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 2010.

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 keywords. Our model yields lower implied valuations and bidder profits than approaches that ignore uncertainty. 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, which does not arise in models that ignore uncertainty. For the keywords we study, the auction mechanism used in practice is not only less efficient than a Vickrey auction, but also for some keywords it raises less revenue.

* We acknowledge Dmitry Taubinsky, Nikhil Agarwal, Hoan Lee, Daisuke Hirata, Chris Sullivan, Maya Meidan for helpful comments and assistance with this paper. We also thank the participants at Cowles foundation conference (Yale), INFORMS meetings, 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 the competition affects the magnitude of the inefficiency.

The model proposed in this paper differs from existing economic models (e.g. [7], [17]) 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 equilibrium 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 keywords 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; and they may also vary with the characteristics of the individual user query, such as the time of day or the user's location.

[7] and [17] 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 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 that ignore uncertainty, which produce multiplicity of equilibria, we provide sufficient conditions for existence and

[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 never be applied.

uniqueness of equilibria, and we provide evidence that these conditions are satisfied empirically. 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.

Thus, the model that incorporates uncertainty, in addition to being more realistic, is more tractable and in many ways easier to analyze than the no-uncertainty alternative. Uniqueness of equilibria is especially useful for precise inference and counterfactual predictions in empirical applications.

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 several search phrases. We start by comparing the estimates implied by our model to those implied by prior approaches, 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 [7] and [17] show is different from a Vickrey auction. In a model without uncertainty, one of the main results of [7] and [17] 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 two of our search phrases, we find that the Vickrey auction raises up to 4though the efficiency difference is only about .5phrase, the GSP raises slightly more revenue.

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 [7] 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 of the ad in position 1. These effects are substantial, in that the "click-through rates" for the highest slot can be tens to hundreds of times higher than for slots a few positions below. The way in which CTRs diminish with position depends on the search phrase in question.

The mapping between bids, ranks and prices has varied over time. Initially, Yahoo! used first-price (pay-your-bid) auctions. As [8] show, in a static model, this auction does not have a pure strategy ex post equilibrium.[3] Since bidders can adjust their bids in real time, a configuration of bids that is not an ex post equilibrium creates incentives for some players to adjust their bids. Indeed, [8] demonstrate that in practice, bidding behavior was characterized by cycling.

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

[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.

[3] By ex post equilibrium, we mean a Nash equilibrium where each bidder's bid is a best response to the realized bids of other players; in a setting where each player is certain of her own value per click (no common values), the concepts of ex post equilibrium, complete information Nash equilibrium, and dominant strategy equilibrium all coincide.

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 generalized second price 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 platform, but as our empirical application generally has non-binding reserve prices, we ignore them for the baseline model in this paper.

3 The Model

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

We begin with a static model, where 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 few hours, a day or a week), and bidders pay their per-click bids when users actually click on their ads. 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.[4]

[4]Empirically, this assumption can be rejected for many search phrases, 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.

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 generalized second price 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 j to index slots of the position auction and i to index bidders. However, if the bidders are arranged in order of their score-weighted bids $b_i\, s_i$, bidder i will be in slot $j = i$. The per-click price p_i that the bidder in position i pays is determined as the minimum price such that the bidder remains in her position

$$p_i = \min\{b_i : s_i b_i \geq s_{i+1} b_{i+1}\} = \frac{s_{i+1} b_{i+1}}{s_i}.$$

Note that advertiser i 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 i in position i is $\alpha_i \gamma_{i+1} b_{i+1}$, which is what the platform would receive if instead, it had placed bidder $i+1$ in position i and charged bidder $i+1$ her per-click bid, b_{i+1}, for each click. So bidder i pays, in expectation, the per-impression revenue that would have been received from the next lowest bidder.

We assume that advertisers are interested in consumer clicks and each advertiser i has a value v_i associated with the consumer click. The profile of advertiser valuations in a particular market $\mathbf{V} = (v_1, \ldots, v_I)'$ 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.

3.2 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 [7] and [17]. We can write the expected surplus of advertiser i from occupying the slot j as

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

where bidder k_{j+1} is in slot $j+1$.

The existing literature, including [7] and [17], 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

$$\alpha_j \left(v_{k_j} - \frac{s_{k_{j+1}} b_{k_{j+1}}}{s_{k_j}} \right) \geq \alpha_l \left(v_{k_j} - \frac{s_{k_{l+1}} b_{k_{l+1}}}{s_{k_j}} \right), \; l > j$$
$$\alpha_j \left(v_{k_j} - \frac{s_{k_{j+1}} b_{k_{j+1}}}{s_{k_j}} \right) \geq \alpha_l \left(v_{k_j} - \frac{s_{k_l} b_{k_l}}{s_{k_j}} \right), \; l < j.$$

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.

[7] and [17] define a refinement of the set of equilibria, which [7] 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.

[7] consider a narrower class of envy-free equilibria in the same information environment as [17], 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

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

[7] 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.

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. In addition, the statistical algorithms can use characteristics of the user query as inputs, including features such as the time of day or the location of the user, and this will lead to different scores for different users. 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.

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.

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 "pause" and "reactivate" their campaigns in order to manage their own advertising expenditures, as well. (We do not formally model such constraints and objectives in this paper.) Second, many 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 the bidder 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 constructed 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.

