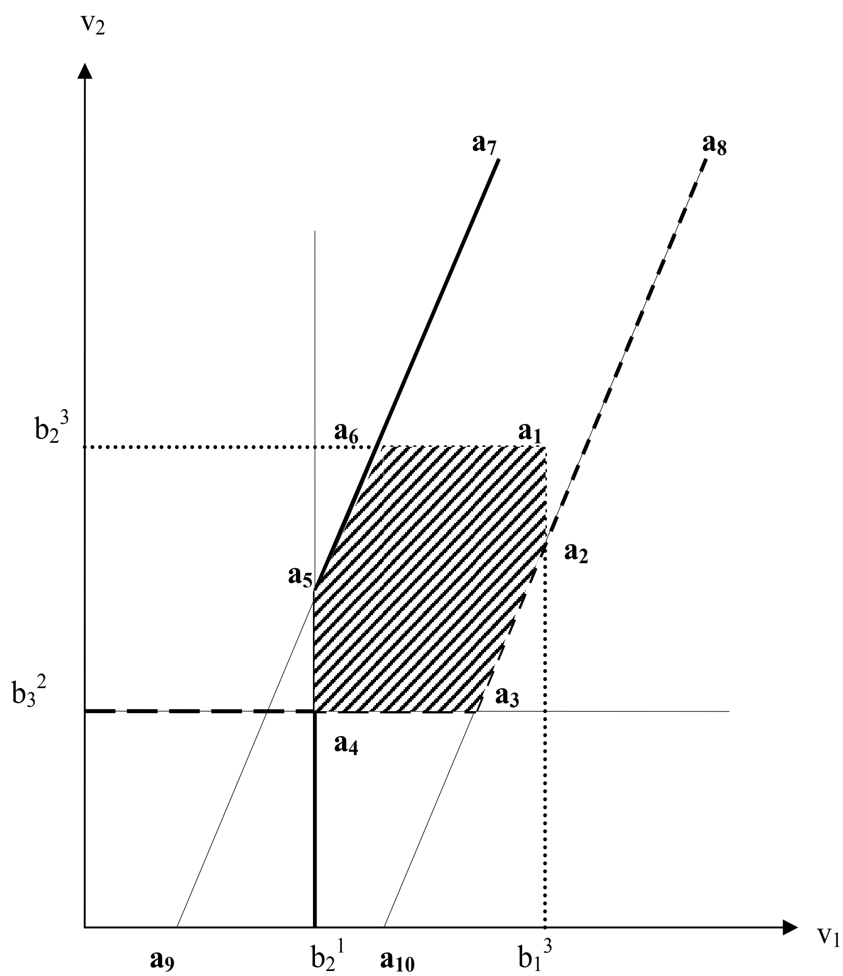

Figure 1: Valuations consistent with hypothetical bids

2.[14] The dark shaded area with boundary points a_1, a_2, a_3, a_4, a_5, and a_6 is consistent with three hypothetical bid vectors b^1, b^2, b^3 where the superscript in the bid vector indicates that bidder i wins item 1, item 2, or no item, respectively. Item 1 is won in the area south-east of the solid line segments $b^1_{\phi(2)}$, a_5 and a_7.[15] Item 2 is won in the area north-west of the dashed line segments $b^2_{\phi(3)}$, a_3 and a_8.[16] No position is won in the area south-west of the dotted line-segments going through the points $b^3_{\phi(3)}$, a_1, and $b^3_{\phi(1)}$.

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 $a_{10} = \min_{t \in T^2}(b^t_{\phi(1)} - (c^2_i/c^1_i)b^t_{\phi(3)})$, and the solid line segments are defined by $b^1_{\phi(1)} = \max_{t \in T^1}(b^t_{\phi(2)})$ and $a_9 = \max_{t \in T^1}(b^t_{\phi(2)} - (c^2_i/c^1_i)b^t_{\phi(3)})$. 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.[17]

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

[14] In Figure 1, we write "b^1_2" for $b^1_{\phi(2)}$ etc.

[15] The line going through the points a_5 and a_7 has slope c^1_i/c^2_i and intercept $b^1_{\phi(3)} - \left(c^1_i/c^2_i\right) b^1_{\phi(2)}$.

