A Structural Model of Sponsored Search Advertising Auctions*

Susan Athey[a] and Denis Nekipelov[b].

[a] Department of Economics, Harvard University, Cambridge, MA & Microsoft Research

[b] Department of Economics, UC Berkeley, Berkeley, CA & Microsoft Research

This Version: February 2011

Abstract

Sponsored links that appear beside Internet search results on the major search engines are sold using real-time auctions. Advertisers place standing bids, and each time a user enters a search query, the search engine holds an auction. Ranks and prices depend on advertiser bids as well as "quality scores" that are assigned for each advertisement and user query. Existing models assume that bids are customized for a single user query. In practice queries arrive more quickly than advertisers can change their bids, and quality scores vary over time and across user queries. This paper develops a new model that incorporates these features. In contrast to prior models, which produce multiplicity of equilibria, we provide sufficient conditions for existence and uniqueness of equilibria. In addition, we propose a homotopy-based method for computing equilibria.

We propose a structural econometric model. With sufficient uncertainty in the environment, the valuations are point-identified; otherwise, we consider bounds on valuations. We develop an estimator which we show is consistent and asymptotically normal, and we assess the small sample properties of the estimator using Monte Carlo.

We apply the model to historical data for several search phrases. Our model yields lower implied valuations and bidder profits than those suggested by the leading alternative model that ignores uncertainty. We find that bidders have substantial strategic incentives to reduce their expressed demand in order to reduce the unit prices they pay in the auctions. These incentives are asymmetric across bidders, leading to inefficient allocation, which does not arise in models that ignore uncertainty. Even on the highly competitive search phrases we study, bidders earn substantial profits from the sponsored search auctions: values per click are two to three times as high as the price paid per click. We also find that for the search phrases we study, the auction mechanism used in practice is less efficient than a Vickrey auction by a few percent, but the revenue effects are positive for some search phrases and negative for others.

---

* We acknowledge Anton Schwaighofer, Michael Ostrovsky, Dmitry Taubinsky, Nikhil Agarwal, Hoan Lee, Daisuke Hirata, Chris Sullivan, Andrew Bacher-Hicks, and Maya Meidan for helpful comments and assistance with this paper. We also thank the participants at Cowles foundation conference (Yale), INFORMS meetings, NBER winter meeting, the Ad Auctions Workshop at EC 2010, and seminar audiences at Harvard, MIT, UC Berkeley and Microsoft Research for helpful comments.

1 Introduction

Online advertising is a big business. Search advertising is an important way for businesses, both online and offline, to attract qualified leads; Google revenues from search advertising auctions top $20 billion per year.

This paper develops and analyzes original theoretical and econometric models of advertiser behavior in the auctions, and applies these models to a real-world dataset. The methods can be used to infer bidder valuations from their observed bids, and to reliably and quickly compute counterfactual equilibrium outcomes for differing economic environments (e.g. different auction format, altered competitive environment). We apply the tools to address economic questions. For example, we quantify the extent to which existing auction rules lead to inefficient allocation as compared to a Vickrey auction, as well as the way in which competition affects the magnitude of the inefficiency.

The model proposed in this paper differs from existing economic models (e.g. Borgers et al. (2007), Edelman et al. (2007), Ghose and Yang (2009), Mela and Yao (2009), Varian (2007)) by incorporating more realistic features of the real-world bidding environment. We show that our more realistic model has several advantages in terms of tractability, ability to rationalize bidding data in an equilibrium framework, and in the specificity of the predictions it generates: it simultaneously avoids the problems of multiplicity of equilibria and lack of point-identification of values that are the focus of much of the existing literature.

Sponsored links that appear beside Internet search results on the major search engines are sold using real-time auctions. Advertisers place standing bids that are stored in a database, where bids are associated with search phrases that form part or all of a user's search query. Each time a user enters a search query, applicable bids from the database are entered in an auction. The ranking of advertisements and the prices paid depend on advertiser bids as well as "quality scores" that are assigned for each advertisement and user query. These quality scores vary over time, as the statistical algorithms incorporate the most recent data about user clicking behavior on this and related advertisements and search queries.

Edelman et al. (2007) and Varian (2007) assume that bids are customized for a single user query and the associated quality scores; alternatively, one can interpret the models as applying to a situation where quality scores, advertisement texts, and user behavior are static over a long period of time which is known to advertisers. However, in practice quality scores do vary from query to query, queries arrive more quickly than advertisers can change their bids,[1] and advertisers cannot perfectly predict changes

[1] Although bids can be changed in real time, the system that runs the real-time auction is updated only periodically based on the state at the time of the update, so that if bids are adjusted in rapid succession, some values of the bids might

in quality scores. This paper develops a new model where bids apply to many user queries, while the quality scores and the set of competing advertisements may vary from query to query. In contrast to existing models, which produce multiplicity of equilibria, we provide sufficient conditions for existence and uniqueness of equilibria, and we provide evidence that these conditions are satisfied empirically for the search phrases we study. One requirement is sufficient uncertainty about quality scores relative to the gaps between bids. We show that the necessary conditions for equilibrium bids can be expressed as an ordinary differential equation, and we develop a homotopy-based method for calculating equilibria given bidder valuations and the distribution of uncertainty.

We then propose a structural econometric model. With sufficient uncertainty in the environment, valuations are point-identified; otherwise, we propose a bounds approach. We develop an estimator for bidder valuations, establish consistency and asymptotically normality, and use Monte Carlo simulation to assess the small sample properties of the estimator.

In the last part of the paper, we apply the model to historical data for two search phrases. We start by comparing the estimates implied by our model to those implied by prior approach focusing on envy-free equilibria and ignoring uncertainty, showing that our model yields lower implied valuations and bidder profits. We then use our estimates to examine the magnitude of bidders' incentives to shade their bids and reduce their expressed demands in order to maximize profits, focusing on the degree to which such incentives are asymmetric across bidders with high versus low valuations. We demonstrate that differential bid-shading leads to inefficient allocation.

The incentives for "demand-reduction" are created by the use of a "generalized second-price auction" (GSP), which Edelman et al. (2007) and Varian (2007) show is different from a Vickrey auction. In a model without uncertainty, one of the main results of Edelman et al. (2007) and Varian (2007) is that the GSP auction is outcome-equivalent to a Vickrey auction for a particular equilibrium selection, which we refer to as the "EOS" equilibrium; however, we show that the equivalence breaks down when bidders use differential bid shading and the same bids apply to many user queries with varying quality scores.

Because a Vickrey auction, run query by query, would lead bidders to bid their values and thus would result in efficient allocation in each auction even when quality scores vary query by query, our findings suggest that there is a non-trivial role for auction format to make a difference in this setting, a finding that would not be possible without uncertainty and using the EOS equilibrium, since then, auction format plays no role.

In our model, the revenue ranking of the GSP and the Vickrey auction is ambiguous. For our first

never be applied.

search phrase, we find that the Vickrey auction raises 3.4% more revenue than the GSP, even though the efficiency difference is only about 2.3%. For our second search phrase, the GSP raises 12% more revenue, even though efficiency difference is about 0.6%.

We analyze the elasticities of residual supply curve for clicks faced by each bidder, which determine the equilibrium gap between each bidder's value and its price per click (the Lerner index, i.e., the profit per click' as a percentage of the price per click). We find that for both search phrases we consider, the Lerner index is higher for the top positions than the bottom ones, and for one of our search phrases, the top positions have substantially higher profits. The higher profits correspond to two relatively high-value advertisers who each dominate one of the top two positions with fairly substantial gaps between their bids, while the lower ranks are more competitive in the sense that many advertisers with similar values compete for them.

Finally, we show that our computational approach is tractable in practice, and we use it to compute counterfactual equilibria in order to evaluate the impact of an increase in entry of advertisers.

2 Overview of Sponsored Search Auctions

Auction design for sponsored search auctions has evolved over time; see Edelman et al. (2007) for a brief history. Since the late 1990s, most sponsored search in the U.S. has been sold at real-time auctions. Advertisers enter per-click bids into a database of standing bids. They pay the search engine only when a user clicks on their ad. Each time a user enters a search query, bids from the database of standing bids compete in an auction. Applicable bids are collected, the bids are ranked, per-click prices are determined as a function of the bids (where the function varies with auction design), and advertisements are displayed in rank order, for a fixed number of slots J. Clicks are counted by the ad platform and the advertiser pays the per-click price for each click. For simplicity, we will focus exposition on a single search phrase, where all advertisers place a distinct bid that is applicable to that search phrase alone.[2]

In this setting, even when there are fewer bidders than positions, bidders are motivated to bid more aggressively in order to get to a higher position and receive more clicks. Empirically, it has been well established that appearing in a higher position on the screen causes the advertisement to receive more clicks. We let $\alpha_{i,j}$ be the ratio of the "click-through rate" (CTR, or the probability that a given query makes a click on the ad) that advertiser i would receive if its ad appears in position j, and the CTR

[2]In general, bidders can place "broad match" bids that apply to any search phrase that includes a specified set of "keywords," but for very high-value search phrases, such as the ones we study here, most advertisers who appear on the first page use exact match bidding.

of the ad in position 1. The CTR for the highest slot can be tens to hundreds of times higher than for lower slots. The way in which CTRs diminish with position depends on the search phrase in question.

In 2002, Google introduced the "generalized second price" auction. The main idea of this auction is that advertisements are ranked in order of the per-click bids (say, $b_1, .., b_N$ with $b_i > b_{i+1}$) and a bidder pays the minimum per-click price required to keep the bidder in her position (so bidder i has position i and pays b_{i+1}). When there is only a single slot, this auction is equivalent to a second-price auction, but with multiple slots, it differs. Subsequently, Google modified the auction to include weights, called "quality scores," for each advertisement, where scores are calculated separately for each advertisement on each search phrase. These scores were initially based primarily on the estimated click-through rate the bidder would attain if it were in the first position. The logic behind this design is straightforward: allocating an advertisement to a given slot yields expected revenue equal to the product of the price charged per click, and the click-through rate. Thus, ranking advertisements by the product of the click-through rate and the per-click bid is equivalent to ranking them by the expected revenue per impression (that is, the revenue from displaying the ad). Later, Google introduced additional variables into the determination of the weights, including measures of the match between the advertisement and the query. Although the formulas used by each search advertising platform are proprietary information and can change at any time, the initial introduction of quality scores by Microsoft and Yahoo! was described in the industry as a GSP auction using the "click-weighting" version of quality scores, that is, quality scores reflect primarily the expected click-through rate of the advertisement.

In practice, there are also a number of reserve prices that apply for the different advertising platforms. Our empirical application generally has non-binding reserve prices, but we include them in the theory.

3 The Model

3.1 A Static Model of a Score-Weighted Generalized Second-Price Auction

We begin with a static model, where each of I advertisers simultaneously place per-click bids b_i on a single search phrase. The bids are then held fixed and applied to all of the users who enter that search phrase over a pre-specified time period (e.g. a day or a week). There is a fixed number of advertising slots J in the search results page.

We model consumer searches as an exogenous process, where each consumer's clicking behavior is random and $\bar{c}_{i,j}$, the average probability that a consumer clicks on a particular ad in a given position, is the same for all consumers. It will greatly simplify exposition and analysis to maintain the assumption that the parameters $\alpha_{i,j}$ (the ratio of advertisement i's CTR in position j to its CTR in position 1) satisfy $\alpha_{i,j} = \alpha_{i',j} \equiv \alpha_j$ for all advertisements i, i'; we will maintain that assumption throughout the paper.[3] That is, there exists a vector of advertisement effects, γ_i, $i = 1, .., I$, and position effects α_j, $j = 1, .., J$, with $\alpha_1 = 1$, such that $\bar{c}_{i,j}$ can be written

$$\bar{c}_{i,j} = \alpha_j \gamma_i.$$

The ad platform conducts a click-weighted GSP auction. Each advertisement i is assigned score s_i, and bids are ranked in order of the product $b_i\, s_i$. In general discussion we will use i to index bidders and j to index positions (slots). We will use the double index notation k_j to denote the bidder occupying slot j. The per-click price p_{k_j} that bidder k_j in position j pays is determined as the minimum price such that the bidder remains in her position

$$p_{k_j} = \min\{b_{k_j} : s_{k_j} b_{k_j} \geq s_{k_{j+1}} b_{k_{j+1}}\} = \frac{s_{k_{j+1}} b_{k_{j+1}}}{s_{k_j}}.$$

Note that advertiser k_j does not directly influence the price that she pays, except when it causes her to change positions, so in effect an advertiser's choice of bid determines which position she attains, where the price per click for each position is exogenous to the bidder and rises with position.

To interpret this auction, observe that if for each i, $s_i = \gamma_i$, then the expected revenue the ad platform receives from placing bidder k_j in position j is $\alpha_j \gamma_{k_{j+1}} b_{k_{j+1}}$ which is what the platform would receive if instead, it had placed bidder in slot $j + 1$ in position j and charged bidder in slot $j + 1$ her per-click bid, $b_{k_{j+1}}$, for each click. So each bidder pays, in expectation, the per-impression revenue that would have been received from the next lowest bidder.

We include the possibility of reserve prices in the auction. The reserve price is assumed to be set in units of score-weighted bids, so that an advertisement is considered only if $s_i b_i > r$. We introduce the reserve price in the model by adding an artificial bidder $I + 1$ such that $b_{I+1} = r$ and $s_{I+1} = 1$.

We assume that advertisers are interested in consumer clicks and each advertiser i has a value v_i associated with a consumer click. The profile of advertiser valuations in a particular market $(v_1, \ldots, v_I)$ is fixed, and advertisers know their valuations with certainty. Each click provides the advertiser i with the surplus $v_i - p_i$. The advertisers are assumed to be risk-neutral.

[3] Empirically, this assumption can be rejected for many search phrases (see, e.g., Jeziorski and Segal (2009)), but the deviations are often small, and the assumption is more likely to hold when the advertisements are fairly similar, as is the case for the search phrases in our sample.

3.2 Equilibrium Behavior with No Uncertainty (NU)

The structure of Nash equilibria in the environment similar to that described in the previous subsection has been considered in Edelman et al. (2007) and Varian (2007). We can write the expected surplus of advertiser i from occupying the slot j as

$$\bar{c}_{i,j}\left(v_i - p_j\right) = \alpha_j \gamma_i \left(v_i - \frac{s_{k_{j+1}} b_{k_{j+1}}}{s_i}\right).$$

The existing literature, including Edelman et al. (2007) and Varian (2007), focus on the case where the bidders know the set of competitors as well as the score-weighted bids of the opponents, and they consider ex post equilibria, where each bidder's score-weighted bid must be a best response to the realizations of $s_{k_{j+1}} b_{k_{j+1}}$ (and recall we have also assumed that the $\bar{c}_{i,j}$ are known). Let us start with this case, which we will refer to as the "No Uncertainty" (NU) case.

The set of bids constituting a full-information Nash equilibrium in the NU model, where each bidder finds it unprofitable to deviate from her assigned slot, are those that satisfy

$$\begin{aligned} \alpha_j \left(v_{k_j} - \tfrac{s_{k_{j+1}} b_{k_{j+1}}}{s_{k_j}}\right) &\geq \alpha_l \left(v_{k_j} - \tfrac{s_{k_{l+1}} b_{k_{l+1}}}{s_{k_j}}\right), \ l > j \\ \alpha_j \left(v_{k_j} - \tfrac{s_{k_{j+1}} b_{k_{j+1}}}{s_{k_j}}\right) &\geq \alpha_l \left(v_{k_j} - \tfrac{s_{k_l} b_{k_l}}{s_{k_j}}\right), \ l < j. \end{aligned}$$

It will sometimes be more convenient to express these inequalities in terms of score-weighted values, as follows:

$$\min_{l<j} \frac{s_{k_l} b_{k_l} \alpha_j - s_{k_{j+1}} b_{k_{j+1}} \alpha_l}{\alpha_l - \alpha_j} \geq s_{k_j} v_{k_j} \geq \max_{l>j} \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_j - s_{k_{l+1}} b_{k_{l+1}} \alpha_l}{\alpha_j - \alpha_l}.$$

An equilibrium always exists, but it is typically not unique, and equilibria may not be monotone: bidders with higher score-weighted values may not be ranked higher.

Edelman et al. (2007) and Varian (2007) define a refinement of the set of equilibria, which Edelman et al. (2007) call "envy-free": no bidder wants to exchange positions and bids with another bidder. The set of envy-free equilibria is characterized by a tighter set of inequalities:

$$s_{k_j} v_{k_j} \geq \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_j - s_{k_{j+2}} b_{k_{j+2}} \alpha_{j+1}}{\alpha_j - \alpha_{j+1}} \geq s_{k_{j+1}} v_{k_{j+1}}. \tag{3.1}$$

The term in between the two inequalities is interpreted as the incremental costs divided by the incremental clicks from changing position, or the "incremental cost per click" $ICC_{j,j+1}$:

$$ICC_{j,j+1} = \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_j - s_{k_{j+2}} b_{k_{j+2}} \alpha_{j+1}}{\alpha_j - \alpha_{j+1}}.$$

Envy-free equilibria are monotone, in that bidders are ranked by their score-weighted valuations, and have the property that local deviations are the most attractive–the equilibria can be characterized by incentive constraints that ensure that a bidder does not want to exchange positions and bids with either the next-highest or the next-lowest bidder.

Edelman et al. (2007) consider a narrower class of envy-free equilibria, the one with the lowest revenue for the auctioneer and the one that coincides with Vickrey payoffs as well as the equilibrium of a related ascending auction game. They require

$$s_{k_j} v_{k_j} \geq ICC_{j,j+1} = s_{k_{j+1}} v_{k_{j+1}}. \tag{3.2}$$

Edelman et al. (2007) show that despite the fact that payoffs coincide with Vickrey payoffs, bidding strategies are not truthful: bidders shade their bids, trading off higher price per click in a higher position against the incremental clicks they obtain from the higher position.

We refer to the equilibrium defined by $ICC_{j,j+1} = s_{k_{j+1}} v_{k_{j+1}}$ as the EOS equilibrium, and the equilibrium defined by $s_{k_j} v_{k_j} = ICC_{j,j+1}$ as the NU-EFLB equilibrium (for "envy-free lower bound").

3.3 Equilibrium Behavior with Score and Entry Uncertainty (SEU)

In reality, advertiser bids apply to many unique queries by users. Each time a query is entered by a user, the set of applicable bids is identified, scores are computed, and the auction is conducted as described above. In practice, both the set of applicable bids and the scores vary from query to query. This section describes this uncertainty in more detail and analyzes its impact on bidding behavior.

3.3.1 Uncertainty in Scores and Entry in the Real-World Environment

The ad platform produces scores at the advertisement-query level using a statistical algorithm. A key component of quality scores is the click-through rate that the platform predicts the advertisement will attain. In practice, the distribution of consumers associated with a given search query and/or their preferences for given advertisers (or for advertisements relative to algorithmic links) can change over time, and so the statistical algorithms are continually updated with new data. Google has stated publicly that it uses individual search history to customize results to individual users; to the extent that Google continues to use the GSP, ranking ads differently for different users can be accomplished by customizing the quality scores for individual users.

In Appendix B, we illustrate how introducing small amounts of uncertainty affects equilibrium in the NU model. However, because the real-world environment incorporates substantial uncertainty, we focus our exposition in the text on non-trivial uncertainty.