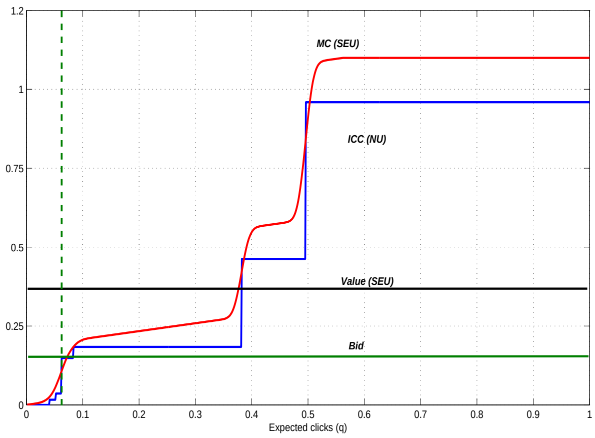
This smoothness reflects the general variability of the environment faced by the advertisers. For the search phrases we consider, the most commonly observed advertisements have a standard deviation of their position number ranging from about one third of a position, to about 2 positions.

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

Figure 1: Marginal/Incremental Cost Per Click in NU and SEU case



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^j_{ik} to be an indicator for the event that bidder i is in slot j and bidder k is in slot $j+1$, and let $C^i_{j,k}$ be a subset of C^i with cardinality j that contains k, representing the set of bidders above bidder i and as well as k. Let $b, \overline{s}, \varepsilon$ be vectors of bids, mean scores, and shocks to scores, respectively Then:

$$\Phi^j_{ik}\left(b,\overline{s},\varepsilon;C^i\right) = \sum_{C^i_{j,k}} \prod_{m\in C^i_{j,k}\setminus\{k\}} \mathbf{1}\left\{b_m\overline{s}_m\varepsilon_m > b_i\overline{s}_i\varepsilon_i\right\} \prod_{m\in C^i\setminus C^i_j} \mathbf{1}\left\{b_m\overline{s}_m\varepsilon_m < b_k\overline{s}_k\varepsilon_k\right\}\mathbf{1}\left\{b_i\overline{s}_i\varepsilon_i > b_k\overline{s}_k\varepsilon_k\right\}.$$

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

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

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

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

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 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. A pure strategy ex post equilibrium in this model may or may not exist; in the case where scores have a wide support, bidders will typically have a unique best response, and equilibrium existence is determined according to whether a solution exists to a system of nonlinear equations (the first-order conditions). In addition, the equilibrium may or may not be unique. The next section analyzes these issues.

3.4 Existence, Uniqueness, and Computation of Equilibrium 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. *Assume that shocks to the scores are i.i.d. across bidders with distribution $F_\varepsilon(\cdot)$, which does not have mass points and has an absolutely continuous density $f_\varepsilon(\cdot)$ with a compact support on $[\underline{\varepsilon}, \overline{\varepsilon}]$. $f_\varepsilon(\cdot)$ is twice continuously differentiable and strictly positive on $[\underline{\varepsilon}, \overline{\varepsilon}]$.*

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:

$$\frac{d}{d\tau}\, EU_i\left(\tau b, \overline{s}\right)\big|_{\tau=1} = -TE_i\left(b, \overline{s}\right), \tag{3.4}$$

that is, a proportional increase in all bids decreases bidder i's utility at the rate $TE_i\left(b, \overline{s}\right)$, 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. We formally establish the identity below in the proof of the next lemma.

The system of first-order conditions that are necessary for equilibrium is given by

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

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 *Let $EU(b, \overline{s})$ be the vector function of the expected utilities for bidders, let $TE(b, \overline{s})$ be the vector of total expenditure functions, and assume that $\frac{\partial}{\partial b'} EU(b, \overline{s})$ and TE are continuous in b. Suppose that $Q_i(v, \overline{s}_i) > 0$. Then a vector of bids b satisfies the first order necessary conditions for equilibrium (3.5) if and only if*

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

Proof: Denote the vector of mean scores of bidders $\overline{s}$. Denote the probability of bidders i and k from configuration $C^i \cup \{i\}$ being in positions j and $j+1$ by $G^j_{ik}(b, \overline{s}, C^i)$. Then $G^j_{ik}(b, \overline{s}, C^i) = \int \Phi^j_{ik}(b, \overline{s}, \varepsilon; C^i)\, dF_\varepsilon(\varepsilon)$, 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

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

The total expenditure can be computed as

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

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

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

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

which can in turn be rewritten as, for each i,

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

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

Q.E.D.

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

$$D(b, \overline{s}) = \frac{\partial}{\partial b'} EU(b, \overline{s}).$$

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

$$D_0(b, \overline{s})\, b = -TE(b, \overline{s}).$$

Lemma 1 transforms the system of first-order conditions into an equivalent form. We can then define a mapping $\beta(\tau)$ which, under some regularity conditions imposed on the payoff function, will exist in some neighborhood of $\tau = 1$:

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

for all bidders i. If the vector of expected utilities maintains regularity for all τ it will be possible to consider $\tau \in [0, 1]$. The next theorem establishes the conditions under which the mapping $\beta(\tau)$ 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 position auction in the SEU environment with a reserve price $r > 0$. Let $EU(b, \overline{s})$ be the vector function of the expected utilities for bidders, and let $TE(b, \overline{s})$ be the vector of total expenditure functions, and assume that D_0 and TE are continuous in b. Suppose that $Q_i(v, \overline{s}_i) > 0$, and that each EU_i is quasi-concave in b_i on V and for each b its gradient 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 $b \in V$ 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 *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 $i = 1, \ldots, I-1$. D_0 is non-singular on V if for each bidder her utility is strictly locally monotone in the bid of either bidder above or below her in the ranking or both.