[16] Here the line going through the points a_3 and a_8 has slope c^1_i/c^2_i and intercept $b^2_{\phi(3)} - \left(c^1_i/c^2_i\right) b^2_{\phi(1)}$.

[17] If T^1 is empty, then the left boundary of the shaded area will equal the vertical line $\left(0, v^2_i\right)$ as by assumption $v^1_i > 0$. If T^2 is empty, then the bottom boundary of the shaded area will equal the horizontal line $\left(v^1_i, 0\right)$. 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 an 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-emptyness 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, the necessary pairwise boundary condition 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 the necessary pairwise boundary condition. 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 than 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.

7.2 Revealed Preference Test Results

This section examines the revealed preference hypothesis for our data. We also comment on the shape of the valuation profile for observations that satisfy the revealed preference hypothesis.

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.[18] 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

[18] 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.

is satisfied on each sub-period. Notice that period T_{i1} starts at the point of time 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-emptyness 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.[19]

The results are the following: For 50 percent of observations the set $\mathbf{V_i^T}$ is

[19]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, $\min_{t\in T^{11}} b^t_{(i)}$, and we take as the lower bound the price paid when position k is won, $\max_{t\in T^k} b^t_{\phi(k+1)}$. When the upper bound does not exist, we replace it with 15. When the lower bound does not exist, we set it to 0.

non-empty. We can conclude that for about half the observations the revealed preference hypothesis is satisfied.

Next, we explore the shape of the valuation profiles that are consistent with revealed preference.

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 multivariate 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 97 percent of observations.

The test results indicate that the willingness to pay decreases with the position. We conclude this section with a caveat of the revealed preference approach as 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.

8 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, Susan and Glenn Ellison. 2008. "Position Auctions with Consumer Search." http://kuznets.harvard.edu/ athey/position.pdf.

Blume, Andreas and Paul Heidhues. 2004. "All Equilibria of the Vickrey Auction." *Journal of Economic Theory*, 114(1): 170-177.

Edelman, Benjamin and Michael Ostrovsky. 2007. "Strategic Bidder Behavior in Sponsored Search Auctions." *Decision Support Systems*, 43(1): 192-198.

Edelman, Benjamin, Michael Ostrovsky and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review*, 97(1): 242-259.

Jehiel, Phililppe and Benny Moldovanu. 2006. "Allocative and Informational Externalities in Auctions and Related Mechanisms," In *The Proceedings of the 9th World Congress of the Econometric Society*, ed. Richard Blundell, Whhitney Newey, and Torsten Persson, 102-135. Cambridge: Cambridge University Press.

Lahaie, Sébastien. 2006. "An Analysis of Alternative Slot Auction Designs for Sponsored Search." In *Proceedings of the 7th ACM Conference on Electronic Commerce*, ed. John Feigenbaum, John Chuang and David Pennock, 218-227. Ann Arbor: Association for Computing Machinery.

Milgrom, Paul. 2000. "Putting Auction Theory to Work." *Journal of Political Economy*, 108(2): 245-272.

Moldovanu, Benny and Aner Sela. "The Optimal Allocation of Prizes in Contests." *American Economic Review*, 91(3): 542-558.

Moldovanu, Benny, Aner Sela and Xianwen Shi. 2007. "Contests for Status." *Journal of Political Economy*, 115(2): 338-363.

Varian, Hal. 2007. "Position Auctions," *International Journal of Industrial Organization*, 25(6): 1163-1178.