We assume that the score of a particular bidder i for a user query is a random variable, denoted s_i, which is equal to

$$s_i = \overline{s}_i \varepsilon_i,$$

where ε_i is a shock to the score induced by random variation in the algorithm's estimates.[4]

Now consider uncertainty in bidder entry. There are many sources of variation in the set of advertisements that are considered for the auction for a particular query. First, some bidders specify budgets (limits on total spending at the level of the account, campaign, or keyword), which the ad platforms respect in part by spreading out the advertiser's participation in auctions over time, withholding participation in a fraction of auctions. Bidders may also "pause" and "reactivate" their campaigns. Second, bidders experiment with multiple advertisements and with different ad text. These advertisements will have distinct click-through rates, and so will appear to other bidders as distinct competitors. For new advertisements, it takes some time for the system to learn the click-through rates; and the ad platform's statistical algorithm may "experiment" with new ads in order to learn. Third, some bidders may target their advertisements at certain demographic categories, and they may enter different bids for those categories (platforms make certain demographic categories available for customized bidding, such as gender, time of day, day of week, and user location).

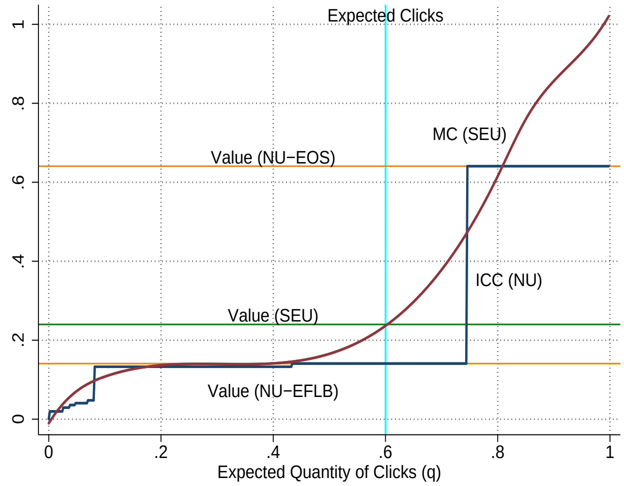
For these and other reasons, it is typical for the configuration of ads to vary on the same search phrase; this variation is substantial for all three major search ad platforms in the U.S., as can be readily verified by repeating the same query from different computers or over time.

The role of the score and entry uncertainty can be illustrated by Figure 1. The x-axis gives the (expected) click-through rate a bidder receives (the "click share"), relative to the click-through rate it would attain in the top position (that is, the average of α_j over the positions a bidder with a given bid experiences). The step function in the figure shows the relationship between the incremental cost per click and expected number of clicks for a single user query, with a commonly observed configuration of advertisements and associated bids, and assuming that each advertisement is assigned a score equal to its average score from the week. As the bidder in question's score-weighted bid increases and crosses the score-weighted bid of each opponent, the bidder moves to a higher position, receiving a higher average CTR. Given a value of $\alpha \in [\alpha_{j+1}, \alpha_j]$, the associated incremental cost per click is $ICC_{j,j+1}$.

The smooth curve shows how uncertainty affects the incremental cost per click. The curve is con-

[4]In the subsequent discussion we often refer to $\overline{s}_i$ as the "mean score" of the bidder. We will further make an identifying assumption $E[\log \varepsilon_i] = 0$ which may not imply that the mean score is equal to $\overline{s}_i$. We will indicate specific points where this distinction is important.

Figure 1: Marginal and Incremental Cost and Implied Valuations for Alternative Models



structed by varying the bid of a given advertisement. For each value of the bid, we calculate the expectation of the share of possible clicks the advertisement receives, where the expectation is taken over possible realizations of quality scores, using the distribution of these scores we estimate below. Corresponding to each expected click share, we calculate the marginal cost of increasing the click share and plot that on the y-axis (details of the computation are provided below). The marginal cost curve increases smoothly rather than in discrete steps because the same advertisement with the same bid would appear in different positions for different user queries, and changing the bid slightly affects outcomes on a small but non-zero share of user queries.

This smoothness reflects the degree of uncertainty faced by the advertisers. For the search phrases we consider, the most commonly observed advertisements have a standard deviation of their position number close to 1 position.

3.3.2 Formalizing the Score and Entry Uncertainty (SEU) Model

Start with the NU model, and consider the following modifications. Bids are fixed for a large set of user queries on the same search phrase, but the game is still a simultaneous-move game: bidders simultaneously select their bids, and then they are held fixed for a pre-specified period of time. Let $\tilde{C}^i$ be a random subset of advertisers excluding advertiser i, with typical realization C^i, and consider shocks to scores as defined in the last subsection.

We use the solution concept of ex post Nash equilibrium. In the environment with uncertainty, we need to specify bidder beliefs. Since our environment has private values (bidders would not learn anything about their own values from observing the others' information) and we model the game as static, an ex post Nash equilibrium merely requires that each bidder correctly anticipates the mapping from his own bids to expected quantities and prices, taking as given opponent bids. Note that the major search engines provide this feedback to bidders through advertiser tools (that is, bidders can enter a hypothetical bid on a keyword and receive estimates of clicks and cost).

Despite these weak information requirements, for simplicity of exposition, we endow the bidders with information about the primitive distributions of uncertainty in the environment. That is, we assume that advertisers correctly anticipate the share of user queries where each configuration of opposing bidders C^i will appear; the mean of each opponent's score-weighted bid, $b_i\overline{s}_i$; and the distribution of shocks to scores, $F_\varepsilon(\cdot)$.

Define Φ_{ik}^j to be an indicator for the event that bidder i is in slot j and bidder k is in slot $j+1$. Let $\mathcal{C}_{j,k}^i$ be the set of all subsets of C^i with cardinality j that contains k, with typical element $C_{j,k}^i$, representing the set of bidders above bidder i as well as k. Also recall that we introduce the reserve price as an additional bidder $I+1$ with bid $b_{I+1}=r$, $\overline{s}_{I+1}=1$ and $\varepsilon_{I+1}=1$. Let $b,\overline{s},\varepsilon$ be vectors of bids, mean scores, and shocks to scores, respectively. Then:

$$\Phi_{ik}^j\left(b,\overline{s},\varepsilon;C^i\right)=\sum_{C_{j,k}^i\in\mathcal{C}_{j,k}^i}\prod_{m\in C_{j,k}^i\backslash\{k\}}\mathbf{1}\left\{b_m\overline{s}_m\varepsilon_m>b_i\overline{s}_i\varepsilon_i\right\}\prod_{m\in C^i\backslash C_{j,k}^i}\mathbf{1}\left\{b_m\overline{s}_m\varepsilon_m<b_k\overline{s}_k\varepsilon_k\right\}\mathbf{1}\left\{b_i\overline{s}_i\varepsilon_i>b_k\overline{s}_k\varepsilon_k\right\}.$$

We can then write the expected number of clicks a bidder will receive as a function of her bid b_i as follows:

$$Q_i(b_i;\, b_{-i},\overline{s})=\mathbb{E}_{\tilde{C}^i,\varepsilon}\left[\sum_{j=1,..,J}\sum_{k\in\tilde{C}^i}\Pr(\Phi_{ik}^j\left(b,\overline{s},\varepsilon;\tilde{C}^i\right)=1)\cdot\alpha_j\cdot\gamma_i\right].$$

The expected total expenditure of the advertiser for the clicks received with bid b_i can be written

$$TE_i(b_i;\, b_{-i},\overline{s})=\mathbb{E}_{\tilde{C}^i,\varepsilon}\left[\sum_{j=1,..,J}\sum_{k\in\tilde{C}^i}\Pr(\Phi_{ik}^j\left(b,\overline{s},\varepsilon;\tilde{C}^i\right)=1)\cdot\alpha_j\cdot\gamma_i\cdot\frac{\overline{s}_k\varepsilon_k b_k}{\overline{s}_i\varepsilon_i}\right].$$

Then, the bidder's problem is to choose b_i to maximize

$$EU_i(b_i;\, b_{-i},\overline{s})\equiv\ v_i\cdot Q_i(b_i;\, b_{-i},\overline{s})-TE_i(b_i;\, b_{-i},\overline{s}). \tag{3.3}$$

We let $EU(b,\overline{s})$ and $TE(b,\overline{s})$ be vector functions where the ith elements are $EU_i(b_i;b_{-i},\overline{s})$ and $TE_i(b_i;b_{-i},\overline{s})$, respectively, for $i=1,..,I$.

We assume that the distributions of the scores have bounded supports. In general, this can lead to a scenario where expected clicks, expenditures and thus profits are constant in bids over certain ranges, since there can be a range of bids that maintain the same average position.

3.4 Existence, Uniqueness, and Computation of Equilibria in the SEU Model

In this section, we derive a particularly convenient representation of the conditions that characterize equilibria in the SEU model, and then we show that standard results from the theory of ordinary differential equations can be used to provide necessary and sufficient conditions for existence and uniqueness of equilibrium. We start by making the following assumption, which we maintain throughout the paper.

ASSUMPTION 1. *The vector of shocks to the scores $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{I+1})$ has the following properties: the components are independent; the distribution of ε_{I+1} is degenerate at 1; the remaining I components are identically distributed with distribution $F_\varepsilon(\cdot)$, which does not have mass points and has an absolutely continuous density $f_\varepsilon(\cdot)$ that has a finite second moment and that is twice continuously differentiable and strictly positive on its support.*

Many of the results in the paper carry over if this assumption is relaxed, but they simplify the analysis substantially.

To begin, we present a simple but powerful identity, proved below in 3.4:

$$\frac{d}{d\tau} EU_i(\tau b, \overline{s})|_{\tau=1} = -TE_i(b, \overline{s}), \quad \text{for } i < I+1, \tag{3.4}$$

that is, a proportional increase in all bids decreases bidder i's utility at the rate $TE_i(b, \overline{s})$, the amount bidder i is spending. The intuition is that ranks and prices depend on the ratios of bids, so a proportional change in all bids simply increases costs proportionally.

The system of first-order conditions that are necessary for equilibrium when all bidders win a strictly positive click share is given by

$$v_i \frac{\partial}{\partial b_i} Q_i(b, \overline{s}) = \frac{\partial}{\partial b_i} TE_i(b, \overline{s}) \quad \text{for all } i. \tag{3.5}$$

Our next result works by combining (3.4) with the first-order conditions, to conclude that a proportional increase in *opponent bids only* decreases utility at the rate $TE_i(b, \overline{s})$; this follows because when bidder i is optimizing, a small change in her own bid has negligible impact.

LEMMA 1 *Assume that $\frac{\partial}{\partial b'_{-(I+1)}} EU(b, \overline{s})$ and TE are continuous in b Then a vector of strictly positive bids b satisfies the first order necessary conditions for equilibrium (3.5) if and only if*

$$\frac{d}{d\tau} EU_i(b_i, \tau b_{-i}, \overline{s})|_{\tau=1} = -TE_i(b, \overline{s}) \text{ for all } i < I+1. \tag{3.6}$$

Proof: Denote the probability of bidders i and k from configuration $C^i \cup \{i\}$ being in positions j and $j+1$ by $G^j_{ik}\left(b, \overline{s}, C^i\right)$. Then $G^j_{ik}\left(b, \overline{s}, C^i\right) = \int \Phi^j_{ik}\left(b, \overline{s}, \varepsilon; C^i\right) dF_\varepsilon\left(\varepsilon\right)$, recalling that Φ^j_{ik} is an indicator for the event that bidder i is in slot j and bidder k is in slot $j+1$. The total quantity of clicks for bidder i can be computed as

$$Q_i\left(b_i, b_{-i}; \overline{s}\right) = \sum_{C^i} \sum_{k \in C^i} \sum_{j=1}^{J} \alpha_j \gamma_i G^j_{ik}\left(b, \overline{s}, C^i\right).$$

The total expenditure can be computed as

$$TE_i\left(b_i, b_{-i}; \overline{s}\right) = b_i \sum_{C^i} \sum_{k \in C^i} \sum_{j=1}^{J} \alpha_j \gamma_i \int \frac{\overline{s}_k b_k \varepsilon_k}{\overline{s}_i b_i \varepsilon_i} \Phi^j_{ik}\left(b, \overline{s}, \varepsilon; C^i\right) dF_\varepsilon\left(\varepsilon\right).$$

Note that $TE_i\left(b_i, b_{-i}; \overline{s}\right)/b_i$ is homogeneous of degree zero in b.

The function $G^j_{ik}\left(b, \overline{s}, C^i\right)$ is homogeneous of degree zero in b as well. As a result, $\sum_{k'=1}^{K} b_{k'} \frac{\partial}{\partial b_{k'}} G^j_{ik}\left(b, \overline{s}, C^i\right) = 0$. Then, the following identity holds

$$\frac{\partial}{\partial b'} EU\left(b, \overline{s}\right) b = -TE\left(b, \overline{s}\right), \tag{3.7}$$

which can in turn be rewritten as, for each $i < I+1$,

$$\frac{d}{d\tau} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right)\Big|_{\tau=1} + b_i \frac{\partial}{\partial b_i} EU_i\left(b_i, b_{-i}, \overline{s}\right) = -TE_i\left(b, \overline{s}\right).$$

Thus, (3.6) is equivalent to $\frac{\partial}{\partial b_i} EU_i\left(b_i, b_{-i}, \overline{s}\right) = 0$ whenever $b_i > 0$.
Q.E.D.

We now build on 3.4 to analyze existence and uniqueness of equilibrium. To do so, we introduce some additional notation.

Let $EU\left(b, \overline{s}\right)$ be the vector of bidder expected utilities, and let $D\left(b, \overline{s}\right)$ the matrix of partial derivatives

$$D\left(b, \overline{s}\right) = \frac{\partial}{\partial b'_{-(I+1)}} EU\left(b, \overline{s}\right).$$

Let $D_0\left(b, \overline{s}\right)$ be the matrix obtained by replacing the diagonal elements of $D\left(b, \overline{s}\right)$ with zeros. Then, (3.6) can be rewritten in matrix notation as

$$D_0\left(b, \overline{s}\right) b_{-(I+1)} = -TE\left(b, \overline{s}\right) - r \frac{\partial EU(b, \overline{s})}{\partial b_{I+1}}.$$

We can then implicitly define a mapping $\beta : [0,1] \to 1$ where, under some regularity conditions imposed on the payoff function, the solution $\beta(\tau)$ will exist in some neighborhood of $\tau = 1$:

$$\tau \frac{d}{d\tau} EU_i\left(\beta_i(\tau), \tau\, \beta_{-i}(\tau), \overline{s}\right) = -TE_i\left(\beta_i(\tau), \tau\, \beta_{-i}(\tau), \overline{s}\right), \tag{3.8}$$

for all bidders $i < I+1$. The next theorem establishes the conditions under which a solution $\beta(\tau)$ exists locally around $\tau = 1$ and globally for $\tau \in [0,\, 1]$. To state the theorem, let $V = [0, v_1] \times \cdots \times [0, v_I]$ be the support of potential bids when bidders bid less than their values, as will be optimal in this game.

THEOREM 1. *Consider a GSP auction in the SEU environment with a reserve price $r > 0$. Assume that D_0 and TE are continuous in b. Suppose that for each i=1,..,I, $Q_i\,(v, \overline{s}) > 0$, and that each EU_i is quasi-concave in b_i on V and for each b the gradient of EU_i contains at least one non-zero element. Then:*

(i) An equilibrium exists if and only if for some $\delta > 0$ the system of equations (3.8) has a solution on $\tau \in [1-\delta,\, 1]$.

(ii) The conditions from part (i) are satisfied for all $\delta \in [0, 1]$, and so an equilibrium exists, if $D_0\,(b, \overline{s})$ is locally Lipschitz and non-singular for all $b \in V$ except a finite number of points.

(iii) There is a unique equilibrium if and only if for some $\delta > 0$ the system of equations (3.8) has a unique solution on $\tau \in [1-\delta,\, 1]$.

(iv) The conditions from part (iii) are satisfied for all $\delta \in [0, 1]$, so that there is a unique equilibrium, if each element of $\frac{\partial}{\partial b'} EU\,(b, \overline{s})$ is Lipschitz in b and non-singular for $b \in V$.

The full proof of this theorem is provided in the Appendix. Quasi-concavity is assumed to ensure that solutions to the first-order condition are always global maxima; it is not otherwise necessary.

Theorem 1 makes use of a high-level assumption that the matrix D_0 is non-singular. In the following lemma we provide more primitive conditions outlining empirically relevant cases where this assumption is satisfied.

LEMMA 2 *(i) For a given $b \in V$ and a given $\overline{s}$, suppose that the bidders are arranged according to their mean score weighted values $\overline{s}_i b_i \geq \overline{s}_{i+1} b_{i+1}$ for $i = 1, \ldots, I-1$. Then $D_0(b, \overline{s})$ is non-singular if for each bidder her utility is strictly locally monotone in the bid of either the bidder above or below her in the ranking or both.*

(ii) Given V, there exist values $\underline{\varepsilon}$ and $\overline{\varepsilon}$ such that (i) is satisfied for all $b \in V$ if the support of ε contains $[\underline{\varepsilon}, \overline{\varepsilon}]$.

Part (i) of 3.4 is satisfied, e.g., if $\frac{\partial EU_i}{\partial b_{i-1}} \neq 0$ for $i = 2, \ldots, I$ and $\frac{\partial EU_1}{\partial b_2} \neq 0$. To see this, note that the diagonal elements of the matrix $D_0\,(b, \overline{s})$ are zero. Therefore, we can compute the determinant

$\det\left(D_0\left(b,\overline{s}\right)\right)=-\frac{\partial EU_1}{\partial b_2}\prod\limits_{i>2}\frac{\partial EU_i}{\partial b_{i-1}}\neq 0$. For part (ii) we note that we can find $\underline{\varepsilon}$ and $\overline{\varepsilon}$ such that for each bidder i we can find bidder i' such that $b_i\overline{s}_i\underline{\varepsilon}>b_{i'}\overline{s}_{i'}\overline{\varepsilon}$ and $b_i\overline{s}_i\overline{\varepsilon}<b_{i'}\overline{s}_{i'}\underline{\varepsilon}$. Then the probability that bidder i' is ranked below bidder i is positive and depends on the bid of bidder i'. Thus, the derivative of bidder i's utility with respect to the bid of bidder i' is not equal to zero.

Equation (3.8) plays a central role in determining the equilibrium bid profile. Now we show that it can be used as a practical device to compute the equilibrium bids. Suppose that functions TE_i and EU_i are known for all bidders. Then, initializing $\beta(0)=0$, we treat the system of equations (3.8) as a system of ordinary differential equations for $\beta(\tau)$. We can use standard methods for numerical integration of ODE if a closed-form solution is not available. Then the vector $\beta(1)$ will correspond to the vector of equilibrium bids.