We can prove this lemma, for instance, in case where $\frac{\partial EU_i}{\partial b_{i-1}} \neq 0$ for $i = 2, \ldots, I$ and $\frac{\partial EU_1}{\partial b_2} \neq 0$. The diagonal elements of the matrix $D_0(b, \overline{s})$ are zero. Therefore, we can compute the determinant $\det(D_0(b, \overline{s})) = -\frac{\partial EU_1}{\partial b_2} \prod_{i>2} \frac{\partial EU_i}{\partial b_{i-1}} \neq 0$, i.e. the matrix $D_0(\cdot)$ is non singular.

Equation (3.8) plays the 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 the 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 $\mathbf{F}(\mathbf{b}, \tau)$ 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 $\mathbf{F}$ using (3.8). If the payoff function is twice continuously differentiable and the equilibrium existence conditions are satisfied, then $\mathbf{F}$ has the desired properties. Details of the application of this method to our problem are in Appendix F.

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 $b_{-i}, \bar{s}$ 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 $AC_i(q_i)$, marginal cost curve $MC_i(q_i)$, and the required bid curve $Q_i^{-1}(q_i)$ 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 $MC_i(q_i)$. We then calculate the bidder's implied profits, illustrated in the figure with the shaded area.

Figure 2: Average cost, marginal cost, inverse click curve and value for a bidder on keyword # 1

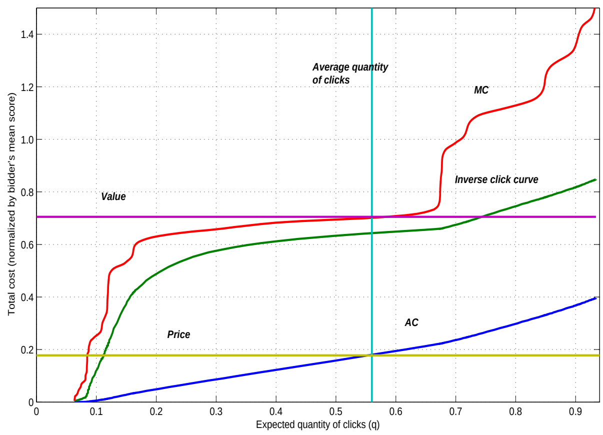
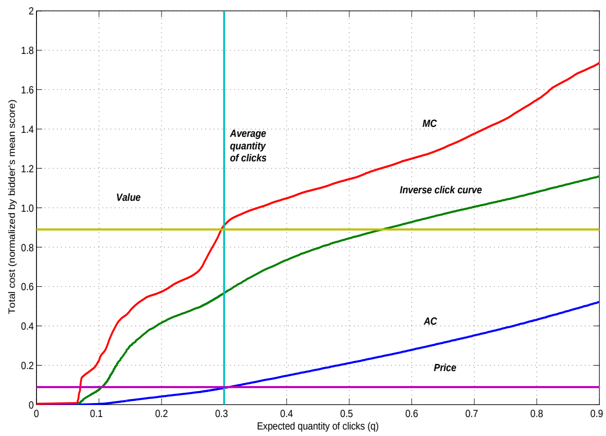


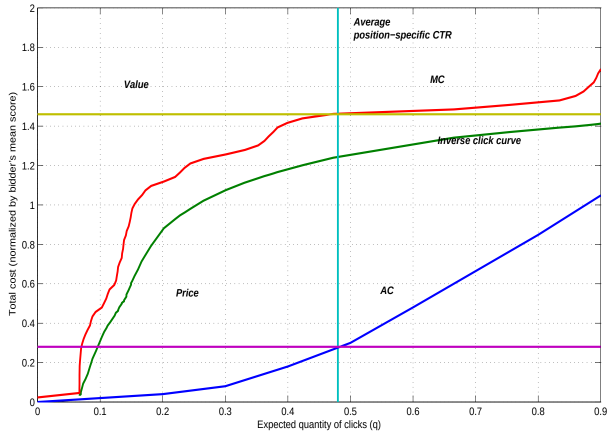
Figure 3: Average cost, marginal cost, inverse click curve and value for a bidder on keyword # 2



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.

Figure 4: Average cost, marginal cost, inverse click curve and value for a bidder on keyword # 3



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 criteria, although it should be noted that there can be Nash equilibria that are not envy-free and where bidders are not ranked in order of the score-weighted valuations, where (4.11) still holds, so the test does not definitively distinguish envy-free Nash equilibrium from ex post Nash equilibrium.