Appendix: Proof of Proposition 4

Observe that we can choose b_1 arbitrarily high and thus ensure that no bidder has an incentive to bid for position 1. Therefore, we can find a Nash equilibrium of the required type if and only if we can find non-negative bids b_2 and b_3 such that four incentive constraints hold. Firstly, bidder 1 does not want to bid for position 2:

$$\begin{aligned} c_1^1(v_1^1 - b_2) \geq c_1^2(v_1^2 - b_3) &\Leftrightarrow \\ b_2 - \frac{c_1^2}{c_1^1} b_3 \leq v_1^1 - \frac{c_1^2}{c_1^1} v_1^2 & \end{aligned} \tag{31}$$

Secondly, bidder 1 does not want to bid for position 3:

$$\begin{aligned} c_1^1(v_1^1 - b_2) \geq c_1^3 v_1^3 &\Leftrightarrow \\ b_2 \leq v_1^1 - \frac{c_1^3}{c_1^1} v_1^3 & \end{aligned} \tag{32}$$

Next, bidder 2 does not want to bid for position 3:

$$\begin{aligned} c_2^2(v_2^2 - b_3) \geq c_2^3 v_2^3 &\Leftrightarrow \\ b_3 \leq v_2^2 - \frac{c_2^3}{c_2^2} v_2^3 & \end{aligned} \tag{33}$$

Finally, bidder 3 does not want to bid for position 2:

$$\begin{aligned} c_3^3 v_3^3 \geq c_3^2 (v_3^2 - b_2) \Leftrightarrow \\ b_2 \geq v_3^2 - \frac{c_3^3}{c_3^2} v_3^3 \end{aligned} \tag{34}$$

We now distinguish two cases. The first case is that the lower bound in (34) is not-positive.

$$\begin{aligned} v_3^2 - \frac{c_3^3}{c_3^2} v_3^3 \leq 0 \Leftrightarrow \\ c_3^2 v_3^2 \leq c_3^3 v_3^3 \end{aligned} \tag{35}$$

In this case, a necessary and sufficient condition for the existence of a non-negative solution to (31)-(33) is that the right hand sides of (32) and (33) are non-negative. The necessity is obvious. To see sufficiency note that $b_2 = b_3 = 0$ will solve (31)-(33) in this case. The upper boundary in (32) is non-negative if:

$$\begin{aligned} v_1^1 - \frac{c_1^3}{c_1^1} v_1^3 \geq 0 \Leftrightarrow \\ c_1^1 v_1^1 \geq c_1^3 v_1^3 \end{aligned} \tag{36}$$

The upper boundary in (33) is non-negative if:

$$\begin{aligned} v_2^2 - \frac{c_2^3}{c_2^2} v_2^3 \geq 0 \Leftrightarrow \\ c_2^2 v_2^2 \geq c_2^3 v_2^3 \end{aligned} \tag{37}$$

Inequalities (36) and (37) are the first two conditions in Proposition 4.

Now suppose that the lower bound in (34) is positive.

$$c_3^2 v_3^2 > c_3^3 v_3^3 \tag{38}$$

Obviously, (36) and (37) remain necessary. But we also need that the upper boundary in (32) is not less than the lower boundary in (34):

$$\begin{aligned} v_1^1 - \frac{c_1^3}{c_1^1} v_1^3 \geq v_3^2 - \frac{c_3^3}{c_3^2} v_3^3 \Leftrightarrow \\ c_1^1 v_1^1 - c_1^3 v_1^3 \geq \frac{c_1^1}{c_3^2}(c_3^2 v_3^2 - c_3^3 v_3^3) \end{aligned} \tag{39}$$

If (36), (37) and (39) hold, then the difference on the left hand side of (31) is minimized when b_2 is at the lower bound given by (34), and b_3 is at the upper bound given by (33). Thus, a necessary and sufficient condition for the existence of a non-negative solution is that for these choices of b_2 and b_3 inequality (31) holds:

$$\begin{aligned} v_3^2 - \frac{c_3^3}{c_3^2} v_3^3 - \frac{c_1^2}{c_1^1}\left(v_2^2 - \frac{c_2^3}{c_2^2} v_2^3\right) \leq v_1^1 - \frac{c_1^2}{c_1^1} v_1^2 \Leftrightarrow \\ (c_1^1 v_1^1 - c_1^2 v_1^2) + \frac{c_1^2}{c_2^2}(c_2^2 v_2^2 - c_2^3 v_2^3) \geq \frac{c_1^1}{c_3^2}(c_3^2 v_3^2 - c_3^3 v_3^3) \end{aligned} \tag{40}$$

Inequalities (39) and (40) are the second pair of conditions in Proposition 4.

□