This suggests a computational approach, which can be described as follows. Suppose that one needs to solve a system of non-linear equations

$$\mathbf{H}(\mathbf{b})=\mathbf{0},$$

where $\mathbf{H}:\mathbb{R}^N\mapsto\mathbb{R}^N$ and $\mathbf{b}\in\mathbb{R}^N$. This system may be hard to solve directly because of significant non-linearities. However, suppose that there exists a function $\mathbf{F}\left(\mathbf{b},\tau\right)$ such that $\mathbf{F}:\mathbb{R}^N\times[0,\,1]\mapsto\mathbb{R}^N$ with the following properties. If $\tau=0$, then the system

$$\mathbf{F}(\mathbf{b},0)=0$$

has an easy-to-find solution, and if $\tau=1$ then

$$\mathbf{F}(\mathbf{b},1)=\mathbf{H}(\mathbf{b})=0.$$

Denote the solution of the system $\mathbf{F}(\mathbf{b},0)=0$ by $\mathbf{b}_0$. If $\mathbf{F}$ is smooth and has a non-singular Jacobi matrix, then the solution of the system

$$\mathbf{F}(\mathbf{b},\tau)=0$$

will be a smooth function of τ. As a result, we can take the derivative of this equation with respect to τ to obtain

$$\frac{\partial\mathbf{F}}{\partial\mathbf{b}'}\dot{\mathbf{b}}+\frac{\partial\mathbf{F}}{\partial\tau}=0,$$

where $\dot{\mathbf{b}}=\left(\frac{db_1}{d\tau},\ldots,\frac{db_N}{d\tau}\right)'$. This expression can be finally re-written in the form

$$\dot{\mathbf{b}}=-\left(\frac{\partial\mathbf{F}}{\partial\mathbf{b}'}\right)^{-1}\frac{\partial\mathbf{F}}{\partial\tau}. \tag{3.9}$$

Equation (3.9) can be used to solve for $\beta(\tau)$. $\beta(0)=\mathbf{b}_0$ is assumed to be known, and $\beta(1)$ corresponds to the solution of the system of equations of interest. Systems of ordinary differential equations are usually easier to solve than non-linear equations.

The computational approach we propose is to define **F** using (3.8). If the payoff function is twice continuously differentiable and the equilibrium existence conditions are satisfied, then **F** has the desired properties. Details of the application of this method to our problem are in Appendix G.

3.5 Bidder Incentives in the SEU Model

It is easier to understand the bidder's incentives in terms of general economic principles if we introduce a change of variables. When bidding, the advertiser implicitly selects an expected quantity of clicks, and a total cost for those clicks. Fix $b_{-i}, \overline{s}$ and suppress them in the notation, and define

$$Q_i^{-1}(q_i) = \inf\{b_i : Q_i(b_i) \geq q_i\},$$

and define

$$\begin{aligned} TC_i(q_i) &= TE_i(Q_i^{-1}(q_i)). \\ AC_i(q_i) &= TE_i(Q_i^{-1}(q_i))/q_i. \end{aligned}$$

Then, the bidder's objective can be rewritten as

$$\max_{q_i} \; q_i(v_i - AC_i(q_i)).$$

This is isomorphic to the objective function faced by an oligopsonist in an imperfectly competitive market. As usual, the solution will be to set marginal cost equal to marginal value, when the average cost curve is differentiable in the relevant range (assume it is for the moment).

$$v_i = q_i AC_i'(q_i) + AC_i(q_i) \equiv MC_i(q_i). \tag{3.10}$$

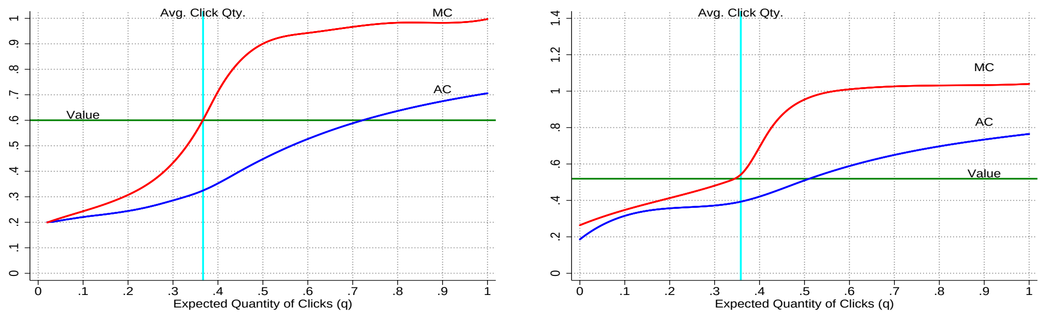
The bidder trades off selecting a higher expected CTR (q_i) and receiving the average per-click profit $v_i - AC_i(q_i)$ on more units, against the increase in the average cost per click that will be felt on all existing clicks. The optimality conditions can be rewritten in the standard way:

$$\frac{v_i - AC_i(q_i)}{AC_i(q_i)} = \frac{d \ln AC_i(q_i)}{d \ln(q_i)}.$$

The bidder's profit as a percentage of cost depends on the elasticity of the average cost per click curve.

To see how this works in practice, consider the following figure, which illustrates the average cost curve $AC_i(q_i)$ and the marginal cost curve $MC_i(q_i)$ for a given search phrase. We select a particular bidder, call it i. Given the actual bid of the advertiser, b_i, we calculate $q_i = Q_i(b_i; b_{-i}, \overline{s})$. We then calculate $MC_i(q_i)$. If the bidder selects q_i optimally, then $v_i = MC_i(q_i)$, as illustrated in the figure. Thus, under the assumption of equilibrium bidding, we infer that the bidder's valuation must have been $MC_i(q_i)$. In Figure 2 we illustrate the structure of the marginal cost and average cost functions for bidders for

Figure 2: Average cost, marginal cost, and value for frequent bidders



(a) Search phrase #1

(b) Search phrase #2

high-value search phrases that most frequently appear in the top position as compared to other bidders. The horizontal line corresponds to the value per click of the considered bidders and the vertical line reflects the quantity of clicks receieved by the bidder provided her actual bid.

This approach to inferring a bidder's valuation from her bid and the average cost curve she faces is the main approach we use in our empirical work. The case where the average cost curve is not differentiable is considered below.

4 Identification of Valuations under Alternative Models

In this section, we consider identification and inference in the following environment. We assume that position-specific click-through rates α_j are known; identification of these is discussed in the appendix.

For the SEU model, we consider observing a large number of queries for a given set of potential bidders, and consider the question of whether the valuations of the bidders can be identified. For each query, we assume that we observe bids, the set of entrants, and the scores. For the NU model, it is more subtle to define the problem, given the disconnect between the model and the real-world bidding environment. The model treats each query as separate, and so in principle, we could imagine that a bidder's valuation changes query to query. In that case, we consider identification of the valuation for each query.

4.1 The No Uncertainty Model

Recall the condition for envy-free Nash equilibrium in the NU model, that score weighted value for bidder j is bounded by incremental cost per clicks $ICC_{j-1,j}$ and $ICC_{j,j+1}$. This implies that observed scores, bids and α_j's are consistent with envy-free Nash equilibrium bidding for some valuations, if and only if

$$ICC_{j,j+1} = \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_j - s_{k_{j+2}} b_{k_{j+2}} \alpha_{j+1}}{\alpha_j - \alpha_{j+1}} \quad \text{is nonincreasing in } j. \tag{4.11}$$

This is a testable restriction of the envy-free Nash equilibrium.

Following Varian (2007), we can illustrate the requirements of the envy-free equilibrium with a figure. Recall Figure 1. The envy-free equilibrium refinement requires that a bidder j selects the position (that is, a feasible click share α_j) that yields the highest value of $s_j v_j \alpha_j - s_j TC_j(\alpha_j)$. This is equivalent to requiring that the score-weighted value is bounded by $ICC_{j,j-1}$ and $ICC_{j,j+1}$.

The requirement that $ICC_{j,j+1}$ is nonincreasing in j corresponds to the total expenditure curve being convex. If (4.11) holds, then we can solve for valuations that satisfy (3.2): we can find score-weighted valuations for each bidder that lie between the steps of the ICC curve. In general, if the inequalities in (3.2) are strict, there will be a set of valuations for the bidder in each position.
Thus, (3.2) determines bounds on the bidder's valuation, as follows:

$$s_{k_j} v_{k_j} \in \left[ICC_{j,j+1}, ICC_{j-1,j}\right].$$

For the highest excluded bidder, $v_{k_{J+1}} = b_{k_{J+1}}$, and for the highest position,

$$s_{k_1} v_{k_1} \in \left[\frac{s_{k_2} b_{k_2} \alpha_1 - s_{k_3} b_{k_3} \alpha_2}{\alpha_1 - \alpha_2}, \infty\right).$$

The EOS equilibrium selection requires $s_{k_j} v_{k_j} = ICC_{j-1,j}$.

4.2 The Score and Entry Uncertainty Model

For the case where $Q_i(\cdot)$ and $TE_i(\cdot)$ are strictly increasing and differentiable in the bid, we can recover the valuation of each bid using the necessary condition for optimality

$$v_i = MC_i(Q_i(b_i)), \tag{4.12}$$

given that all of the distributions required to calculate $MC_i(q_i)$ are assumed to be observable. Note that the local optimality condition is necessary but not sufficient for b_i to be a best response bid for a bidder with value v_i; a testable restriction of the model is that the bid is globally optimal for the valuation that satisfies local optimality. One requirement for global optimality is that the objective function is locally concave at the chosen bid: $MC_i'(Q_i(b_i)) \geq 0$. A sufficient (but not necessary) condition for global optimality is that MC_i is increasing everywhere, since this implies that the bidder's objective function (given opponent bids) is globally concave, and we can conclude that indeed, b_i is an optimal bid for a bidder with value v_i. If MC_i is nonmonotone, then global optimality of the bid should be verified directly.

Now consider the case where $TE_i(\cdot)$ is not differentiable everywhere. This occurs when score uncertainty has limited support, and when there is not too much uncertainty about entry. This analysis parallels the "kinked demand curve" theory from intermediate microeconomics. Note that $Q_i(\cdot)$ is nondecreasing and continuous from the left, so it must be differentiable almost everywhere. If $Q_i(\cdot)$ is constant on $[b_i', b_i'')$ and then increasing at b_i'', then $Q_i^{-1}(q_i) = b_i'$ for $q_i \in [Q_i(b_i'), Q_i(b_i''))$, while $Q_i^{-1}(Q_i(b_i'')) = b_i''$. This implies in turn that $TC_i(\cdot)$ is non-differentiable at $Q_i(b_i'')$, and that $MC_i(\cdot)$ jumps up at that point. Thus, if we observe any b_i on $[b_i', b_i'')$, the assumption that this bid is a best response implies only that

$$v_i \in [MC_i(Q_i(b_i')), MC_i(Q_i(b_i''))]. \tag{4.13}$$

Summarizing:

THEOREM 2. *Consider the SEU model, where bids are fixed over a large number of queries. Suppose that we observe the bids of I bidders $(b_1, .., b_I)$, the joint distribution of their scores s, and the set of entrants in each query. Then:*

(i) Bounds on the valuation for bidder i are given by (4.13), where $b_i' = Q_i^{-1}(Q_i(b_i); b_{-i}, \overline{s})$, and $b_i'' = \sup\{b_i''' : Q_i(b_i'''; b_{-i}, \overline{s}) = Q_i(b_i; b_{-i}, \overline{s})\}$.

(ii) A necessary and sufficient condition for the observed bids to be consistent with ex post equilibrium is that for some $(v_1, ..., v_I)$ within the bounds from (i), the observed bids $(b_1, .., b_I)$ are globally optimal for a bidder solving (3.3). A sufficient condition is that $MC_i(\cdot)$ is nondecreasing for each i.

The proof follows directly from the discussion above and the fact that the functions Q_i and MC_i are uniquely defined from the observed bids and the distribution of scores and entrants.

Equilibria in the SEU environment are not necessarily envy-free, and further, they are not necessarily monotone in the sense that bidders with higher score-weighted valuations place higher score-weighted bids. However, if there are many bidders and substantial uncertainty, each bidder will face a similar marginal cost curve, and monotonicity will follow.

4.3 Comparing Inferences From Alternative Models

A natural question concerns how the inferences from the NU and SEU models compare, given the same auction environment. In this subsection, we show that if the NU model gives bounds on valuations that are consistent across queries (that is, the intersection of the bounds are non-empty), then those bounds will be contained in the bounds from the SEU model. However, in practice, we find that consistency typically fails–the bounds implied by the NU model for one query do not intersect with the bounds from another.

THEOREM 3. *Consider a dataset with repeated observations of search queries, where bids are constant throughout the sample. Consider two alternative models for inference, the NU model and the SEU model. Assume that the NU model produces bounds on valuations that are consistent for a given bidder across the different observations of search queries in the dataset where the advertiser's bid is entered, and consider the intersection of these bounds for each advertiser. This intersection is contained in the bounds on valuations obtained using the SEU model.*

Proof: Fix a vector of bids b and the distributions of scores and entrants. Let $u_i(v_i, b_i'; b_{-i}, \overline{s}_i, \varepsilon, C)$ be the bidder's utility for a particular user query when the bidder's valuation is v_i and bids b_i', and for this proof include explicitly each bidder's valuation as an argument of EU_i. Let v^{NU} be a vector of valuations that is consistent with b being a Nash equilibrium bidding profile in the NU model for all possible realizations of scores and participants. Suppose that v^{NU} is not in the bounds for valuations in the SEU model. Then,

$$\begin{aligned} EU_i(v_i^{NU}, b_i; b_{-i}, \overline{s}) &= \mathbb{E}_{\varepsilon,C}[\max_{b_i'} u_i(v_i^{NU}, b_i'; b_{-i}, \overline{s}_i, \varepsilon, C)] \\ &\geq \max_{b_i'} \mathbb{E}_{\varepsilon,C}[u_i(v_i^{NU}, b_i'; b_{-i}, \overline{s}_i, \varepsilon, C)] \\ &> EU_i(v_i^{NU}, b_i; b_{-i}, \overline{s}). \end{aligned}$$

This is a contradiction. Thus, we conclude that v_i^{NU} is in the bounds for valuations in the SEU model.

5 Estimation of Bidder Valuations

In this section we demonstrate how the expected payoff of a bidder in a position auction can be recovered from the data. The structure of the data for position auctions makes the estimation procedure different from the standard empirical analyses of auctions. In the setup of online position auctions the same set of bids will be used in a set of auctions. In our historical data sample most bidders keep their bids unchanged during the considered time period.

In our data sample a portion of advertisers have multiple advertisements. Bidders submit a separate bid for each ad. Our analysis will be facilitated by the fact that the search engine has a policy of not showing two ads by the same advertiser on the same page. We will use a simplifying assumption that bidders maximize an expected profit from bidding for each ad separately. We associate the bidder with a unique combination of the advertisment and the bid and the goal of our structural estimation will be to recover the latent values per click of the advertisers bidding to place the ads.

Previously we assumed that de-meaned scores have the same distribution across advertisers. We use an additional subscript t to indicate individual user queries with bidder configurations. We assume that configurations C_t of the bidders who were considered for user query t are observed. We assume that the number of advertisers I is fixed and denote $N_i = \sum_t \mathbf{1}\{i \in C_t\}$ the number of queries for which advertiser i was considered. Our further inference is based on the assumption that $N_i \to \infty$ for all bidders $i = 1, \ldots, I$. We denote the total number of user queries in the dataset by T.

We impose the following assumption regarding the joint distribution of shocks to the scores and configurations.

ASSUMPTION 2. *The shocks to the scores are independent from the configurations: $\varepsilon_{it} \perp C_t$ for $i = 1, \ldots, I$. Configurations of advertisers are i.i.d. across queries and the shocks to the scores are i.i.d. across queries and advertisers with expectation $E[\log\, \varepsilon_{it}] = 0$.*

Assumption 2 is used in the identification and estimation of the uncertainty in the score distribution.[5]

To analyze the uncertainty of the scores we use their empirical distribution. In our model for bidder i the score in query t is determined as $s_{it} = \overline{s}_i \varepsilon_{it}$. We note that from Assumption 2 it follows that $E[\log\, s_{it}] = \log\, \overline{s}_i$. Using this observation, we estimate the mean score from the observed realizations of scores for bidder i for impressions t as $\widehat{\overline{s}}_i = \exp\left(\frac{1}{N_i}\sum_t \mathbf{1}\{i \in C_t\} \log\, s_{i,t}\right)$. Then by Assumption 2 and the Slutsky theorem it follows that $\frac{1}{T}N_i \stackrel{p}{\longrightarrow} P(i \in C_t)$. Similarly, we find that $\frac{1}{T}\sum_t \mathbf{1}\{i \in C_t\} \log\, s_{i,t} \stackrel{p}{\longrightarrow} P(i \in C_t) \log\, \overline{s}_i$. Then the consistency of the mean score estimator follows from the continuous mapping theorem.

Then we form the sample of estimated shocks to the scores using $\widehat{\varepsilon}_{it} = \frac{s_{it}}{\widehat{\overline{s}}_i}$. As an estimator for the

[5]The assumption that shocks to scores are i.i.d. across advertisers is actually stronger than what is necessary. There could be a component of the shocks to scores that is common to all advertisements on a specific user query, and the distribution of shocks to scores would still be identified following the literature on measurement error. See Appendix F for details. The assumption that $\varepsilon_{it} \perp C_t$ is explored further from the perspective of the empirical application in Appendix H.

distribution of the shocks to the scores we use the empirical distribution

$$\widehat{F}_{\varepsilon}(\varepsilon)=\frac{1}{I}\sum_{i=1}^{I}\frac{1}{N_i}\sum_{t}\mathbf{1}\{i\in C_t\}\mathbf{1}\left\{\widehat{\varepsilon}_{it}\leq\varepsilon\right\}.$$

Using Assumption 2 and stochastic equicontinuity of the empirical distribution function, the estimator can be expressed by

$$\frac{1}{T}\sum_{t}\mathbf{1}\{i\in C_t\}\mathbf{1}\left\{\frac{\overline{s}_i\varepsilon_{it}}{\widehat{\overline{s}}_i}\leq\varepsilon\right\}=\frac{1}{T}\sum_{t}\mathbf{1}\{i\in C_t\}\mathbf{1}\left\{\varepsilon_{it}-\varepsilon\leq 0\right\}$$

$$+f_{\varepsilon}(0)\frac{\overline{s}_i-\widehat{\overline{s}}_i}{\overline{s}_i^2}\frac{1}{T}\sum_{t}\mathbf{1}\{i\in C_t\}\varepsilon_{it}+o_p(1)=E\left[\mathbf{1}\{\varepsilon_{it}\leq\varepsilon\}\right]P\left(i\in C_t\right)+o_p(1).$$

Combining this with our previous result we find that $\widehat{F}_{\varepsilon}(\varepsilon)$ is a consistent estimator for the distribution of the shocks to the scores $F_{\varepsilon}(\cdot)$.