Following [17], 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 lowest position, $v_{k_J} = b_{k_J}$, 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 $Q_i(b_i)$ and $TE_i(b_i)$ are strictly increasing and differentiable, 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 $TE_i(b_i)$ 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 $Q_i(b_i)$ 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 entrants in each query. Then:*
(i) Bounds on the valuation for bidder i are given by (4.13), where $b_i' = Q_i^{-1}(Q_i(b_i); b_{-i}, \bar{s})$, and $b_i'' = \sup\{b_i''' : Q_i(b_i'''; b_{-i}, \bar{s}) = Q_i(b_i; b_{-i}, \bar{s})\}$.
(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's objective function will be similar once bids and valuations are adjusted for scores, 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 and the distributions of scores and entrants. Let $u_i(v_i^{NU}, b_i'; b_{-i}, \overline{s}_i, \varepsilon, C)$ be the bidder's utility for a particular user query, and for this proof include explicitly the bidders 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.

4.3.1 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 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

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

except for the lowest-ranked bidder who bids her valuation. This contrasts with the [7] 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

$$\alpha_2(v_2 - b_2) f_{1/\tilde{s}_2}\left(\frac{b_3 s_3}{b_2}\right) + \left[\alpha_1(v_2 - b_2) - \alpha_2\left(v_2 - \frac{b_3 s_3 b_2}{b_1 s_1}\right)\right] f_{1/\tilde{s}_2}\left(\frac{b_1 s_1}{b_2}\right) = 0 \tag{4.14}$$

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.

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 5 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.

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

Figure 5: Marginal cost, and recovered values for bidder in a monotone configuration

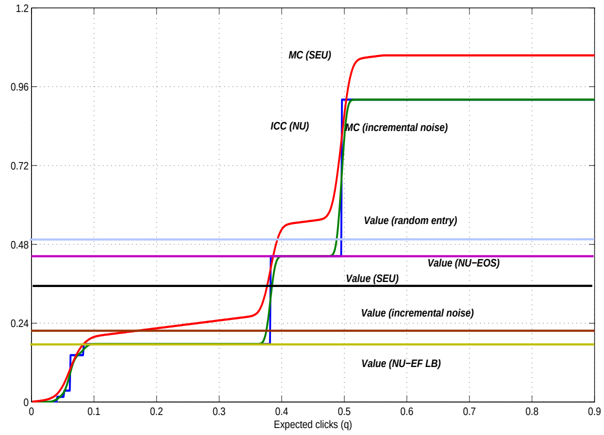
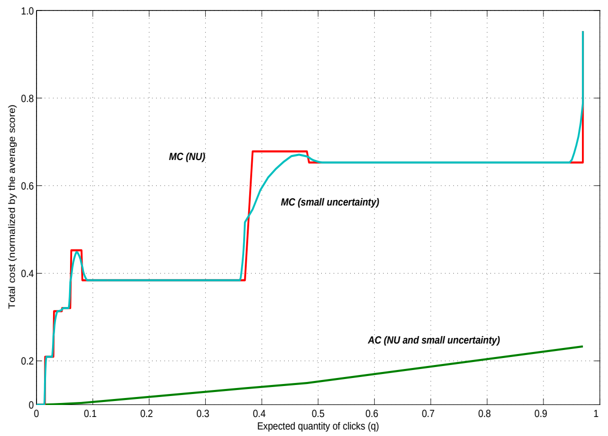


Figure 6: Marginal cost and total cost curves for bidder in a frequent configuration



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 simultaneous ads. 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 will assume that bidders have a separate valuation for each ad and the goal of our numerical procedure will be to recover valuations of the bidders corresponding to each ad.[5]

[5]Our empirical analysis shows that valuations for the ads of the same bidder are very close which can suggest that our empirical analysis is meaningful

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 $i = 1, \dots, I$. 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* $i = 1, \dots, I$. *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.

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 $\frac{1}{T} N_i \xrightarrow{p} P(i \in C_t)$. Similarly, we find that $\frac{1}{T}\sum_t \mathbf{1}\{i \in C_t\} \log s_{i,t} \xrightarrow{p} P(i \in C_t) \log \overline{s}_i$. 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 distribution of the shocks to the scores we use the empirical distribution

$$\widehat{F}_\varepsilon(\varepsilon) = \frac{1}{I}\sum_{i=1}^{I} \frac{1}{N_i} \sum_t \mathbf{1}\{i \in C_t\} \mathbf{1}\{\widehat{\varepsilon}_{it} \le \varepsilon\}.$$

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

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

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

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 we can use a simpler approach to evaluation of the bidder's first-order condition. We associated this case with the case of a substantial overlap of the click-weighted bids. We found that in this case we can characterize the first-order condition of the bidder as

$$v_i \frac{\partial Q_i\left(b_i, b_{-i}, \overline{s}\right)}{\partial b_i} - \frac{\partial TE_i\left(b_i, b_{-i}, \overline{s}\right)}{\partial b_i} = 0.$$

As a result, the value can be computed as a function of the own and rival bid as the marginal expected cost per click.

Each of the functions needed to recover the value can be estimated from the data. We use 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 make independent sampling from the empirical sample of observed configurations and estimated shocks to the scores $\left\{C_t^i,\, \widehat{\varepsilon}_{kt}\right\}_{t,k=1,\dots,I}$ 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 $C^i_{t^*}$ and the shocks to the scores for all bidders $\widehat{\varepsilon}_{1t^*}, \dots, \widehat{\varepsilon}_{It^*}$. For consistent inference we require that $\frac{N_i}{T^*} \to 0$ for all $i = 1, \dots, I$. Then for each such draw we compute the rank of the bidder of interest i as

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

We also compute the price paid by bidder i as

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

Then we estimate the total expenditure function as

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

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 $i = 1, \dots, I$ 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 [11] 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 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:*

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

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 C.