In the case where the expected payoff function has a unique maximum for each value of the bidder (as we find empirically) the bidder's first-order condition can be used to derive the bidder's valuation, as follows:

$$v_i=\frac{\frac{\partial TE_i(b_i,b_{-i},\overline{s})}{\partial b_i}}{\frac{\partial Q_i(b_i,b_{-i},\overline{s})}{\partial b_i}}.$$

Each of the functions needed to recover the value can be estimated from the data. We use the empirical distribution of the scores to approximate the uncertainty in the scores and use the observed bidder configurations to approximate the uncertainty in bidder configurations. To compute the approximation we take independent samples from the empirical sample of observed configurations and estimated shocks to the scores $\left\{C_t^i,\,\widehat{\varepsilon}_{kt}\right\}_{t,k=1,\ldots,I}$ excluding the bidder of interest i from the sample (recall that we denoted by C^i the configuration excluding bidder i). Following the literature on bootstrap we index the draws from this empirical sample by t^* and denote the simulated sample size T^*. A single draw t^* will include the configuration $C_{t^*}^i$ and the shocks to the scores for all bidders $\widehat{\varepsilon}_{1t^*},\ldots,\widehat{\varepsilon}_{It^*}$. For consistent inference we require that $\frac{N_i}{T^*}\to 0$ for all $i=1,\ldots,I$. Then for each such draw we compute the rank of the bidder of interest i as

$$rank_i\left(C_{t^*}^i\right)=rank\{b_i\widehat{\overline{s}}_i\widehat{\varepsilon}_{it^*}\,;\,b_k\widehat{\overline{s}}_k\widehat{\varepsilon}_{kt^*},\ \forall k\in C_{t^*}^i\}.$$

We also compute the price paid by bidder i as

$$Price_i\left(C_{t^*}^i\right)=\frac{b_k\widehat{\overline{s}}_k\widehat{\varepsilon}_{kt^*}}{\widehat{\overline{s}}_i\widehat{\varepsilon}_{it^*}},\ \text{ for the } k \text{ such that }\ rank_k\left(C_{t^*}^i\right)=rank_i\left(C_{t^*}^i\right)+1.$$

Then we estimate the total expenditure function as

$$\widehat{TE}_i\left(b_i,b_{-i},\overline{s}\right)=\frac{1}{T^*}\sum_{t^*=1}^{T^*}\widehat{\alpha}_{rank_i\left(C_{t^*}^i\right)}\,Price_i\left(C_{t^*}^i\right),$$

and the expected quantity of clicks as

$$\widehat{Q}_i\left(b_i, b_{-i}, \overline{s}\right) = \frac{1}{T^*}\sum_{t^*=1}^{T^*} \widehat{\alpha}_{rank_i\left(C^i_{t^*}\right)}.$$

At the next step we estimate the derivatives. To do that we use a higher-order numerical derivative formula. For a step-size τ_N, depending on the sample size, we compute the implied value as

$$\hat{v}_i = \frac{-\widehat{TE}_i\left(b_i - 2\tau_N, b_{-i}, \overline{s}\right) + 8\widehat{TE}_i\left(b_i - \tau_N, b_{-i}, \overline{s}\right) - 8\widehat{TE}_i\left(b_i + \tau_N, b_{-i}, \overline{s}\right) + \widehat{TE}_i\left(b_i + 2\tau_N, b_{-i}, \overline{s}\right)}{-\widehat{Q}_i\left(b_i - 2\tau_N, b_{-i}, \overline{s}\right) + 8\widehat{Q}_i\left(b_i - \tau_N, b_{-i}, \overline{s}\right) - 8\widehat{Q}_i\left(b_i + \tau_N, b_{-i}, \overline{s}\right) + \widehat{Q}_i\left(b_i + 2\tau_N, b_{-i}, \overline{s}\right)}.$$

The choice of τ_N such that $\sqrt{N_i\,\tau_N} \to \infty$, $\sqrt{N_i}\,\tau_N^3 \to 0$ and $\tau_N \to 0$ for all $i = 1, \ldots, I$ assures that the empirical numerical derivative above converges to the slope of the population marginal cost function. We use this formula to recover the implied valuations. The following result is based on the derivation in Hong et al. (2010) and its proof is given in the Appendix.

THEOREM 4. *Under the sufficient conditions of Theorem 1 and Assumption 2, the derivative of the total expenditure function with respect to the bid vector satisfies the Lindeberg condition and has a finite variance in the limit, while if the derivative of the total quantity of clicks with respect to the bid vector is non-vanishing in the limit our estimator of valuations is asymptotically normal:*

$$\sqrt{N_i\,\tau_N}\left(\widehat{v}_i - v_i\right) \xrightarrow{d} N\left(0, \frac{324\,\Omega}{\left(Q_i'\left(b_i, b_{-i}, \overline{s}\right)\right)^2}\right),$$

where

$$\Omega = Var\left(\frac{u_i\left(v_i, b_i + \tau_N;\, b_{-i},\, \overline{s}_i, \varepsilon_{it}, C_t\right) - u_i\left(v_i, b_i - \tau_N;\, b_{-i},\, \overline{s}_i, \varepsilon_{it}, C_t\right)}{\sqrt{\tau_N}}\right)$$

This shows that with the increasing number of impressions, the estimates of advertiser's valuations will be asymptotically normal and their asymptotic variance will be determined by the variance of the profit per click for the advertiser of interest.

Our analysis extends to the case where the objective function of the bidder can have a set of optimal points. An empirical approach to this case is discussed in Appendix D.

6 Data

For estimation we use a sample of data of auctions for two high-value search phrases (these are among the highest revenue search phrases on the advertising platform). The data is historical, for a three-month period sometime between 2006 and 2009, and it has been preserved for research purposes. The specific time period and the specific search phrases are kept confidential to avoid revealing any proprietary information, and all bids are normalized to a single scale in order to avoid revealing information about the specific revenue of the search phrases. We analyze each search phrase entirely separately, and we compare the results.

We begin with describing the main dataset. There are more than 75,000 searches per week between the two search phrases. We focus on impressions from the first page of advertising results. In the page showing the results of the consumer's search query up to 8 ads are displayed: some in the space above the algorithmic search results and some to the side. In our empirical analysis we control for the position of the advertisement. For consistency of the bidding data with our static analytic framework, we use the data only from one week at a time. However, we compare results across weeks for various specification tests and to validate our general approach.

The following variables are observed for each user query (individual auction): the advertiser account associated with each advertisement; the specific advertisement (characterized by ad text, a bid and a landing page where a user is redirected after clicking on the ad); the positions in which the advertisements were displayed on the screen; the per-click bids and system-assigned scores for the advertisements on the individual query; the per-click prices charged for each advertisement; and the clicks received by each of the advertisements.

A complication that we did not emphasize in the theoretical section is that each advertiser can have multiple active advertisements (possibly with distinct bids) on a given search phrase, while the advertising platform only allows one advertisement per bidder to appear. The different advertisements receive different scores by the system, and thus even if advertiser bids are the same across advertisements, the rotation among different advertisements will create fluctuations in outcomes for opposing bidders. Thus, the variation in advertisements is an important source of uncertainty. They also create complications for thinking about bidder optimization. Why does a bidder have multiple active advertisements, and do the motivations conflict with our assumptions about optimal bidding? In practice, bidders tend to test out variations on advertisements to see whether different ad texts perform better and/or are scored better by the advertising platform.

We chose to handle the multiple advertisements by first treating them separately, and assuming that the advertiser takes the existence of multiple advertisements (and the system's selection among them) as exogenous. Since two advertisements by the same advertiser cannot appear in the same auction, it is possible to treat the advertiser's objective function as additively separable in its bids. We estimate separately the valuations for the different advertisements. We find that valuations and profits are very close for different advertisements by the same bidder. In the empirical application, two-thirds to three quarters of advertisers have only one advertisement, while 90quarters of ad impressions come from advertisers with less than three advertisements, and more than three quarters of advertisers have most of their ad imporessions with the same advertisement (Herfindahl indices of more than .85, measuring the "market shares" of the advertisements in terms of user queries).

Another complication that arises with our data is that in our research data set, we observe only the ads that were actually displayed. We also can infer the product of the bid and the quality score for the last ad that was **not** displayed, because it is used to set the price for the last displayed ad. The missing data potentially creates problems for our estimation, because shocks to the quality scores of excluded ads can potentially put them onto the screen, but without knowledge of these advertisements, we exclude that possibility. Another problem is that it can bias our estimation of the distribution of quality scores, because very low draws of quality scores might push an advertisement off of the page, removing it from our sample. In Appendix H we evaluate the biases created by this problem, and find that they are economically and statistically insignificant.

7 Estimates from Alternative Models

We use several alternative models for estimation: NU-EOS, NU-EFLB, and SEU.

7.1 Modeling Details for NU Models

In the NU Models, we treat each auction as separate, envisioning that bids and valuations might change from auction to auction. We then empirically characterize whether bounds on valuations are consistent across auctions, and how implied valuations change over time. In particular, for each auction we recover the bounds on valuations using the constructed $ICC_{j,j+1}$ curves for positions. We notice that in a large fraction of cases the ICC curve fails to be monotone auction-by-auction. Varian (2007) suggested computing an approximate weighted solution. We consider a weighted ICC as

$$ICC^{d}_{j,j+1} = \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_j d_j - s_{k_{j+2}} b_{k_{j+2}} \alpha_{j+1} d_{j+1}}{\alpha_j d_j - \alpha_{j+1} d_{j+1}},$$

where weights minimize $\sum\limits_{j=1}^{J} (1-d_j)^2$ such that a weighted ICC is monotone. In the empirical study we perform this procedure for all considered search phrases. We recover the values of the advertisers from the re-weighted ICC curve as

$$s_{k_j} v_{k_j} \in \left[ICC^{d^*}_{j,j+1}, ICC^{d^*}_{j-1,j} \right],$$

where weights d^* solve the minimization problem above. We abuse notation and omit the index of user query t that should subscript the weights and the score. The selected weights are tailored to each specific auction and vary auction by auction.

Similar to Varian (2007) we find a large number of violations from monotonicity, and we correct them using weighting. We also find that the bounds on valuations fluctuate substantially in the NU models. The fluctuations occur from query to query, oscillating back and forth between bounds for commonly observed sets of entrants and scores, so that it is difficult to imagine rationalizing the fluctuations on the basis changing valuations (and bids do not change at that frequency, and often don't change at all). The median standard deviation of the recovered value for a single bid across queries ranges from approximately 11% to 23% for the NU-EFLB model and approximately from 18% to 30% for the NU-EOS model. Moreover, the number of auctions that violate the value monotonicity auction-by-auction exceeds 25% with most violations occurring in the middle and the lowest positions.

The weights aimed at making the ICC curves monotone vary from auction to auction, depending on how far a particular configuration is from the configurations with the monotone ICC.

Throughout, in our analysis of the NU models, we use the variant where ICC curves are adjusted to be monotone using this approach.

Now consider the estimation of values. Across all of the advertisements that have auctions for which the monotonicity of the implied score-weighted values is not violated, we could not find any examples where the bounds have a non-empty intersection. Even restricting the dataset to a very limited period of time (2 hours) allows us to find only 5% of advertisements where the bounds intersect. In this case the number of observations per advertisement ranges from 1 to a few hundred.

To address the issue of non-overlapping bounds, we take the following approach. The set of recovered values corresponds to the bounds constructed from the incremental cost curve, with the lower bound corresponding to NU-EFLB and the upper bound corresponding to NU-EOS. For each bidder we can collect a set of values corresponding to different user queries. We use the median over different values of the lower and upper bounds for each bidder as estimates of valuations from the full-information models NU-EFLB and NU-EOS, respectively (the median is used rather than the mean, to reduce the sensitivity of our results to assumptions we make about the upper bound of valuations for the NU-EOS model for the bidder in the top position, where the upper bound is not identified from the data). We should note that this approach may result in negative implied profit per click for some queries, since we infer values using the median values across user queries.

7.2 Modeling Details for the SEU Model

The exact procedure for estimation of the first-order condition has been described in the previous section. We observe empirically that the SEU model yields very tight bounds or point estimates for almost all advertisers. As a result, we will focus on the lower bound of SEU valuations, and refer to them as if they are unique. We already illustrated in Figure 2 the estimated average cost curves, marginal cost curves, and implied valuations for an individual bidder for a given search phrase. The figures illustrate how valuations are inferred from bids: the vertical line shows the expected CTR the bidder attains with the bid she places in the data, and the place where that line intersects the marginal cost curve defines the implied valuation for this bidder on this search phrase.

7.3 Empirical Results

We find empirically that estimated marginal cost curves are strictly increasing for each of the observed advertisements on both search phrases, which implies that the implied valuations and bids comprise an ex post Nash equilibrium in the SEU model. We formally test this by considering a grid of bids (with 600 grid points). At each point we test the hypothesis that the marginal cost for a sample of score realizations is equal to zero. This test rejected the null at 5% level for all grid points and phrases (we constructed the grid such that the maximum bid guarantees that the bidder always attains the highest position, corresponding to the maximum achievable clickthrough rate).

Using our three alternative models, we recover valuations for all advertisements featured in the auctions in the selected week of data for a selected search phrase. We present our results normalizing recovered valuations and profits per click using the mean of the bid for the search phrase # 1 as a numeraire. We use the same normalizing factor for the values and profits per click recovered for both considered search phrases. In Table 1 we display basic statistics for log valuations for three analyzed search phrases. We notice that the search phrase #2 is the higher value phrase out of the two phrases that we analyze. However, the range of valuations remains comparable across both search phrases. Assuming that the valuations corresponding to different advertisements are stable within the period of analysis, we can compute the standard errors for the recovered values using the asymptotic formula. It turns out that the recovered valuations have very tight standard errors due to large number of auctions in the sample.

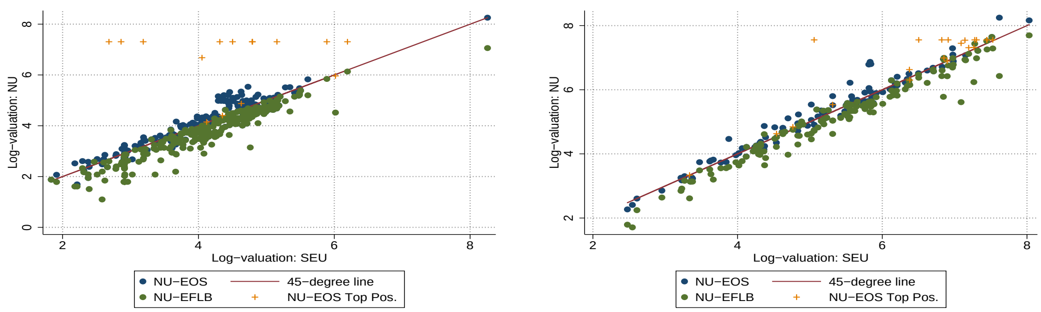
Table 1: Mean of valuations for different models across search phrases

	All bidders		Top bidders in >5% queries excluded	
Model; Search phrase	#1	#2	#1	#2
NU-EFLB	0.043	0.229	0.041	0.181
NU-EOS	0.100	0.350	0.063	0.255
SEU	0.063	0.287	0.061	0.226

We report the means of recovered valuations across search phrases and bidders for both of the models, where for the NU models we use the modified (monotone) ICC curves. The values are normalized by the highest observed bid for search phrase #1.

Not surprisingly, the values computed from the NU-EFLB model in columns 1-2 of Table 1 tend to under-estimate the values recovered in the SEU framework while the values computed from the NU-EOS framework over-estimate the SEU values. We also notice that a substantial disadvantage of the NU-EOS framework is that it does not provide a meaningful upper bound for the bidders that are in the top position. To compute the numbers in the first two columns of Table 1 we imputed an upper bound on valuations equal to the largest bid in the sample for each search phrase. The results turn out to be sensitive to the definition of the upper bound on valuations in the NU-EOS framework and in columns 3-4 of Table 1 we report the mean log-values omitting the bidders who appear in the top position in more than 5% cases. As expected, the average values for this group of bidders are smaller than overall mean values, while the average estimated values in NU-EFLB and NU-EOS frameworks are closer to the estimated values in the NU framework.

Figure 3: Log of Valuations in NU-models Versus SEU Model (with indicated imputed NU-EOS values)



(a) Search phrase #1

(b) Search phrase #2

Co-location of estimated values can be represented graphically. Figure 3 displays the implied valuations for alternative models (or their bounds) for all advertisements against the implied valuations from the SEU model, in logarithmic scale, for both search phrases. NU-EFLB underestimates the values for most of the advertisements, while the NU-EOS overerestimates the values. To understand why, recall first from Theorem 3 that if the NU models has an interval of valuations that is in the bounds on valuations across all queries, then those valuations will also be within the bounds for the SEU model. Since there is no such interval in our data for any advertiser, thus it is an empirical question as to how the NU model bounds will relate to the SEU valuations. We separately indicate the values where the median value was implied from the upper bound on the values for the top bidders in NU-EOS framework (this occurs for bidders in the top position for more than half the time that they appear on a query). Figure 3 demonstrates that these imputed values can induce large biases in the estimated average values.

Combining the recovered values with the data, we can compute the implied ex-post profits per click

Table 2: Lerner Index $\frac{\text{Avg.}(value - CPC)}{\text{Avg.}(CPC)}$ (advertisement is unit of analysis)

	All bidders				Top bidders in >5% queries excluded			
Model	Mean	25%	50%	75%	Mean	25%	50%	75%
Search phrase #1								
NU-EFLB	0.231	0.010	0.129	0.341	0.22	0.006	0.109	0.341
NU-EOS	2.947	0.180	0.707	1.546	0.892	0.162	0.631	1.430
SEU	0.789	0.337	0.479	0.950	0.773	0.335	0.476	0.950
Search phrase #2								
NU-EFLB	0.394	0.116	0.328	0.595	0.335	0.078	0.272	0.514
NU-EOS	1.300	0.426	0.661	1.214	0.914	0.384	0.553	0.938
SEU	0.795	0.405	0.594	0.988	0.693	0.371	0.543	0.865

across the bidders by averaging the per-impression profit per click across different impressions. In the NU framework, the value obtained under the NU-EFLB assumption under-estimates the valuation and the value obtained under the NU-EOS assumption over-estimates the valuation; the same comparisons hold for profit per click.

We begin with the analysis of the profit per click per advertisement. Table 2 illustrates per-query profit per click relative to the average cost per click for each of the different models (the Lerner Index for the advertiser), weighting each advertisement equally. We compute each advertiser's Lerner Index by weighting both the profit per click and the cost per click by the score and the position-specific clickthrough rate.

We can note from columns 1-4 of Table 2 that the surplus predicted by SEU and NU framework can be substantially different. The largest deviations from the surplus in SEU framework is demonstrated by the surplus evaluated from NU-EOS framework. We have recognized that this result is very sensitive to the value estimates for the top bidders (whose NU-EOS values are inferred from the upper bound on the distribution of the values). To isolate this effect, in columns 5-8 of Table 2 we report the profit PC excluding the bidders that appear in the top position in more than 5% of queries. It clear that the last four columns of Table 2 provide a more reasonable base for comparison between the frameworks, showing that NU-EOS framework tends to over-estimated the Lerner Index relative to SEU framework, while NU-EFLB tends to underestimate the Lerner Index. However, NU-EOS Lerner Indices are more spread out across advertisers. The magnitude of the Lerner Index remains very close across the considered search phrases.

8 Counterfactual experiments

8.1 Alternative Models and the Role of Uncertainty

We begin by looking at how the alternative models do in terms of predicting behavior out of sample. We proceed by taking implied valuations from each model using one week of data (taking the valuations corresponding to the mean values for the NU models), and then predicting revenue on an auction-by-auction basis in the next week of data using the same model to generate counterfactual predictions. For the advertisements which do not appear in the first week of data, we hold the counterfactual bids equal to the observed bids. Figure 4 illustrates the results, where the x-axis is the expected revenue given the actual bids and prices in the auction, and the y-axis shows the predictions (or bounds on predictions) for the SEU model. Note that the SEU model provides a very good fit for the data. One reason for that is that the sample of advertisers and their bids do not change substantially from week to week, creating a similar competitive environment for the bidders in each week. As a result, our model predicts very similar bids for the same advertisers in the second week of data.

Details of the computational algorithm for the SEU model are given in Appendix G.

Figure 4: Log of Predicted Revenues v. Log of Actual Revenues for Search Phrase 1

(a) SEU Model (b) NU-EFLB (c) NU-EOS

Note that NU-EOS and NU-EFLB produce bounds on valuations when drawing inferences, and then each valuation profile generates a range of equilibria, expanding again the range of possible outcomes in the prediction. The revenue predicted by the NU-EFLB model tends to understate the actual revenue. On the other hand, the revenue predicted by the NU-EOS model tends to under-state the revenue. In most cases, however, when the NU bounds straddle the actual revenue, the revenue in the SEU case remains is also the bounds.

Table 3: Mean squared deviation of the predicted revenues per query from the true revenues (normalized by true mean revenues per query)

Model for equilibrium (Model for values)	Mean	25%	50%	75%
Search phrase #1				
SEU (SEU)	0.078103	0.000287	0.002876	0.068488
NU-EFLB (SEU)	0.152973	0.005861	0.036122	0.164296
NU-EOS (SEU)	0.163367	0.011001	0.058934	0.178699
NU-EFLB (NU-EFLB)	0.204177	0.013794	0.098821	0.251827
NU-EOS (NU-EOS)	0.189599	0.037119	0.089006	0.200281
Search phrase #2				
SEU (SEU)	0.115878	0.000024	0.002207	0.01245
NU-EFLB (SEU)	0.155453	0.003919	0.021243	0.089429
NU-EOS (SEU)	0.154876	0.003709	0.021950	0.091917
NU-EFLB (NU-EFLB)	0.200494	0.010863	0.045751	0.158933
NU-EOS (NU-EOS)	0.155696	0.004090	0.026585	0.107583

Table 3 describes the distribution of the standard deviations of the predicted revenues from the actual revenue in week 2 normalized by the mean actual revenues.

8.2 Competition, Elasticities and Profits PC

In this section, we examine the properties and implications of the estimated elasticities for the average cost curve for clicks. First, we observe that there is substantial variation across bidders and across search phrases in the elasticity of the average cost curve. Table 4 provides summary statistics on the inverse of the elasticity faced by all the advertisements, grouping bidders together by the average ranking the advertisements received. The table also shows the gaps between value and bid, and between bid and payment, each normalized by the bid, for bidders in each category (recall that the Lerner Index $\frac{\text{Value-CPC}}{\text{CPC}}$ will be equal to the inverse of the elasticity). The gap between bid and payment is large and implies that the bids substantially exceed the payment. The bid tends to be between .43 and .83 of the value for all positions, typically around two-thirds of the value on search phrase #2. The results for the second position for search phrase #1 are skewed by an unusually high value estimate by a bidder who dominates the second position.

Table 4: Characteristics of competition

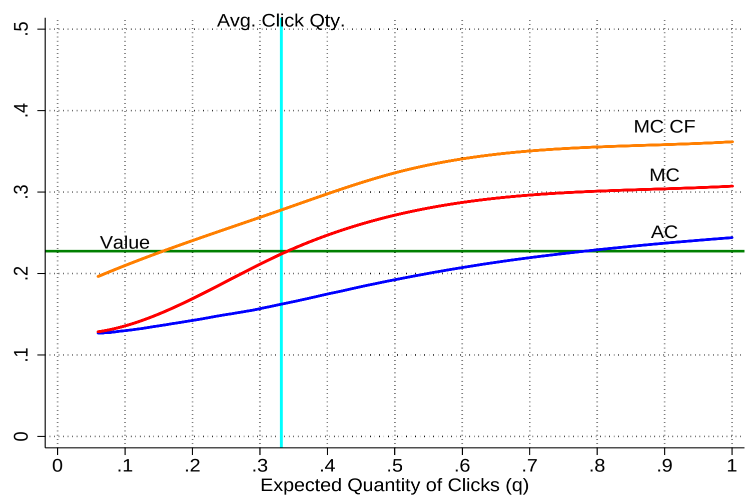
Avg. ranking	Mean $\frac{\text{Value-Bid}}{\text{Value}}$	Mean $\frac{\text{Value-Bid}}{\text{CPC}}$	Mean $\frac{\text{Bid-CPC}}{\text{CPC}}$	Inverse Elasticity Mean	Inverse Elasticity 25%	Inverse Elasticity 50%	Inverse Elasticity 75%
Search phrase #1							
$[1, 1.5)$	0.182264	0.583629	1.607820	2.191449	2.245929	2.245929	2.245929
$[1.5, 2.5)$	0.572825	3.708380	1.301450	5.009830	5.650886	5.650886	5.650886
$[2.5, 4)$	0.191422	0.347059	0.409290	0.756349	0.628711	0.628711	1.039009
$[4, 5.5)$	0.225535	0.463235	0.507975	0.971210	0.768171	0.945691	1.174055
$[5.5, 8)$	0.164800	0.307044	0.452509	0.759553	0.607790	0.648143	0.727018
Search phrase #2							
$[1, 1.5)$	0.339828	0.758487	0.431421	1.189908	0.797125	1.288512	1.288512
$[1.5, 2.5)$	0.337085	0.753351	0.462675	1.216027	1.167568	1.213478	1.383097
$[2.5, 4)$	0.328453	0.719037	0.454364	1.173401	1.152048	1.221465	1.306378
$[4, 5.5)$	0.315977	0.658021	0.385349	1.043370	0.903537	1.054401	1.229759
$[5.5, 8)$	0.305704	0.660775	0.417593	1.078368	0.799619	1.032054	1.099350

We report mean elasticities of the MC curve from the SEU model corresponding to bidders whose average position is in the displayed bracket. The statistics are first computed at the advertisement level, and then weighted by the frequency of user queries in which they appear in the relevant position buckets.

We study the impact of competition on advertiser behavior and profits using a counterfactual experiment. We focus on a high-value bidder who appears most frequently in the top 3 positions, and consider increasing the number of rival bidders by 20%. We want to study the impact of this event on the high-value (and, hence, high-revenue) bidders, social welfare and the revenue of the auction platform.

The new bidders and their values and average scores are generated randomly from the empirical distribution of bidders in the bottom positions, and the score shocks are generated from the score distribution. We illustrate the impact of the new bidders on a high-value bidder on Figure 5. The entry of new bidders tends to increase the marginal cost uniformly over positions leading to the shift of the marginal cost up (to MC CF). We quantify the effect of rival entry by considering, first, the best response of a high-value bidder to the changing number of rivals, and, second, we re-compute the bids in the SEU environment allowing all bidders to re-optimize their bids.

Figure 5: An impact of rival entry on a high-value bidder



8.3 Auction Design: Comparing Vickrey Auctions and the Generalized Second Price Auction

In a model without uncertainty, EOS and Varian have shown that the EOS equilibrium implements the same allocation and the same prices as a Vickrey auction. Thus, the choice of auction design does not matter. However, once real-world uncertainty is incorporated, this equivalence breaks down. If the auctioneer held a separate Vickrey auction for each user query, it would be optimal for each advertiser to bid its value, even if the same bid was to be applied across many different user queries. If we take the quality scores calculated for each impression as the best estimate of the efficient scores (that is, efficiency requires ads to be ranked according to the product of value and quality score), then the ads will always be ranked efficiently, query by query, in the Vickrey auction, even as quality scores change over time.

In contrast, in the GSP auction used in practice, if bids apply to many queries (as in the SEU model) and scores and entry vary across queries, then different bidders will have different gaps between their bids and values. This implies that the ads will not be ranked efficiently in many cases. Therefore, the GSP auction is strictly less efficient than the Vickrey auction, so long as there is sufficient uncertainty in the environment.

Table 5 shows the results of a counterfactual comparison of the two mechanisms. We used the values estimated in the SEU model, and computed counterfactual equilibria in each auction format: Vickrey and GSP auction. To simplify, we ignored reserve prices, which were rarely binding in any case. Note that the Vickrey auction gives the same results as if the NU-EOS model is used, since in a world where bidders change their bids to play the NU-EOS equilibrium in each query, the allocation and prices are the same as Vickrey prices. The SEU model equilibrium gives the outcome of the generalized second price auction under uncertainty.

We see that the Vickrey auction always gives higher efficiency, which is necessarily the case, and the efficiency differences are small.

The revenue comparison between the mechanisms is theoretically ambiguous, so it is an empirical question as to which one performs better. We see that for search phrase 1, the revenue differences are larger in magnitude than the efficiency differences, in the same direction: the Vickrey auction is more efficient, and raises 3.4% more revenue.

In contrast, for search phrase 2, the Vickrey auction raises 11% less revenue, despite being 0.6% more efficient.

Thus, we see a benefit of using the structural model to obtain estimates of values and the distribution of quality scores in the environment: we can do counterfactual experiments to compare auction designs in a scenario where theory is ambiguous about the revenue comparison. Our estimates show that the efficiency gains from a Vickrey auction are small for some search phrases, but more substantial for others, and that the revenue comparison will likely vary from search phrase to search phrase. Thus, further research is required to assess the best choice for the platform as a whole from a revenue perspective, while from an efficiency perspective, Vickrey auctions offer the potential for modest improvements.

Table 5: Predicted revenues and welfare for the SEU generalized second price auction model versus the NU-EOS model (equivalent to query-by-query Vickrey auctions), using SEU values and actual bidder configurations in both cases

		Percentiles		
Model (values)	Mean	25%	50%	75%
Search phrase #1				
Revenue SEU (SEU)	235.004970	248.280609	259.453125	267.584167
Revenue Vickrey=NU-EOS(SEU)	243.040010	232.857819	263.601868	285.036224
Welfare SEU (SEU)	760.519819	837.780518	862.425690	875.496826
Welfare Vickrey=NU-EOS (SEU)	778.474910	852.110107	871.899078	885.075439
Search phrase #2				
Revenue SEU (SEU)	317.688514	267.728424	329.609161	408.292572
Revenue Vickrey=NU-EOS(SEU)	283.921230	223.360092	315.655853	366.069519
Welfare SEU (SEU)	650.797979	548.630920	738.566711	805.667175
Welfare Vickrey=NU-EOS (SEU)	654.896761	550.820801	751.610657	810.383545

This table represents the expected total per query revenue of the auction platform and the expected per query social welfare. To compute the expected revenue we used the estimated position clickthrough rates and the scores of the advertisers. To compute the expected welfare we used the values estimated from the SEU model and the estimated position clickthrough rates and scores of the advertisers. Values and cost per click are normalized by the maximum bid for the search phrase # 1.

9 Conclusion

In this paper we develop and estimate a new model of advertiser behavior under uncertainty in the sponsored search advertising auctions. Unlike the existing models which assume that bids are customized for a single user query, we model the fact that queries arrive more quickly than advertisers can change their bids, and advertisers cannot perfectly predict quality scores. We present theoretical characterizations of existence and uniqueness, and propose a computational algorithm for computing equilibria given primitives of the model. We develop an estimator for bidder valuations, establish its properties, and apply it to historical data from Microsoft's search advertising auctions. Our model yields lower implied valuations and bidder profits than approaches that ignore uncertainty.

The empirical application provides insight into the economics of search advertising auctions. We find that bidders have substantial profits per click, even on some of the industry's most competitive search phrases: bidder values are typically 33% to 100% higher than their cost per click, even on very competitive search phrases. Further, profits vary with the equilibrium position on the page, and on the search phrases we study, they appear to be higher in the top positions, for a few high-value advertisers. Profits are determined by the rate at which clicks decline with the position on the page, and by the dispersion of bidder values around a given bidder's value.

We find that bidders have substantial strategic incentives to reduce their expressed demand in order to reduce the unit prices they pay in the auctions, and these incentives are asymmetric across bidders, leading to inefficient allocation. We quantify the inefficiency as being fairly small (less than 3We show that a Vickrey auction would eliminate the inefficiency, but the impact of switching to a Vickrey auction for revenue is ambiguous. Since equilibrium Vickrey bids (equal to bidder values) are much higher than GSP bids (33% to 100%), unless a mechanism was devised to transition GSP bids automatically, there could be substantial transition costs from changing the mechanism as well.

We also show how important it is to account for bidder responses when a search advertising platform contemplates a change: in a counterfactual experiment, bidders increased their bids in response to an increase in competition. Thus, structural models like the one in this paper can be useful for forecasting the longer-term impact of changes to the ad platform.

References

Agarwal, N. and Athey, S. and Yang, D. (2009): "Skewed Bidding in Pay-per-Action Auctions for Online Advertising", American Economic Review: Paper and Proceedings, 441–447.

Athey, S. and Ellison, G. (2009): "Position auctions with consumer search", Harvard and MIT Working paper.

Athey, S., and P. Haile (2002): "Identification of standard auction models", Econometrica, 2107—2140.

Borgers, T. and Cox, I. and Pesendorfer, M. and Petricek, V. (2007): "Equilibrium bids in sponsored search auctions: Theory and evidence", Working paper.

Boltyanskii, V., R. Gamkrelidze, and L. Pontryagin (1960): "Theory of optimal processes. I. The maximum principle", Izv. Akad. Nauk SSSR Ser. Mat, 24(1), 3—42.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Parameter set inference in a class of econometric models", Econometrica, 75, 1243—1284.

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(1), 242—259.

Edelman, B., and M. Ostrovsky (2007): "Strategic bidder behavior in sponsored search auctions", Decision support systems, 43(1), 192—198.

Ghose, A., and S. Yang. (2009): "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets", Management Science, 55 (10), 1605—1622.

Haile, P., and E. Tamer (2003): "Inference with an incomplete model of English auctions", Journal of Political Economy, 111(1), 1—51.

Hendricks, K. and Paarsch, H.J. (1995): "A survey of recent empirical work concerning auctions", The Canadian Journal of Economics, 28(2), 403—426.

Hashimoto, T. (2009): "Equilibrium Selection and Inefficiency in Internet Advertising Auctions", SSRN Working paper.

Hong, H., Mahajan, A., and D. Nekipelov (2010): " Extremum Estimation and Numerical Derivatives", Stanford and UC Berkeley Working paper.

Imbens, G., and C. Manski (2004): "Confidence intervals for partially identified parameters", Econometrica, 1845—1857.

Jeziorski, P. and I. Segal, "What Makes them Click: Empirical Analysis of Consumer Demand for Internet Search Advertising," working paper, Stanford University, February 2009.

Kosorok, M. (2008): "Introduction to empirical processes and semiparametric inference", Springer.

McFadden, D., and W. Newey (1994): "Large Sample Estimation and Hypothesis Testing", Handbook of econometrics, 4, 2111—2245.

Mela, C.F., and Yao, S. (2009): "A dynamic model of sponsored search advertising", SSRN Working paper.

Paarsch, H.J. (1997): "Deriving an estimate of the optimal reserve price: an application to British Columbian timber sales", Journal of Econometrics, 78(2), 333—357.

Pontryagin, L. (1966): "On the theory of differential games", Russian Mathematical Surveys, 21(4), 193—246.

Pollard, D. (1990) "Empirical processes: theory and applications", *Institute of Mathematical Statistic.*

Varian, H. (2007): "Position auctions", International Journal of Industrial Organization, 25(6), 1163—1178.

Appendix

A Proof of Theorem 1

Throughout the proof, we abuse notation by writing $\frac{\partial}{\partial b_j} EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right)$for $\frac{\partial}{\partial b_j} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right)\Big|_{b=\beta(\tau)}$. We start with proving parts (i) and (iii). First, we prove the sufficiency of these conditions. Suppose that for some $\delta > 0$ with $\tau \in [1-\delta, 1]$, there exists a (unique) solution $\beta(\tau)$ to the equation

$$\tau \frac{d}{d\tau} EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right) = TE_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right), \text{ for all } i. \tag{A.14}$$

If $\delta = 1$, we define at the origin

$$\frac{\partial}{\partial b_j} EU_i\left(0, 0, \overline{s}\right) = \lim_{\varepsilon \downarrow 0} \frac{\partial}{\partial b_j} EU_i\left(\varepsilon, 0, \overline{s}\right). \tag{A.15}$$

Lemma 1 establishes that (A.14) holds at $\tau = 1$ if and only if the bidders' first-order conditions hold. The results of Lemma 1 apply to the case where the auction has a positive reserve price.When the reserve price is equal to r, then both the expected utility and the total expenditure become functions of r. Homogeneity of the utility function will also be preserved when we consider the vector of bids accompanied by r. As a result, equation (3.7) will take the form

$$\frac{\partial}{\partial b'} EU\left(b, \overline{s}, r\right) b + \frac{\partial}{\partial r} EU\left(b, \overline{s}, r\right) r = -TE\left(b, \overline{s}, r\right). \tag{A.16}$$

As a result, we can re-write our key equation as

$$\frac{d}{d\,\tau} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right)\big|_{\tau=1} = -TE_i\left(b, \overline{s}\right) - r\frac{\partial}{\partial r} EU_i\left(b, \overline{s}, r\right). \tag{A.17}$$

Our results for τ in the neighborhood of $\tau = 1$ will apply with total expenditure function corrected by the influence of the reserve price. In the further analysis we can simply use the modified total expenditure function

$$\widetilde{TE}_i\left(b_i, \tau b_{-i}, \overline{s}\right) = TE_i\left(b, \overline{s}\right) + r\frac{\partial}{\partial r}EU_i\left(b, \overline{s}, r\right). \tag{A.18}$$

In the case where the vector of the payoff functions has a non-singular Jacobi matrix globally in the support of bids, we can also extend the results for $\tau \in [0, 1]$ to the case with the reserve price. In this case, the initial condition for $\tau = 0$ will solve

$$\widetilde{TE}_i(\beta_i(0), 0, \overline{s}) = 0.$$

Note that for all bidders $i = 1, \ldots, N$ this is a non-linear equation with a scalar argument $b_i(0)$, which can be solved numerically. This will allow us to construct a starting value for the system of differential equations. The solution $\beta(\tau)$ exists at $\tau = 1$ by assumption. Due to quasi-concavity of the objective function, it will correspond to the maximum of the payoff function. This means that there will exist an equilibrium in the considered auction b^* corresponding to $\beta(1)$. This proves the sufficiency.

Second, we prove necessity. Suppose that there exists an equilibrium vector of bids b^*. Then it solves the system of the first-order conditions

$$\frac{\partial EU_i\left(b_i^*, b_{-i}^*, \overline{s}\right)}{\partial b_i} = 0.$$

Define the mapping $\beta(\tau)$ such that

$$\frac{\partial EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right)}{\partial b_i} = 0, \tag{A.19}$$

which coincides with the system of the first-order conditions at $\tau = 1$ meaning that $\beta(1) = b^*$. We prove the existence of such mapping by the following manipulations. Due to the smoothness of the objective function, if the mapping exists, it is continuous. From homogeneity of $Q_i(\cdot)$ function and $TE_i(\cdot)/b_i$ (established in the proof of Lemma 1), it follows that

$$\sum_j b_j \frac{\partial}{\partial b_j} EU_i\left(b_i, b_{-i}, \overline{s}\right) = -\widetilde{TE}_i\left(b_i, b_{-i}, \overline{s}\right), \tag{A.20}$$

for any b in the support of these functions (with the derivative of the payoff function continuously extended to the origin by (A.15)). Function $\widetilde{TE}_i(\cdot)$ is defined in (A.18).