6 Data

For estimation we use a sample of data of auctions for three high-value search phrases (within the top several thousand search phrases on the advertising platform). The data is historical, for a three-month period sometime between 2006 and 2008, 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 500,000 searches per week between the three 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 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 (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. In addition, the ad platform's initial scores may be higher than the long-run scores for idiosyncratic reasons related to the scoring algorithms, in which case there is an incentive to continually create new advertisements (and if the converse holds, they will stick to existing advertisements).

We chose to handle the multiple advertisements by first treating them separately, and assuming that the advertiser takes the existence of multiple advertisements 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. 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 particular we find that the median (across advertisers) standard deviation of recovered valuations corresponding to the ads of the same advertiser is 14 times smaller than the standard deviation of valuations across advertisers. The median standard deviation of the per-advertisement profit per click (we will refer to this quantity from now on as *profit PC*) across advertisements of the same advertiser is 5 times smaller than the overall standard deviation. It is also possible that bidders change their bids during the course of the week, but this is surprisingly uncommon in our dataset. Indeed, the bids corresponding to the same advertisement are very stable even in the cross-week data. The maximum standard deviation of bids corresponding to a particular advertisement is less than a half of the standard deviation of bids in the entire set of data (the median standard deviation for the bids of the same advertiser is approximately 14 times smaller).

7 Estimates from Alternative Models

We use several alternative models for estimation: NU-Nash, NU-Envy-Free, NU-EOS, and SEU.

As a baseline case we use the SEU model. We recover valuations from this model using the empirical analog of the bidder's first-order condition.[6] The exact procedure for estimation of the first-order condition has been described in the previous section. We observe 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 Figures 2, 3, and 4 the estimated total 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.

We find empirically that estimated marginal cost curves are strictly increasing for each of the observed advertisements on each of the three 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). In each point we run the test testing 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 corresponds to the maximum achievable clickthrough rate).

Using our empirical algorithm we recover valuations for all advertisements featured in the auctions in the selected week of data for a selected keyword. We will show our results by 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 all three considered search phrases. In Table 4[7] we display basic statistics for log valuations for three analyzed search phrases. We notice that the search phrase #2 is the highest value phrase out of phrases that we analyze. However, the range of valuations remains comparable across the 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.

We also recover valuations under alternative information assumptions. In the no uncertainty cases and their refinements, 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

[6]The estimates were obtained under the assumption that displayed ads coincide with the ads that were considered for a particular query. This assumption will lead to some under-estimation of competition for the bottom positions and over-estimation of the values of the bottom bidders. We are in the process of accounting for this bias by considering additional data on advertiser participation.

[7]The tables with results for this and subsequent sections are displayed in Appendix G

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. [17] suggested computing an approximate weighted solution. We consider a weighted ICC as

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

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. In principle, similarly to [17] we find a large number of violations from monotonicity, all of them were corrected by the weighting. We 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 lower bound and approximately from 18% to 30% for the upper bound corresponding to the NU-EOS model for three considered search phrases. Moreover, the number of auctions that violate the value monotonicity auction-by-auction is quite high. For the considered search phrases it exceeds 25% with most violations occurring in the middle and the lowest positions. 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 examples of the intersection of the bounds for the same bidder. However, restriction of the dataset to a very limited period of time allows us to find up to 5% of cases (out of those not violating monotonicity) where the bounds intersect for the shortest considered period of time, which is equivalent to approximately 2 hours of search query logs. In this case the number of observations per advertisement ranges from 1 to 285 in our sample. For consistency of our analysis we choose the following approach. The set of recovered values corresponds to the bounds constructed from the incremental cost curve. We refer to the lower bound in this discussion as the NU-Envy-Free bound and we refer to the upper bound as the NU-EOS bound. For each bidder we can collect a set of values corresponding to different impressions. We choose to use the average over different values of the bounds for each bidder as estimates of valuations from the full-information model.

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. On Figures 8-10 we report the histograms for the mean absolute deviations for the weights from 1: $w = \frac{1}{J}\sum_{j=1}^{J}|1 - d_j^*|$ across the auctions. The observed deviations remain large across all keywords.[8] As it was mentioned before, we chose to use the average bid for the key phrase #1 as a numeraire. In Table 5 we report the means across bidders for the values recovered under different information assumptions. As a confirmation of results in Table 4, we can see that the key phrase #2 has the highest value. We can also see that this tendency maintains under different information assumptions. However, we can see that the values computed from the lower bound in the NU-EF framework 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.

Co-location of estimated values can be represented graphically. For values we use logarithmic scale for convenience of presentation. Figures 11-13 display the implied valuations for alternative models (or their bounds) for all advertisements observable in the selected subset of data for 3 search phrases, against the implied valuations from the SEU model, in logarithmic scale, for each of the three search phrases. In constructing the figures, we drop all auctions where bidding was inconsistent with an envy-free Nash equilibrium (that is, where the incremental costs are not increasing with higher positions); this occurs in percentages of queries ranging from 47% to 71% across search phrases. In general, it turns out that NU-Envy Free underestimates the values for most of the advertisements. On the other hand the NU-EOS underestimates the values for keyword #1 but overestimates them for keyword #2. We notice that across key phrases from 79% to 95% of SEU values are within the bounds provided by the NU framework. 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. However, 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.