In particular, the support of bids includes all vectors $(b_i, \tau b_{-i})$ for some $\delta > 0$ and $\tau \in [1 - \delta, 1]$. Given that (A.20) is a direct consequence of homogeneity, it will be satisfied for any τ and any b in the support of bids

$$b_i \frac{\partial}{\partial b_i} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right) + \sum_{j \neq i} \tau b_j \frac{\partial}{\partial b_j} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right) = -\widetilde{TE}_i\left(b_i, \tau b_{-i}, \overline{s}\right). \tag{A.21}$$

This equation will also be valid for $\beta(\tau)$ defined by (A.19) (if it exists). Substituting $b = \beta(\tau)$ into (A.21), we conclude that

$$\sum_{j\neq i} \tau b_j(\tau)\frac{\partial}{\partial b_j} EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right) = -\widetilde{TE}_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right)$$

is equivalent to the definition of $\beta(\tau)$ by (A.19). This can be re-written as

$$\tau\frac{d}{d\tau} EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right) = -\widetilde{TE}_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right), \text{ for all } i.$$

This equation has solution $\beta(1) = b^*$ by our assumption and equation (A.19). By our assumption, the Jacobi matrix for the vector of payoffs is non-singular at $\tau = 1$ while each $\widetilde{TE}_i(\cdot)$ is continuous. By Boltyanskii et al. (1960) this means that the differential equation (A.14) has a continuous solution in some neighborhood of $\tau = 1$. This proves the necessity of the statement.

Thus, we have proved that existence and uniqueness of the equilibrium bid vector is equivalent to the existence and uniqueness of the solution to the differential equation (A.14). This establishes (i) and (iii) in Theorem 1.

Now we proceed with proving (ii) and (iv) and establish the result for the global existence of the solution to (A.14) under stronger conditions for the payoff functions. Assume that $D_0(b, \overline{s})$ is locally Lipschitz and non-singular. From equation (3.8) for each τ we will be able to find bid vectors $\beta(\tau)$ which solve the system (A.19), which will transform to the system of equilibrium first-order conditions for $\tau = 1$. We can now verify that the vector of bids $\beta(0) = 0$ solves the system of differential equation (3.8) corresponding to $\tau = 0$. This will allow us to characterize the equilibrium as a solution to differential equation (3.8) with the initial value $\beta(0) = 0$. Bidder i's cost will be equal to zero if all other bidders bid zero. Therefore, for all b_i in the support of bids $TE_i\left(b_i, 0, \overline{s}\right) = 0$. As a result, $\beta(0) = 0$ will solve equation (3.8) for $\tau = 0$. From our previous result, it follows that $\beta(1)$ is the equilibrium vector of bids. Equation (3.8) states that for the mapping $\beta(\tau)$ that is defined by the first-order conditions for all bidders and all $\tau \in [0, 1]$ satisfies

$$\tau\frac{d}{d\tau} EU_i\left(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s}\right) = -\widetilde{TE}_i\left(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s}\right).$$

Given that $\beta(\tau)$ is a function of τ, we can apply the chain rule and express the total derivative of the payoff function in terms of the derivative of $\beta(\tau)$ with respect to τ:

$$\sum_j \left(\tau + (1-\tau)\mathbf{1}_{j=i}\right)\frac{\partial}{\partial b_j} EU_i\left(b_i, \tau\, b_{-i}, \overline{s}\right)\dot{\beta}_j = \frac{TE_i\left(b_i, \tau\, b_{-i}, \overline{s}\right)}{\tau} - \sum_{j\neq i}\frac{\partial}{\partial b_j} EU_i\left(b_i, \tau\, b_{-i}, \overline{s}\right) b_j, \tag{A.22}$$

where $\dot{\beta}_j$ stands for $\frac{d\dot{\beta}_j(\tau)}{d\tau}$. This equation is equivalent to equation (3.8) with the added initial condition $\beta(0) = 0$.

Equation (A.22) determines the derivative of function $\beta(\tau)$ with respect to τ. It will define a continuous function

$\beta(\tau)$ if the left-hand-side expression is continuous and non-singular. We will show that we can make a change of variables under which continuity and non-singularity of the left-hand side is clear.

In fact, note that the matrix of coefficients for $\dot{\beta}(\tau)$ can potentially become singular in the vicinity of $\tau = 0$. We assure that it is not the case by proving that the solution of (A.22) can be represented as a product of the vector function $x(\tau)$ that solves a non-singular system of differential equations and a matrix $M(\tau)$ that is not degenerate. Define function $x(\tau)$ and matrix $M(\tau)$ such that $\beta(\tau) = M(\tau)x(\tau)$, where matrix $M(\tau)$ is non-degenerate for each $\tau \in [0, 1]$, and as a function of τ $M(\cdot)$ satisfies

$$\frac{\partial}{\partial b'} EU\left(M\, x, \overline{s}\right)\, \dot{M} = -\frac{1-\tau}{\tau} \text{diag}\left\{\frac{\partial}{\partial b_i} EU_i\left(M\, x, \overline{s}\right)\right\} M(\tau). \tag{A.23}$$

Here $\dot{M} = \frac{d}{d\tau} M(\tau)$ is the matrix of derivatives of elements of $M(\tau)$ with respect to τ. If such a matrix indeed exists, then we can use the transformation $\beta(\tau) = M(\tau)x(\tau)$, and re-write the equation for $\dot{\beta}$ as a vector condition for $\dot{x}$:

$$\begin{aligned}\frac{\partial}{\partial b'} EU\left(M(\tau)x(\tau), \overline{s}\right) M(\tau)\dot{x} &= \frac{TE\left(M(\tau)x(\tau), \overline{s}\right)}{\tau} \\ &+ \left[\text{diag}\left(\frac{\partial}{\partial b'} EU\left(M(\tau)x(\tau), \overline{s}\right)\right) - \frac{\partial}{\partial b'} EU\left(M(\tau)x(\tau), \overline{s}\right)\right] M(\tau)x(\tau).\end{aligned}$$

Note that this transforms the original problem to the system of differential equations for $x(\tau)$ that is free from singularities in the vicinity of $\tau = 0$ by the assumption of the theorem. Moreover, the right-hand side of this system is Lipschitz-continuous. Therefore, by the standard existence theorem for the systems of nonlinear differential equations in Pontryagin (1966), the function $x(\tau)$ solving the above equation exists and is unique.

To finish the proof we use the following lemma to assure the existence of a non-singular matrix $M(\cdot)$.

LEMMA 1. *Suppose that matrix $M(\tau)$ has elements depending on τ and matrices $Z(M, \tau)$ and $Y(M, \tau)$ are known. Moreover, $Z(M, \tau)$ is non-singular for all M and $\tau \in [0, 1]$ and both $Y(M, \tau)$ and $Z(M, \tau)$ are Lipschitz-continuous in M and τ. Then the system of equations*

$$Z\left(M, \tau\right)\, \dot{M} = Y\left(M, \tau\right)\, M$$

with the boundary condition $M(1) = I_{n \times n}$ (identity matrix) has a unique non-singular solution.

The proof of this lemma can be found in Boltyanskii et al. (1960) and Pontryagin (1966).

In equation (A.23) $\frac{\partial}{\partial b'} EU\left(M\, x, \overline{s}\right)$ is Lipschitz. Therefore, both the right and the left-hand sides are Lipschitz and non-singular. As a result of Lemma 1 we conclude that considered transformation $\beta(\tau) = M(\tau)x(\tau)$ is unique.

This is system of ordinary differential equations without singularities (the vector of payoff functions has a non-singular Jacobi matrix and the considered change of variables is defined by a non-singular matrix $M(\tau)$). Now once we have this representation we proceed in the following steps. First, note that in the considered equation we define the vector of bids as a function of parameters τ. This means that we can represent the given system of differential equations as a system of differential equations for the vector of bids in the form:

$$A\,\dot{x} = c,$$

where matrix A corresponds to the matrix $-D_0(M(\tau)x(\tau), \bar{s})$. Both A and c are functions of x and τ. The set of bids satisfying the first-order condition will correspond to the solution of this equation $x(\tau)$ when $\tau = 1$.

Second, given that the set of equilibrium bids is associated with the solution of the given system, we can analyze the equilibrium by analyzing this solution. Given that matrix $A = D_0(M(\tau)x(\tau), \bar{s})$ is non-singular and the right-hand side c is continuous by the assumption of the Theorem, this system has a unique solution $x(\tau)$.

Third, if c is smooth and bounded, and the matrix of derivatives of the payoffs is strictly monotone, then the representation $\beta(\tau) = M(\tau)x(\tau)$ will hold for all points in the support of the vector of bids. As a result, we can apply Lemma 1 which establishes the sufficient condition for the uniqueness of the solution and proves the results (ii) and (iv) in Theorem 1.

B The Impact of Vanishing Uncertainty on Bidding and Identification

To gain some further intuition for how a model with uncertainty differs from the NU model, consider some limiting cases that are close to the NU model, where a small amount of uncertainty is added that serves as a refinement to the set NU equilibria. (In the empirical application, uncertainty is not small, so this exercise is intended to build intuition only.) First, consider what we call the random entry refinement. Suppose that there is no score uncertainty, but that with probability ϕ, a new advertiser enters with a random bid, and the distribution of the advertiser's score-weighted bid has full support over the relevant region. This is a realistic model of a new entrant or a new advertiser: the initial scores assigned by the system will not stay constant, and an advertiser may appear with a number of different score-weighted bids, each with low probability.

Now consider taking the limit as ϕ approaches zero. Then, taking into account that the entry of the random bidder affects marginal incentives only when it ties with the bidders score-weighted bid, it will be optimal for each advertiser to submit a bid that is an ex post equilibrium in the NU model, and in addition, where the bidder is indifferent between her current position when paying exactly her bid, or taking the next-lower position and paying the bid of the next-lowest bidder. Formally, the equilibrium conditions are the original equilibrium conditions (3.1), plus

$$s_{k_j} v_{k_j} \geq \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_{j+1} - s_{k_{j+2}} b_{k_{j+2}} \alpha_{j+2}}{\alpha_{j+1} - \alpha_{j+2}} = s_{k_{j+1}} v_{k_{j+1}},$$

except for the lowest-ranked bidder who bids her valuation. This contrasts with the Edelman et al. (2007) refinement, that satisfies

$$s_{k_j} v_{k_j} \geq \frac{s_{k_{j+1}} b_{k_{j+1}} \alpha_{j} - s_{k_{j+2}} b_{k_{j+2}} \alpha_{j+1}}{\alpha_{j} - \alpha_{j+1}} = s_{k_{j+1}} v_{k_{j+1}}.$$

The random entry strategies are envy-free if and only if $\alpha_j/\alpha_{j-1} \leq \alpha_{j+1}/\alpha_j$ for all $1 < j < J$ and the equilibrium is monotone. However, in general the random entry equilibrium may not exist in pure strategies. Intuitively, the auction has a "first-price" flavor: with some probability, each bidder pays her bid. Then, two bidders with similar score-weighted valuations will also place similar score-weighted bids; but when an opponent's bid is too close, a bidder's best response may be to drop down a position and take a lower price. This in turn might induce the opponent to change her bid, leading to cycling.

It is somewhat more subtle to consider the effects of small amounts of score uncertainty. We provide some intuition for a special case. Assume that $v_1 s_1 > v_2 s_2 > v_3 s_3$, and suppose there are two slots. Assume that $\tilde{s}_2$ is the stochastic score for bidder 2, and that the scores of the other bidders are fixed at their means. Let $f_{1/\tilde{s}_2}$ be the PDF of $1/\tilde{s}_2$. The local indifference condition defining the optimal bid b_2 (given the bids b_1, b_3) is

$$\alpha_2(v_2 - b_2) f_{1/\tilde{s}_2}\left(\frac{b_3 s_3}{b_2}\right) + \left[\alpha_1(v_2 - b_2) - \alpha_2\left(v_2 - \frac{b_3 s_3 b_2}{b_1 s_1}\right)\right] f_{1/\tilde{s}_2}\left(\frac{b_1 s_1}{b_2}\right) = 0 \tag{B.24}$$

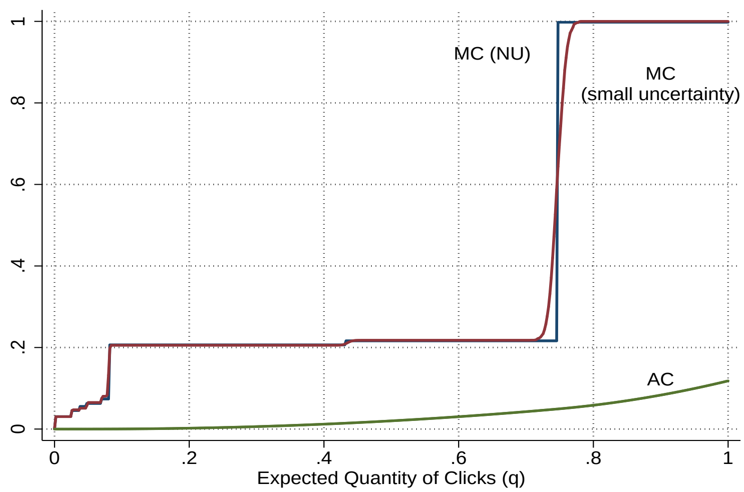
Suppose for a moment that $f_{1/\tilde{s}_2}\left(\frac{b_3 s_3}{b_2}\right) = 0$, so bidder 2 is not at risk for dropping a position. If $\gamma_2^* = \frac{b_1 \gamma_1}{b_2}$ is the critical value of the quality score that makes bidder 2 tie for the top position, the indifference condition reduces to

$$\alpha_1(v_2 - b_2) = \alpha_2\left(v_2 - \frac{b_3 \gamma_3}{\gamma_2^*}\right),$$

which is the EOS condition in the contingency where bidder 2 is tied with bidder 1. In contrast, if $f_{1/\tilde{s}_2}\left(\frac{b_1 s_1}{b_2}\right) = 0$ (no chance of moving up a position), the bidder is always better off by increasing her bid until $b_2 = v_2$, for standard reasons: the bid only matters if it causes the bidder to go from losing to winning. So, a small amount of quality score uncertainty puts upward pressure on bids if a bid is far from moving up to the next position, and we should generally expect to see the lowest position bidder place bids in a region where the bidder has some chance of moving up.[6]

We can also consider a refinement where the bidders face uncertainty, but the probability of a change in score or configuration is very small. Figure 6 below shows an effect of the small noise on the marginal and total cost. We use the actual bid and score data from a top configuration in a particular market. In this picture we assume that the score has a distribution with a mass point in the mean score. The sample for computation is generated by picking the score equal to the mean with probability $1-\varepsilon$ and equal to a random draw from the empirical distribution of scores with probability ε.

Figure 6: Marginal cost and total cost curves for bidder in a frequent configuration



C Proof of Theorem 4

To analyze the properties of the estimate for valuation we use the fact that the empirical profit function converges in probability to the population expected payoff function uniformly in valuation and the bid. Moreover, by our assumption regarding the distribution of the score, the score has a continuous density with a finite support. This implies that the numerical derivative will converge to the true derivative for the population analog of the considered functions. In particular, using Taylor's expansion and assuming that considered functions are twice differentiable with a Lipschitz-continuous residual of the second-order Taylor's expansion we can write:

$$\frac{-TE_i(b_i-2\tau_N,b_{-i},\bar{s})+8TE_i(b_i-\tau_N,b_{-i},\bar{s})-8TE_i(b_i+\tau_N,b_{-i},\bar{s})+TE_i(b_i+2\tau_N,b_{-i},\bar{s})}{-Q_i(b_i-2\tau_T,b_{-i},\bar{s})+8Q_i(b_i-\tau_N,b_{-i},\bar{s})-8Q_i(b_i+\tau_N,b_{-i},\bar{s})+Q_i(b_i+2\tau_N,b_{-i},\bar{s})}$$

$$=\frac{TE_i'(b_i,b_{-i},\bar{s})+L_1\tau_T^3}{Q_i'(b_i,b_{-i},\bar{s})+L_2\tau_N^3}=\frac{TE_i'(b_i,b_{-i},\bar{s})}{Q_i'(b_i,b_{-i},\bar{s})}+L_1\tau_N^3+L_2\tau_N^3+o\left(\tau_N^3\right),$$

[6] A similar result has been idependently obtained in Hashimoto (2009).

where L_1 and L_2 are Lipschitz constants. Next we consider the difference:

$$\widehat{v}_i - v_i = \frac{-\widehat{TE}_i(b_i-2\tau_N,b_{-i},\overline{s})+8\widehat{TE}_i(b_i-\tau_N,b_{-i},\overline{s})-8\widehat{TE}_i(b_i+\tau_N,b_{-i},\overline{s})+\widehat{TE}_i(b_i+2\tau_N,b_{-i},\overline{s})}{-\widehat{Q}_i(b_i-2\tau_N,b_{-i},\overline{s})+8\widehat{Q}_i(b_i-\tau_N,b_{-i},\overline{s})-8\widehat{Q}_i(b_i+\tau_N,b_{-i},\overline{s})+\widehat{Q}_i(b_i+2\tau_N,b_{-i},\overline{s})} - \frac{TE_i'(b_i,b_{-i},\overline{s})}{Q_i'(b_i,b_{-i},\overline{s})}$$

$$= D_1 + D_2 + D_3 + o_p\left(\frac{1}{\sqrt{T\,\tau_T}}\right).$$

Here we use the following decomposition:

$$D_1 = \frac{18}{Q_i'\left(b_i, b_{-i}, \overline{s}\right)} \left[\widehat{TE}_i\left(b_i, b_{-i}, \overline{s}\right) - TE_i\left(b_i, b_{-i}, \overline{s}\right)\right],$$

$$D_2 = -\frac{18\, TE_i'\left(b_i, b_{-i}, \overline{s}\right)}{\left(Q_i'\left(b_i, b_{-i}, \overline{s}\right)\right)^2} \left[\widehat{Q}_i\left(b_i, b_{-i}, \overline{s}\right) - Q_i\left(b_i, b_{-i}, \overline{s}\right)\right],$$

and

$$D_3 = \frac{-TE_i(b_i-2\tau_N,b_{-i},\overline{s})+8TE_i(b_i-\tau_N,b_{-i},\overline{s})-8TE_i(b_i+\tau_N,b_{-i},\overline{s})+TE_i(b_i-2\tau_N,b_{-i},\overline{s})}{-Q_i(b_i-2\tau_T,b_{-i},\overline{s})+8Q_i(b_i-\tau_N,b_{-i},\overline{s})-8Q_i(b_i+\tau_N,b_{-i},\overline{s})+Q_i(b_i-2\tau_N,b_{-i},\overline{s})} - \frac{TE_i'(b_i,b_{-i},\overline{s})}{Q_i'(b_i,b_{-i},\overline{s})}.$$

We omitted all the terms of the smaller order than $o_p((T\tau_T)^{-1/2})$ using the assumption regarding the rate of the numerical differentiation. We can use the result in Pollard (1990) to argue that

$$\sup_{|\delta|<\epsilon} \frac{1}{\sqrt{T^*}} \sum_{t^*} \left[u_i\left(v_i, b_i+\delta;\, b_{-i},\, \overline{s}_i, \widehat{\varepsilon}_{it^*}, C_{t^*}\right) - u_i\left(v_i, b_i+\tau_T;\, b_{-i},\, \overline{s}_i, \widehat{\varepsilon}_{it}, C_t\right)\right] = o_p(\sqrt{\delta}).$$

Finally, using the structure of total expenditure and expected quantity of clicks, we can write:

$$\sqrt{T\tau_T}\left(\widehat{v}_i - v_i\right) = -18 \frac{1}{Q_i'(b_i,b_{-i},\overline{s})} \frac{1}{\sqrt{T}} \sum_t \frac{u_i(v_i,b_i+\tau_T;b_{-i},\overline{s}_i,\widehat{\varepsilon}_{it},C_t)-u_i(v_i,b_i-\tau_T;b_{-i},\overline{s}_i,\widehat{\varepsilon}_{it},C_t)}{\sqrt{\tau_T}},$$

Then if $\Omega = \mathrm{Var}\left(\frac{u_i(v_i,b_i+\tau_T;b_{-i},\overline{s}_i,\widehat{\varepsilon}_{it},C_t)-u_i(v_i,b_i-\tau_T;b_{-i},\overline{s}_i,\widehat{\varepsilon}_{it},C_t)}{\sqrt{\tau_T}}\right)$, it follows that the and i.i.d. Assumption 2, bootstrap is valid by Kosorok (2008) and

$$\sqrt{T\tau_T}\left(\widehat{v}_i - v_i\right) \xrightarrow{d} N\left(0,\, \frac{324\,\Omega}{\left(Q_i'\left(b_i, b_{-i}, \overline{s}\right)\right)^2}\right)$$

D Estimation of valuations in case of set-valued best response correspondences

Even though we can consistently estimate the payoff of the bidder for each valuation and the score, there is no guarantee that for each bid there will be a single valuation which makes this bid consistent with the first-order condition. General results for set inference in the auction settings have been developed for instance in Haile and Tamer (2003), while general results for identification in the auction settings are given in Athey and Haile (2002). This result will display most likely in the situation where score-weighted bids have limited overlap, i.e. for a fixed set of bids we can find positions such that some bidders will never have their ads displayed in these positions. In this case local bid changes may not affect the payoff as they will not affect the relative ranking of the bidders. If $b_k\,\overline{s}_k\,\overline{\varepsilon} < b_i\,\overline{s}_i\underline{\varepsilon}$, then the score-weighted bid of bidder k will always be below the bid of bidder i. Similarly, if $b_k\,\overline{s}_k\underline{\varepsilon} > b_i\overline{s}_i\overline{\varepsilon}$ then the bid of bidder k will always be ranked higher than the bid of bidder i. In the extreme case where for each pair of bidders j and k we have

$$(b_k\,\overline{s}_k\underline{\varepsilon} - b_j\,\overline{s}_j\overline{\varepsilon})\,(b_j\,\overline{s}_j\underline{\varepsilon} - b_k\,\overline{s}_k\overline{\varepsilon}) > 0$$

(i.e. the ranked bids never overlap), then the model substantially simplifies. Assume that the bids are ordered by their ranks using the mean scores: $b_j\overline{s}_j > b_{j-1}\overline{s}_{j-1}$. Also assume that $\pi = 0$ so that all bidders are always present in the auction. A selected bidder will be placed in position k and pay $b_k\overline{s}_k E\left[s^{-1}\right]$ if $b_k\frac{\overline{s}_k\overline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}} < b < b_{k-1}\frac{\overline{s}_{k-1}\underline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}}$. If the bid is $b_k\frac{\overline{s}_k\underline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}} < b < b_k\frac{\overline{s}_k\overline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}}$ or $b_k\frac{\overline{s}_k\underline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}} < b < b_k\frac{\overline{s}_k\overline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}}$, then the probability of being placed in position k is

$$\int F_\varepsilon\left(\frac{bs}{b_k\overline{s}_k}\right) f_\varepsilon\left(\frac{s}{\overline{s}_i}\right)\,ds,$$

and the expected payment is

$$\int\int \mathbf{1}\left\{b_k\,s' < b\,s\right\}\frac{b_k\,s'}{s} f_\varepsilon\left(\frac{s'}{\overline{s}_k}\right) f_\varepsilon\left(\frac{s}{\overline{s}_i}\right)\,ds\,ds'.$$

Similarly if $b_{k-1}\frac{\overline{s}_{k-1}\underline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}} < b < b_{k-1}\frac{\overline{s}_{k-1}\overline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}}$ or $b_{k-1}\frac{\overline{s}_{k-1}\underline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}} < b < b_{k-1}\frac{\overline{s}_{k-1}\overline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}}$, then the probability of being placed in position k is

$$\int\left(1 - F_\varepsilon\left(\frac{bs}{b_{k-1}\overline{s}_{k-1}}\right)\right) f_\varepsilon\left(\frac{s}{\overline{s}_i}\right)\,ds,$$

and the expected payment is

$$\int\int \mathbf{1}\left\{b_{k-1}\,s' > b\,s\right\}\frac{b_k\,\overline{s}_k}{s} f_\varepsilon\left(\frac{s'}{\overline{s}_{k-1}}\right) f_\varepsilon\left(\frac{s}{\overline{s}_i}\right)\,ds\,ds'.$$

Then the objective function of the bidder i will be not strictly monotone. It will have "flat spots" where there is no bid overlap and it will be smooth where score-weighted bids overlap. We can explicitly compute the marginal utility from bidding b as

$$\frac{\partial}{\partial b}\mathbb{E}_{\varepsilon,C}\left[u_i\left(v_i, b_i = b, b_{-i};\, \varepsilon_{it}, C_t^i\right)\right] = \begin{cases} 0, \text{ if } \; b_k\frac{\overline{s}_k\overline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}} < b < b_{k-1}\frac{\overline{s}_{k-1}\underline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}}, \\ \overline{\alpha}_k \int\left(\frac{v_i s}{b_k} - b\right) f_\varepsilon\left(\frac{s}{\overline{s}_i}\right) f_\varepsilon\left(\frac{bs}{b_k\overline{s}_k}\right) \\ -\overline{\alpha}_{k+1}\int\left(\frac{v_i s}{b_k} - \frac{b_{k+1}\overline{s}_{k+1}}{s}\right) f_\varepsilon\left(\frac{s}{\overline{s}_i}\right) f_\varepsilon\left(\frac{bs}{b_k\overline{s}_k}\right)\,ds, \\ \qquad\qquad \text{if } \; b_k\frac{\overline{s}_k\underline{\varepsilon}}{\overline{s}_i} < b < b_k\frac{\overline{s}_k\overline{\varepsilon}}{\overline{s}_i}. \end{cases}$$

In the limited overlap case the numerical algorithm for computation of the best responses will contain 3 steps.

- Step 1 Compute $\frac{\partial}{\partial b}E\left[u_i\left(v_i, b_i=b, b_{-i}, \varepsilon_{it}, C_t^i\right)\right]$ at each of $4\left(N-1\right)$ points $b_k\frac{\overline{s}_k(\times/\div)\varepsilon}{\overline{s}_i(\times/\div)\varepsilon}$
- Step 2 If for some k there are 2 points out of 4 $b_k\frac{\overline{s}_k(\times/\div)\varepsilon}{\overline{s}_i(\times/\div)\varepsilon}$ where the marginal utility has different signs, solve the non-linear equation

$$\overline{\alpha}_k\int\left(\frac{v_i s}{b_k}-b\right)f_\varepsilon\left(s-\overline{s}_i\right)f_\varepsilon\left(\frac{bs}{b_k\overline{s}_k}\right)-\overline{\alpha}_{k+1}\int\left(\frac{v_i s}{b_k}-\frac{b_{k+1}\overline{s}_{k+1}}{s}\right)f_\varepsilon\left(\frac{s}{\overline{s}_i}\right)f_\varepsilon\left(\frac{bs}{b_k\overline{s}_k}\right)ds=0.$$

  Obtain solution b^*.
- Step 3 Compare $\overline{\alpha}_k\left(v_i-\overline{s}_k b_k E\left[s_{it}^{-1}\right]\right)$ for all k and $E\left[u_i\left(v_i, b_i=b^*, b_{-i}, \varepsilon_{it}, C_t^i\right)\right]$ where the latter were computed. If the maximum value is $\overline{\alpha}_k\left(v_i-\overline{s}_k b_k E\left[s_{it}^{-1}\right]\right)$, then the best response is set valued with $b\in\left[b_k\frac{\overline{s}_k\overline{\varepsilon}}{\overline{s}_i\underline{\varepsilon}},\ b_{k-1}\frac{\overline{s}_{k-1}\underline{\varepsilon}}{\overline{s}_i\overline{\varepsilon}}\right]$. Otherwise, the best response is unique and equal to b^*.

To recover valuations in case of limited overlap of the score-ranked bids, we fix the set of observed bids. We also fix the grid which contains the support of valuations. Then for each bidder and each value on the grid we solve for the set of best responses. Given the produced set of best responses we pick the set of valuations for which the set of best responses contains the actually observed best response. Technically this implies that we recover the set:

$$S_i=\left\{(b,v)\,\middle|\,b\in\text{BR}_i\left(v,\,b_{-i}\right),v\in\mathcal{V}\right\}.$$

The estimated valuation is the cut of this set such that

$$\left(\widehat{v}_i,\overline{b}_i\right)\in S_i,$$

where $\overline{b}_i$ is observed in the data.

The structure of our empirical procedure allows us to formulate the following result.

THEOREM 5. *Under Assumption 2 the estimation procedure following the outlined steps 1-3 is numerically equivalent to the statistics inversion procedure in Chernozhukov et al. (2007). As a result, the estimates of identified set of valuations will be described by Theorem 2.1 in Chernozhukov et al. (2007).*

To provide the argument, we consider the following scheme.

1. Consider the sample of all observed bidder configurations over queries t $\{C_t\}_{t=1}^T$ where T is the total number of queries. Uniformly over these sets draw a set C_{t^*}. Select a particular bidder i Construct a set $C_{t^*}^i=C_{t^*}\setminus\{i\}$. In total we construct T^* subsamples of collections of sets of configurations.

2. For a fixed position j make K^* random subsamples $\{C^i_{t^*,k,j-1}\}_{k=1}^{K(T)}$ of $j-1$ bidders out of set C_{t^*}. The number of subsamples K^* needs to grow such that $K^*/\sqrt{T} \to \infty$. For configuration C_{t^*} compute the payoff of bidder i from being placed in position j

$$u^{i,j}_{t^*,k}(b_i, v_i) = \overline{\alpha}_j \sum_{k=1}^{K^*} \int\int \left(v_i\, F_s\left(\frac{s'}{\overline{s}_k}\right) \right.$$
$$\times \prod_{m \in C^i_{t^*,k,j-1}} \left(\frac{1 - F_s\left(\frac{sb}{\overline{s}_m b_m}\right)}{F_s\left(\frac{sb}{\overline{s}_m b_m}\right)} \right) \prod_{n \in C^i_{t^*}} F_s\left(\frac{sb}{\overline{s}_n b_n}\right)$$
$$- \sum_{k \in C^i_{t^*} \setminus C^i_{t^*,k,j-1}} \frac{b_k s'}{s} \mathbf{1}\{b_k\, s' < b\, s\} \frac{F_s\left(\frac{s}{\overline{s}_i}\right)}{F_s\left(\frac{s' b_k}{\overline{s}_i b}\right)}$$
$$\left. \times \prod_{m \in C^i_{t^*,k,j-1}} \left(\frac{1 - F_s\left(\frac{sb}{\overline{s}_m b_m}\right)}{F_s\left(\frac{s' b_k}{\overline{s}_m b_m}\right)} \right) \prod_{n \in C^i_{t^*}} F_s\left(\frac{s' b_k}{\overline{s}_n b_n}\right) \right) d\, \log\, F_s\left(\frac{s}{\overline{s}_i}\right)\, d\, \log\, F_s\left(\frac{s'}{\overline{s}_k}\right).$$

If we use T^* draws of configurations of bidders in the first stage, and K^* draws in the second stage, we need to compute the approximated payoff by rescaling as

$$\widehat{EU}_i\left(b_i = b;\, b_{-i},\, \overline{s}\right) = \sum_{j=1}^{J} \frac{1}{T^*} \sum_{t=1}^{T^*} \frac{\binom{\#C^i_{t^*}}{j}}{K^*} \sum_{k=1}^{K^*} u^{i,j}_{t^*,k}(b_i, v_i).$$

This procedure allows us to evaluate the payoff function of a single bidder using $T^* \times K^*$ total draws. Note that we can "recycle" the draws of sets of configurations to compute the payoff functions for different bidders. We then can compute the numerical derivative

$$\frac{\partial}{\partial b}\widehat{EU}_i\left(b_i = b;\, b_{-i},\, \overline{s}\right) = \sum_{j=1}^{J} \frac{1}{T^*} \sum_{t=1}^{T^*} \frac{\binom{\#C^i_{t^*}}{j}}{K^*} \sum_{k=1}^{K^*} \frac{u^{i,j}_{t^*,k}\left(b+\tau, v_i\right) - u^{i,j}_{t^*,k}\left(b-\tau, v_i\right)}{2\tau}.$$

Given the assumption that bidders set their bids optimally, we ca write the condition

$$\frac{\partial}{\partial b}\widehat{EU}_i\left(b_i = \overline{b}_i,\, b_{-i}\right) = \sum_{j=1}^{J} \frac{1}{T^*} \sum_{t^*} \frac{\binom{\#C^i_{t^*}}{j}}{K^*} \sum_{k=1}^{K^*} \frac{u^{i,j}_{t^*,k}\left(\overline{b}_i+\tau, v_i\right) - u^{i,j}_{t^*,k}\left(\overline{b}_i-\tau, v_i\right)}{2\tau} = o_p\left(1\right),$$

at the observed bid. Then we can recover the set of values that correspond to the observable bid. To do so we form the grid over v and minimize

$$\left(\sum_{j=1}^{J} \frac{1}{T^*} \sum_{t^*} \frac{\binom{\#C^i_{t^*}}{j}}{K^*} \sum_{k=1}^{K^*} \frac{u^{i,j}_{t^*,k}\left(\overline{b}_i+\tau, v_i\right) - u^{i,j}_{t^*,k}\left(\overline{b}_i-\tau, v_i\right)}{2\tau} \right)^2,$$

with respect to v. The set of minimizers will deliver the identified set of valuations $\widehat{\mathcal{F}}_{v,T,J}$. This procedure allows estimation similar to that offered in Chernozhukov et al. (2007). The confidence sets can be recovered using the tools developed in Imbens and Manski (2004).

E Algorithm and description of Monte-Carlo Simulations

In the Monte-Carlo simulations we analyze the stability of our estimation procedure with respect to the sampling noise in the data as well as the width of the support of valuations. The first set of Monte-Carlo simulations was designed to analyze the robustness of the suggested computational procedure to the sampling noise in the observed configurations of advertisers. The setup of the Monte-Carlo simulation was the following. We considered the case where there are 5 advertisers competing for 2 slots. The click-through rates of these slots were fixed at levels 1 and 0.5. The valuations have support on $[0, 1]$ and the scores for all advertisers are uniformly distributed on $[0, .1]$. We consider the cases where the reserve price was equal to 0.1, 0.2 and 0.3. We use the same probability of a binding budget constraint for all bidders. This probability was selected at the levels 0, 0.01, and 0.05. We used 2000 Monte-Carlo replications. Each iteration was organized in the following way. First, we sample valuations for each bidder from $U\,[0, 1]$. Second, for the set of valuations we computed the equilibrium of the model. In case of the uniform distribution of the scores, the problem of computing the equilibrium is equivalent to solving a system of polynomial equations (of order 4 for 5 players) with linear constraints. Then for each bidder we generated uniform random variables and removed the bidders for whom the uniform draw was below the probability of a binding budget constraint. Then we fixed the bids and generated each set of Monte-Carlo draws using the algorithm

- Using uniform draws, remove bidders with binding budget constraint
- Record equilibrium bids for remaining bidders
- Generate scores for the bidders from the uniform distribution
- Allocate bidders to slots and compute the prices

We used three setups where each Monte-Carlo sample had 500, 1000 and 2000 individual draws. For each sample we computed the payoff function, and computed the valuations of the participating bidders by inverting the first-order condition. In the table below we report our results. We report standard deviations of the difference between exact and estimated profits for players from 1 to 5 and the standard deviations for recovered valuations for players from 1 to 5. The following table reports the estimates for the case where the probability of players dropping out due to budget constraints is zero.

This table shows a significant decline in the standard errors of estimation when the Monte-Carlo sample size increases. This supports the formal argument of consistency of our estimation procedure.

Table 6: Results of Monte-Carlo Analysis (no binding budget constraints)

	Profits					Valuations				
Player#	1	2	3	4	5	1	2	3	4	5
Sample size =500										
	.654	.622	.788	.501	.714	.220	.124	.150	.221	.250
Sample size =1000										
	.311	.355	.330	.341	.318	.110	.098	.101	.118	.106
Sample size =2000										
	.122	.110	.114	.164	.142	.055	.068	.060	.071	.062

Table 7: Results of Monte-Carlo Analysis (probability of reaching the budget constraint 1%)

	Profits					Valuations				
Player#	1	2	3	4	5	1	2	3	4	5
Sample size =500										
	1.034	.1.507	1.142	.980	1.450	.320	.215	.345	.318	.343
Sample size =1000										
	.890	1.079	1.120	.760	1.235	.250	.201	.305	.285	.299
Sample size =2000										
	.530	.511	.595	.544	.645	.176	.129	.201	.148	.187

Table 8: Results of Monte-Carlo Analysis (probability of reaching the budget constraint 5%)

	Profits					Valuations				
Player#	1	2	3	4	5	1	2	3	4	5
Sample size =500										
	2.003	3.790	3.202	2.254	2.990	.269	.235	.130	.021	.189
Sample size =1000										
	1.840	1.089	2.044	2.011	2.940	.336	.218	.238	.299	.201
Sample size =2000										
	1.188	1.112	2.230	1.970	1.450	.096	.128	.130	.199	.160

F Recovering distributions of scores and clickthrough rates from the data

Now we will provide a more formal argument for identification of the CTR. First, we consider identification of the distribution of noise in the click-through rates, and subsequently, the distribution of estimated click-through rates. The distribution of the estimated advertiser-specific rate is denoted $F_{\gamma,i}(\cdot\,|\,z)$ and the distribution of the estimated slot-specific click-through rate is denoted $F_{\alpha,j}(\cdot\,|\,z)$. The distribution of bidder valuations is also a common knowledge among bidders. The following proposition establishes the fact that we can recover distributions of the bidder-specific and the slot-specific CTR from observable frequencies of clicks $G_{ij}(\cdot)$ for bidder i in slot j.