Combining the recovered values with the data, we can compute the implied ex-post profits per click across the bidders by averaging the per-impression profit per click across different impressions. We noticed that in the NU framework, the value obtained under the NU-EF assumption under-estimates the valuation and the value obtained under the NU-EOS assumption over-estimates the valuation. This relationship between the recovered values translates into a similar relationship between expected profits per click of the bidders. Figures 14-16 are scatter plots for the profits per click under the NU assumption against the profit per click that is computed under the SEU assumption for 3 considered search phrases.

Profits per click can be aggregated at the level of the advertisement and at the level of query, to assess the overall implied total bidder profit per click in each of the models. We begin with the analysis of the

[8]The figures for this and subsequent section can be found in Appendix H

profit per click per advertisement. Table 6 illustrates mean, 25th, 50th, and 75th percentiles of per-query profit per click relative to the average cost per click for each of the different models. In this table, we use the values for the Envy-Free refinement that are computed from the weighted ICC curves such that the weighted ICC curves become monotone. We compute each entry by weighting both the profit per click and the score per click by the score and the position-specific clickthorugh rate. As a result, for the query-level aggregation we compute:

$$\frac{\text{Avg.}(value - CPC)}{\text{Avg.}(CPC)} = \frac{\sum\limits_{j=1}^{J} \alpha_j \, \bar{s}_{k_j} \left(v_{k_j} - CPC_{k_j,t} \right)}{\sum_t \sum\limits_{j=1}^{J} \alpha_j \, \bar{s}_{k_j} CPC_{k_j,t}}$$

To compute the profit per click in the NU-Envy free case we used the formula for the bounds which we described above. The upper bound will coincide with the NU-EOS case. In Table 6 we normalized the profits per click to the maximum per impression profit per click for the SEU case.

Across all search phrases the advertiser's profit per click per advertisement under the SEU assumption lies within the profit per click bounds provided by the NU framework. One can see that this relationship also maintains quantile-by quantile in most cases. We should note that NU framework is capable of producing negative profits per click. In fact, we attribute values to the bidders equal to the average between the recovered values in the NU framework. This implies that for some impressions this mean value can actually be below the price per click, especially if the true profit per click is small. We can next consider the properties of the profit per click per query shown in Table 7. We can see that qualitatively the results remain similar to the properties of profits per click computed at the advertisement level. However, we can notice that on the query level profits per click tend to have much larger inter-quartile ranges.

We can make an interesting observation that in the SEU framework the recovered per impression profits per click are very similar to the profits per click recovered on the advertisement level. The profits per click recovered from the bounds in the NU framework tend to be larger in the case of the NU-EOS bound and smaller in the case of the NU-EF concept.

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. To make the predictions across the models we drop all of the auctions that do not satisfy the monotonicity of the incremental cost curve. For the advertisements which do not appear in the first week of data, we hold the counterfactual bids equal to the observed bids. Figure 17 illustrates the results, where the x-axis is the expected revenue given the actual bids and prices in the auction, while 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. As a result, our model predicts very similar bids for the same advertisers in the second week of data. The analysis of NU-Envy-Free case is based on computing the bids based on the lower bound for valuations from the Envy-Free case fore all the advertisers who are present in both weeks. For the advertisers that are new in week 2 we fix the bids equal to their actual bids. Computation of the revenue is based on computing the bids from the lower bound for the equilibrium bid. We assume that the upper and the lower bounds on the bids of the new bidder in the NU case coincide and equal to their actual bids.

We recompute the equilibrium using the system of equations formed by bidder's first-order conditions. For the new bidders in the predictions and in the counterfactual exercises we fix the bids at the actual levels. Be begin with the system of equations for all the bidders who change their behavior in the counterfactual situation as

$$\widehat{v}_i \frac{\partial Q_i^* \left(b_i, b_{-i}, \overline{s}\right)}{\partial b_i} - \frac{\partial TE_i^* \left(b_i, b_{-i}, \overline{s}\right)}{\partial b_i} = 0.$$

In this system of equations we use estimated valuations $\widehat{v}_i$ while the quantity of clicks and the total expenditure function are recomputed for the new environment to include new bidders assuming that their scores are generated from the same distribution of scores. Consider the case where the number of bidders changes and new bidders arrive. Suppose that S is the set of additional bidders and the new total number of bidders is $I' = I + \#S$; we add the new bidders to all previously observed configurations in the sample. Then the total expenditure and the quantity of clicks as functions of bids are computed using the simulated sample analogs in the same ways as for estimation by forming a simulated sample of size T^* and averaging over the simulated draws

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

and

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

We also use a higher-order derivative formula to approximate the ratio on the right-hand side of the system of equations.

$$\hat{v}_i = \frac{-TE_i^*(b_i - 2\tau_N, b_{-i}, \overline{s}) + 8TE_i^*(b_i - \tau_N, b_{-i}, \overline{s}) - 8TE_i^*(b_i + \tau_N, b_{-i}, \overline{s}) + TE_i^*(b_i + 2\tau_N, b_{-i}, \overline{s})}{-Q_i^*(b_i - 2\tau_N, b_{-i}, \overline{s}) + 8Q_i^*(b_i - \tau_N, b_{-i}, \overline{s}) - 8Q_i^*(b_i + \tau_N, b_{-i}, \overline{s}) + Q_i^*(b_i + 2\tau_N, b_{-i}, \overline{s})}.$$

Note that the population analogs of the functions on the right-hand side are represented by smooth and continuous functions of the bid profile. The marginal cost is monotone increasing in the own and rival bids. Moreover, the sum of effects of rival bids on the marginal cost is smaller than the own effect. We use the homotopy method described in Appendix F to compute the counterfactual bid vector.