THEOREM 6. *Assume that the distribution of the estimated slot-specific CTR is degenerate at α in slot 1 (where α is a known constant), and the distribution of the noise in the advertiser-specific CTR $F_\gamma(\cdot)$ is the same across advertisers. Moreover, assume that the noise in the estimated slot-specific CTR ε_j^α is independent from the noise in the estimated advertiser-specific CTR ε_i^γ for all advertisers and all slots. Then both the distribution of advertiser-specific CTR and the distribution of slot-specific CTR $F_{\alpha,j}(\cdot)$ for all slots j are identified.*

Proof:

Given that $G_{c,i,j}(x) = E\left[\mathbf{1}\left\{C_{ij} < x\right\}\right]$, then for slot 1

$$G_{c,i,1}(x) = E\left[\mathbf{1}\left\{\alpha\Gamma_i < x\right\}\right] = F_\gamma\left(\frac{x}{\alpha}\right),$$

meaning that the distribution of Γ_i is identified. Denote the distribution of $\log C_{ij}$ by $G^l_{c,i,j}(\cdot)$ and the distribution of $\log A_j$ and $\log \Gamma_i$ by $F^l_{\alpha,i}$ and F^l_γ correspondingly. Then the density of the logarithm of the CTR is expressed through the density of slot-specific CTR and advertiser-specific CTR by the convolution formula

$$g^l_{c,i,j}(x) = \int_{\log \underline{\gamma}}^{\log \overline{\gamma}} f^l_{\gamma}(\gamma)\, f^l_{\alpha,j}(x-\gamma)\, d\gamma.$$

Then the characteristic function for the distribution of A_j can be expressed using deconvolution

$$\chi^l_{\alpha,j}(t) = \frac{\chi^l_{c,i,j}(t)}{\chi^l_{\gamma}(t)}.$$

The characteristic function is computed as

$$\chi^l_{\gamma}(t) = \int_{-\infty}^{+\infty} e^{itx} f^l_{\gamma}(x)\, dx,$$

where $i = \sqrt{-1}$. Then we can recover the distribution of slot-specific CTR for slot j using the inverse Fourier transformation

$$F_{\alpha,j}(x) = \int_{-\infty}^{\log x} dz \int_{-\infty}^{+\infty} e^{-itz} \chi^l_{\alpha,j}(t)\, dt.$$

As a result, for each slot $j = 1, \ldots, J$ starting from the second one we can find the distribution of its slot-specific conversion rate.

Q.E.D.

G Computing equilibria via numerical continuation

For $\tau \in [0,1]$ the system (3.8) can be re-written as

$$\sum_{j \neq i} \frac{\partial EU_i\left(\beta_i(\tau),\, \tau\, \beta_{-i}(\tau), \overline{s}\right)}{\partial b_j}\, \tau\, b_j(\tau) = -TE_i\left(\beta_i(\tau),\, \tau\, \beta_{-i}(\tau), \overline{s}\right), \quad i = 1, \ldots, N. \tag{G.25}$$

If the payoff function is twice continuously differentiable and the equilibrium existence conditions are satisfied, then $\beta(\tau)$ is a smooth function of τ. As a result, we can further differentiate both sides of this expression with respect to τ. For the left-hand side we can obtain

$$\begin{aligned}&\sum_{j,k \neq i} \frac{\partial^2 EU_i(\beta_i(\tau), \tau\, \beta_{-i}(\tau), \overline{s})}{\partial b_j\, \partial b_k} \left[\tau^2\, b_j\, \dot{b}_k + \tau b_j\, b_k\right] + \frac{\partial^2 EU_i(\beta_i(\tau), \tau\, \beta_{-i}(\tau), \overline{s})}{\partial b_j\, \partial b_i}\, \tau\, b_j\, \dot{b}_i \\ &+ \sum_{j \neq i} \frac{\partial EU_i(\beta_i(\tau), \tau\, \beta_{-i}(\tau), \overline{s})}{\partial b_j} \left[\tau\, \dot{b}_j + b_j\right],\end{aligned} \tag{G.26}$$

where $\dot{b} = \frac{db}{d\tau}$. Then using the notation δ_{kj} for the Kronecker symbol, we can re-write the expression of interest as

$$\sum_k a^i_k\, \dot{b}_k = c^i, \tag{G.27}$$

and

$$a_k^i = \left[\tau^2 (1-\delta_{ik}) + \tau\delta_{ik}\right] \sum_{j\neq i} \frac{\partial^2 EU_i(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s})}{\partial b_j\, \partial b_k}\, b_j b_k + \tau (1-\delta_{ik})\, v_i \frac{\partial Q_i(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s})}{\partial b_k} b_k$$
$$+\delta_{ik} \frac{\partial TE_i(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s})}{\partial b_i} b_i$$

and

$$c^i = -\sum\nolimits_k \tau (1-\delta_{ik}) \sum_{j\neq i} \frac{\partial^2 EU_i(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s})}{\partial b_j\, \partial b_k} b_j b_k + (1-\delta_{ik})\, v_i \frac{\partial Q_i(\beta_i(\tau), \tau\,\beta_{-i}(\tau), \overline{s})}{\partial b_k} b_k$$

We make an inverse transformation and express the system of equations of interest in the form

$$A(\mathbf{b}, \tau)\ \dot{\mathbf{b}} = c(\mathbf{b}, \tau),$$

where the elements of matrix $A(\mathbf{b}, \tau)$ can be computed as $A_{ik}(\mathbf{b}, \tau) = a_k^i$. We know that the original system of non-linear equations has the solution $\beta(0) = 0$ corresponding to the point $\tau = 0$. We solve the problem by constructing a grid over $\tau \in [0,\ 1]$ and choosing the tolerance level Δ accordingly to the step of the grid. The set of grid point is $\{\tau_N\}_{t=1}^T$ where $\Delta = \max\limits_{t=2,\ldots,T} \|\tau_N - \tau_{t-1}\|$. The solution at each grid point τ_N will be a vector of bids b_t. Then we can use the modified Euler integration scheme to compute the solution on the extended interval. We can note that the system of differential equation has a singularity of order one at the origin. We use a simple regularization scheme which allows us to avoid the singularity at a cost of an additional approximation error of order Δ^α, where α is the power such that $\lim\limits_{\delta\to+0} \delta^{-\alpha} \frac{\partial^2 EU_i(b_i, b_{-i})}{\partial b_i \partial b_j}\Big|_{\|b\|=\delta} < \infty$ for all i. Note that this condition is satisfied if the Hessian matrix of the payoff function is non-degenerate at the origin. We initialize the system at $b_0 = \Delta/4$ and make a preliminary inverse Euler step by solving

$$\mathbf{b}_{1/2} = b_0 + A\left(\mathbf{b}_{1/2}, \Delta/2\right)^{-1} c\left(\mathbf{b}_{1/2}, \Delta/2\right) \Delta/2 \tag{G.28}$$

with respect to $\mathbf{b}_{1/2}$. Such an inverse step enhances the stability of the algorithm and it will be the most time-consuming part. Then the algorithm proceeds from step t to step $t+1$ in the steps of 1/2. Suppose that $\mathbf{b}_t$ is the solution at step t. Then we make a preliminary Euler step

$$\mathbf{b}_{t+1/2} = \mathbf{b}_t + \frac{\Delta}{2} A(\mathbf{b}_t, \tau_N)^{-1} c(\mathbf{b}_t, \tau_N). \tag{G.29}$$

Then using this preliminary solution we make the final step

$$\mathbf{b}_{t+1} = \mathbf{b}_t + \Delta A\left(\mathbf{b}_{t+1/2}, \tau_N + \frac{1}{2}\Delta\right)^{-1} c\left(\mathbf{b}_{t+1/2}, \tau_N + \frac{1}{2}\Delta\right).$$

Note that the values that are updated only influence the evaluated derivative, while the final step size is still equal to Δ. We can use standard numerical derivative approximation to compute the elements of $A(\mathbf{b}, \tau)$ and $c(b, \tau)$. For the first derivative we use the third-order formula such that

$$\frac{\partial EU_i(\mathbf{b}, \tau, \overline{s})}{\partial b_j} = \frac{EU_i(b_j - 2\delta, b_{-j}, \tau, \overline{s}) - 8EU_i(b_j - \delta, b_{-j}, \tau, \overline{s}) + 8EU_i(b_j + \delta, b_{-j}, \tau, \overline{s}) - EU_i(b_j + 2\delta, b_{-j}, \tau, \overline{s})}{12\delta} + o\left(\delta^5\right),$$

where δ is the step size in the domain of bids[7]. For the second cross-derivatives we can use the "diamond" formula

$$\frac{\partial^2 EU_i(\mathbf{b},\tau)}{\partial b_j \partial b_k} = \frac{1}{12\delta^2}\Big[EU_i\left(b_j-2\delta, b_{-j},\tau,\overline{s}\right) - EU_i\left(b_k-2\delta, b_{-k},\tau,\overline{s}\right)$$
$$-8EU_i\left(b_j-\delta, b_{-j},\tau,\overline{s}\right) + 8EU_i\left(b_k-\delta, b_{-k},\tau,\overline{s}\right)$$
$$+8EU_i\left(b_j+\delta, b_{-j},\tau,\overline{s}\right) - 8EU_i\left(b_k+\delta, b_{-k},\tau,\overline{s}\right)$$
$$-EU_i\left(b_j+2\delta, b_{-j},\tau,\overline{s}\right) + EU_i\left(b_k+2\delta, b_{-k},\tau,\overline{s}\right)\Big] + o\left(\delta^4\right),$$

Then the order of approximation error on the right-hand side is $o(\delta^4)$. For stability of the computational algorithm it is necessary that $\delta^4 = o(\Delta)$. This can be achieved even if one chooses $\delta = \Delta$ (up to scale of the grid). This condition becomes essential if in the sample the function EU_i is not smooth. In that case the minimal step size δ is determined by the granularity of the support of the payoff function. The step size for τ should be chosen appropriately and cannot be too small to avoid the accumulation of numerical error.

Initialization of the system simplifies when the auction has a reserve price. When the reserve price is equal to r, then both the expected utility and the total expenditure become functions of r. Homogeneity of the utility function will also be preserved when we consider the vector of bids accompanied by r. As a result, the system of of equilibrium equations will take the form

$$\frac{\partial}{\partial \mathbf{b}'} EU\left(\mathbf{b},\overline{s},r\right)\mathbf{b} + \frac{\partial}{\partial r} EU\left(\mathbf{b},\overline{s},r\right) r = -TE\left(\mathbf{b},\overline{s},r\right). \tag{G.30}$$

As a result, we can re-write our main result as

$$\frac{d}{d\tau} EU_i\left(b_i,\tau\mathbf{b}_{-i},\overline{s}\right)\big|_{\tau=1} = -TE_i\left(\mathbf{b},\overline{s}\right) - r\frac{\partial}{\partial r} EU_i\left(\mathbf{b},\overline{s},r\right). \tag{G.31}$$

Our results for τ in the neighborhood of $\tau = 1$ will apply with total expenditure function corrected by the influence of the reserve price. In the case where the vector of the payoff functions has a non-singular Jacobi matrix globally in the support of bids, we can also extend the results for $\tau \in [0,1]$ to the case with the reserve price. In this case, the initial condition for $\tau = 0$ will solve

$$-TE_i(b_i(0),0,\overline{s}) - r\frac{\partial}{\partial r} EU_i\left(b_i(0),0,\overline{s},r\right) = 0.$$

Note that for all bidders $i = 1,\ldots,N$ this is a non-linear equation with a scalar argument $b_i(0)$, which can be solved numerically. This will allow us to construct a starting value for the system of differential equations. Note that in this case equilibrium computations simplify because there is no need in the "inverse" Euler step which

[7] We need to emphasize that for the appropriate quality of approximation, when using the fourth-order formula for the numerical derivative, one needs to assure that $\Delta \gg \delta^5$. In other words, the step size for numerical integration should be larger than the step size for numerical derivative.

we used to stabilize the system of differential equations at the origin. The algorithm will start from the standard preliminary Euler step $\frac{1}{2}\Delta$.

H The sources of estimation bias and robustness check

In this Appendix, we discuss the modeling choices we made in light of the data limitations, and we present the empirical results that establish the robustness of our estimation approach to these modeling choices.

There are three main elements used to estimate the marginal cost for a particular advertiser: (i) the distribution of quality scores (mean values and the distribution of shocks); (ii) the set of user queries where the advertisement of the advertiser of interest was considered; (iii) the set of competing advertisements that was considered for a user query. The feature of our historical research dataset is that we do not observe bids and quality scores for advertisements that did not appear on the page. As a result, we do not know the full set of advertisements that was considered for a particular user query. There could be several reasons why an advertisement did not appear in a particular user query. First, the random draw of the quality score was too low and the score-weighted bid of the advertiser was either outbid by other bidders or did not exceed the reserve price. Second, the advertiser has set budget limit for the ad campaign and the budget has been exceeded. Third, the advertiser has set exclusion targeting and a particular user query does not satisfy targeting restrictions.

In our empirical analysis we assume that the observed sets of ads coinsides with the sets of ads considered for user queries. This creates several potential problems for our analysis, which can be discussed in the context of the three components of the marginal cost estimation. First, a selection problem may arise, in that we only observe quality scores that were high enough so that the product of the advertisers per-click bid and their quality score ranked in the top set of advertisements. This could potentially impact our estimates of the mean quality scores as well as the shape of the quality score distribution. Second, we may over-estimate the uncertainty in rival configuration by exposing the ad to the queries for which it was not eligible due to exclusion targeting.

Now consider how we handle these problems. Our approach is loosely motivated by a model (although this is not a completely accurate description of the setting) where advertisers submit multiple advertisements and have budget constraints that determine the fraction of user queries the advertisements might appear on, and the system randomly selects which advertisement is chosen as well as which user queries to assign the advertisement to. We first discuss the choices and provide a comparison between the outcomes of the following empirical analyses. (i) We use our baseline methodology and estimate the distribution of quality scores, ignoring the selection problem and treat the data as if it came from the population of quality scores rather than a selected sample. In addition, we focus only on the first page of advertisements viewed by the user, which account for a very large share of the clicks and revenue for each advertisement. (ii) We assume that each advertisement's bid was considered for all user queries in the sample (that is, even though in practice the advertisement did not appear on many user queries, we assume that a priori the advertisement could have appeared on any of them and the advertiser did not anticipate in advance which subset would be selected). (iii) We assume that the empirical distribution of competing advertisements is the distribution that advertiser anticipates.

We discussed the first approach in our main empirical section. We will now compare the results obtained using our baseline approach with the results obtained under the second and the third sets of assumptions.

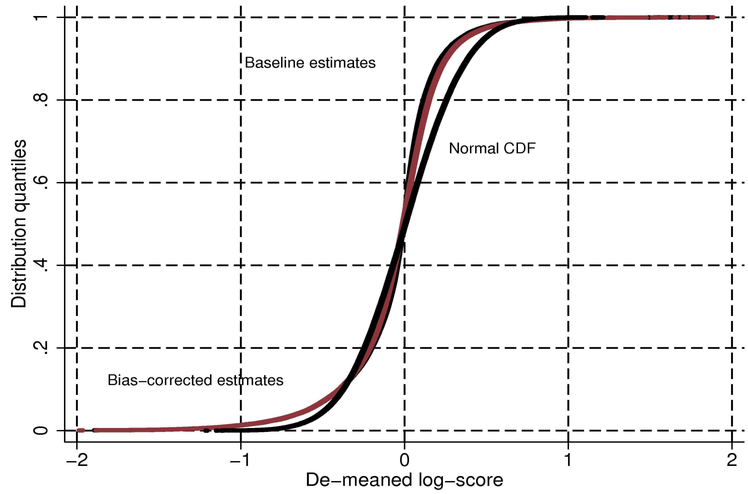
Estimation of the score distribution: To study the effect of the sample selection on the estimate of the distribution of the shocks to the scores, we adapt our estimation methodology to the second assumption, that each advertisement's bid was considered for all user queries in the sample. Therefore, the ads that did not appear in some user queries received low draws of quality scores. Our goal is to assess the robustness of our estimate of the empirical distribution of shocks to the scores to this assumption.

To estimate the marginal cost of advertisers under this assumption we created an additional dataset that contains user sessions where user queries contain pages with the search results beyond the first one. Then the ads that were considered for the first page of the search results but were not placed because of low draws of the scores can be considered for placement in the lower pages of the search results. By creating a database of the ads within the same user session we construct an approximation to the set of ads considered for a certain query. Then we use the sample of such long sessions to construct the empirical distribution of shocks to the scores.

Figure 7 demonstrates the differences between empirical distributions of shocks to the scores in our original dataset and in the new dataset. This graph shows that there is a large overlap between these distributions. An overlay with normal c.d.f. shows that both distributions have very high curtosis with a large probability mass around zero. The Kolmogorov-Smirnov test does not reject the null of the same distribution of the scores.

This similarity between the empirical distributions of shocks to the scores translates into the similarity of the estimated values per click for advertisers demonstrated in Table 9. The deviation of the estimated value with the adjustment for ad eligibility is the largest for the bidders in the bottom positions, which is a feature inherited from the approach taking into account eligibility of ads. The impact of the bias in estimation of the distribution of shocks to the scores is small. We show the histogram for the estimated distribution of logarithm of shocks

Figure 7: Cumulative distribution function of shocks to the scores



excluding top and bottom 1% quantiles. As one can see, even though the distribution has long right and left "tails", most of the distribution mass is concentrated about zero with a much larger kurtosis than the normal distribution. This means that even though the scores may take very small values, the probability of such extreme draws is small and is not sufficient to create large biases in the estimates of values.

Selection of user queries where the advertisement is considered: To study the effect of eligibility of ads for queries, we adapted our empirical methodology to the third assumption that the empirical distribution of observed competing advertisements is the distribution that advertiser anticipates.

We use the additional dataset on long user queries to estimate the scores. Then when we estimate the expected cost per click for the advertisers, we only use rival ad configurations where the ad of the advertiser of interest was observed. Note that the disadvantage of this approach (one reason why we did not adopt it for our baseline methodology) is that we ignore the fact that ads did not appear in certain user queries because their quality scores were low. Therefore, we may underestimate the impact of bids on participation since a higher bid may lead to a higher probability of participation.

We estimate the marginal cost for each advertiser by using only user queries where the advertisement of this advertiser was displayed. Then using the finite-point approximation to the derivative, we estimate the marginal cost for each bidder and recover valuations. The results of the analysis across three analyzed search phrases are demonstrated in Table 9. We show the mean log-values for all bidders and also separate the results for the top bidders (those whose average position is above 2) and the bottom bidders. One can see that the impact of the imposed change in the procedure on the overall mean is below 1%. A bidder-by-bidder analysis shows that for all bidders the confidence intervals for the valuations obtained using our main method and the method adjusted to the ad eligibility overlap. The deviation of the estimated value with the adjustment for ad eligibility is the largest for the bidders in the bottom positions. The main explanation for this result is that many ads that are at the bottom positions are appearing infrequently. This means that the sample sizes that can be used for the method taking into account the ad eligibility are small, leading to larger error in the estimated values.

Table 9: Log-values recovered from alternative estimation procedures

Baseline estimator			
Search phrase	Mean	Avg. position<2	Avg. position>2
#1	-3.25283	-2.67345	-3.47893
#2	-1.98829	-1.32934	-2.65464
Adjustment for ad eligibility			
Search phrase	Mean	Avg. position< 2	Avg. position> 2
#1	-3.32985	-2.84932	-3.68932
#2	-2.02362	-1.36723	-2.73123
Adjustment for selection bias			
Search phrase	Mean	Avg. position< 2	Avg. position> 2
#1	-3.25256	-2.65738	-3.46541
#2	-1.99152	-1.32137	-2.67109