We perform the same computation for the upper bound corresponding to the NU-EOS case. Figures 18 and 19 demonstrate the scatterplot of predicted versus actual revenue for such auctions. Note that NU-Nash and NU-Envy-Free 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. On Figure 17 we show the result of inferring revenues in week 2 from valuations recovered in week 1 in the SEU model. On Figure 18 we show the results of using the values from the NU-Envy-Free lower bound and using NU-EF lower bound to infer bids. On Figure 19 we show the results of using the values from the NU-EOS model and using NU-EOS model. On Figure 19 we show the results of using the values from the SEU model and using NU-EOS model to infer bids. We do the same exercise for all recovered values and two models in the NU context.

The revenue predicted by the NU-Envy-Free model tends to understate the actual revenue. On the other hand, the revenue predicted by the NU-EOS model under-estimates the revenue. In most cases, however, the revenue in the SEU case remains within the bounds. There are two reasons for this effect. First, by adjusting the cost curves forcing their monotonicity, we make the structure of the recovered values comparable across the models. As a result we can expect that in a large proportion of the observed auction the best-response correspondences in the SEU case will be a subset of the best-response correspondences in the NU case. Second, new entry into the auctions and our assumption of the degeneracy of the bounds for the best-responses of the new entrants tightens the bounds and makes them closer to the actual bids. The results turn out to be even more robust for the predictions for the average cost per click for the bidders. We summarize our results in Table 8. It shows the parameters of 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 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 the 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 generalized second-price 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 generalized second price auction is strictly less efficient than the Vickrey auction, so long as there is sufficient uncertainty in the environment.

Table 9 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 generalized second price auction. To simplify the comparisons 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 but not insignificant. For our first two search phrases, the efficiency difference is about half of one percent, while it is about 4% for search phrase 3.

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 phrases 1 and 2, the revenue differences are larger in magnitude than the efficiency differences, in the same direction: the Vickrey auction is more efficient, and raises 6-8% more revenue. The revenue gains appear throughout the distribution of queries.

In contrast, for search phrase 3, the Vickrey auction raises 1.2% less revenue, despite being 4% more efficient. The Vickrey auction does raise higher revenue for the median query (ranked by revenue), but in the lower and higher quantiles, the generalized second price auction is superior.

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.

The source of the inefficiency of the generalized second price auction is the asymmetric gaps between values and bids for different bidders. In the next section, we explore those gaps and the sources of the asymmetries in more detail.

8.3 Competition, Elasticities and Profits PC

In this section, we examine the properties and implications of the estimated elasticities. First, we observe that there is substantial variation across bidders and across search phrases in the elasticity of the average cost curve. Tables 10-12 provide summary statistics on the elasticity faced by all the advertisements related to three analyzed search phrases 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 $\frac{\text{Value-CPC}}{\text{CPC}}$ will be equal to the inverse of the elasticity). Tables 10-12 also show that the average elasticity of cost across advertisers remains similar across positions. The gap between bid and payment is large and implies that the bids substantially exceed the payment. On the other hand, the bid tends to be close to 2/3 of the value for all positions.

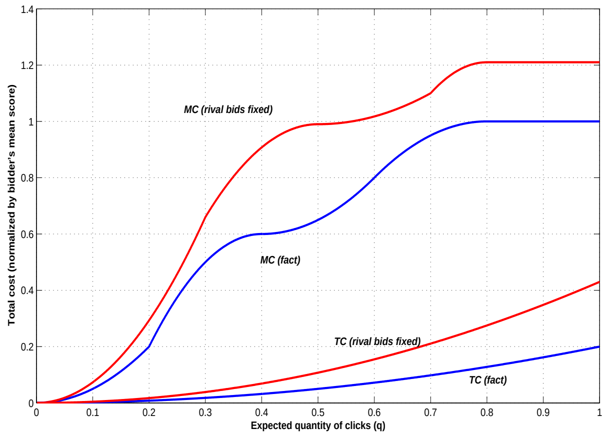
Analysis of the results for the second and the third search phrases demonstrates that the elasticity of the average cost-per-click tends to increase towards the lower positions. One possible reason for that is that higher positions are occupied by large advertisers with high bids and the competition occurs on the lower position level.

We study the structure of the competition in the market for the first search phrase using two counterfactual experiments. In the first experiment we study the role of the assigned scores in the competitiveness of the bidders. We consider the bidder who is most frequently placed in top 1 position. Then we increase the scores for all other bidders by 20% and study the impact of that increase on the marginal cost of the bidder. We first analyze the impact of such change by keeping the bids of all the competing bidders fix and allowing on the considered top bidder to re-adjust her bid.

In the second experiment, we increase the number of rival bidders by 20% considering the same set of auctions and the same top bidder. The bids for the new bidders are generated randomly from the empirical distribution of bids of bidders in the bottom positions and the scores are generated from the score distribution.

The results of the experiments are demonstrated in Tables 13-15. The top line of the table contains the factual observed mark-ups for the bidders and the two bottom lines correspond to the experiments. One can see that in both experiments the profit per click of the considered top bidder has substantially decreased as compared to the baseline. On the other hand the average profits per click and markups remained similar. We can illustrate this table using the marginal cost curve for the bidder. The following figure 7 illustrates the induced changes to the marginal cost curves for one of the top bidder for the first search phrase. We can see this change results in a shift of the total cost per click and the marginal cost

Figure 7: Changes in the marginal cost due to increased number of rivals



up. The change is not sufficient for the bidder to move to a lower position. However, the optimal bid can change by a substantial amount and so is the profit per click of the bidder.

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 utilize the fact that queries arrive more quickly than advertisers can change their bids, and advertisers cannot perfectly predict quality scores. In contrast to existing models that ignore uncertainty, 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 given advertiser valuations and the distribution of uncertainty. We develop an econometric methodology for estimating our structural model. Using our model we can recover the bidder valuations, which we show is consistent and asymptotically normal, and we provide Monte Carlo analysis to assess the small sample properties of the estimator. Finally, we apply the model to historical data for several keywords. Our model yields lower implied valuations and bidder profits than approaches that ignore uncertainty. 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.

References

[1] 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.

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

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

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

[5] 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.

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

[7] 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.

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

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

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

[11] Hong, H., Mahajan, A., and D. Nekipelov (2009): " Extremum Estimation and Numerical Derivatives", Stanford and UC Berkeley Working paper.

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

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

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

[15] 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.

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

[17] 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

$$\tau \frac{d}{d\tau} EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right) = TE_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right), \quad \text{for all } i. \tag{A.15}$$

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

$$\frac{\partial}{\partial b_j} EU_i\left(0, 0, \overline{s}\right) = \lim_{\varepsilon \downarrow 0} \frac{\partial}{\partial b_j} EU_i\left(\varepsilon, 0, \overline{s}\right). \tag{A.16}$$

Lemma 1 establishes that (A.15) 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

$$\frac{\partial}{\partial b'} EU\left(b, \overline{s}, r\right) b + \frac{\partial}{\partial r} EU\left(b, \overline{s}, r\right) r = -TE\left(b, \overline{s}, r\right). \tag{A.17}$$

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

$$\frac{d}{d\,\tau} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right)\big|_{\tau=1} = -TE_i\left(b, \overline{s}\right) - r\frac{\partial}{\partial r} EU_i\left(b, \overline{s}, r\right). \tag{A.18}$$

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

$$\widetilde{TE}_i\left(b_i, \tau b_{-i}, \overline{s}\right) = TE_i\left(b, \overline{s}\right) + r\frac{\partial}{\partial r} EU_i\left(b, \overline{s}, r\right). \tag{A.19}$$

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 $i = 1, \dots, N$ 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 the 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

$$\frac{\partial EU_i\left(\beta_i(\tau), \tau\beta_{-i}(\tau), \overline{s}\right)}{\partial b_i} = 0, \tag{A.20}$$

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

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

for any b in the support of these functions (with the derivative of the payoff function continuously extended to the origin by (A.16)). Function $\widetilde{TE}_i(\cdot)$ is defined in (A.19).

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.21) is a direct consequence of homogeneity, it will be satisfied for any τ and any b in the support of bids

$$b_i \frac{\partial}{\partial b_i} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right) + \sum_{j \neq i} \tau b_j \frac{\partial}{\partial b_j} EU_i\left(b_i, \tau b_{-i}, \overline{s}\right) = -\widetilde{TE}_i\left(b_i, \tau b_{-i}, \overline{s}\right). \tag{A.22}$$

This equation will also be valid for $\beta(\tau)$ defined by (A.20) (if it exists). Substituting $b = \beta(\tau)$ into (A.22), 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.20). 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.20). 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 [5] this means that the differential equation (A.15) has a continuous solution in some neighborhood of $\tau = 1$. This proves the necessity of the statement.

As a result, we 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.15). This proves (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.15) 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.20), 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:

$$\sum_j \left(\tau + (1-\tau)\mathbf{1}_{j=i}\right) \frac{\partial}{\partial b_j} EU_i\left(b_i, \tau\, b_{-i}, \overline{s}\right) \dot{\beta}_j = \frac{TE_i\left(b_i, \tau\, b_{-i}, \overline{s}\right)}{\tau} - \sum_{j \neq i} \frac{\partial}{\partial b_j} EU_i\left(b_i, \tau\, b_{-i}, \overline{s}\right) b_j, \tag{A.23}$$

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

Equation (A.23) 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.23) 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

$$\frac{\partial}{\partial b'} EU\left(M\, x, \overline{s}\right)\, \dot{M} = -\frac{1-\tau}{\tau} \operatorname{diag}\left\{\frac{\partial}{\partial b_i} EU_i\left(M\, x, \overline{s}\right)\right\} M(\tau). \tag{A.24}$$

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 $\dot{x}$:

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

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 [16], 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 [5] and [16].

In equation (A.24) $\frac{\partial}{\partial b'} EU\left(M\,x,\bar{s}\right)$ 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 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:

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

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

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

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

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

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

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

We omitted all the terms of the smaller order than $o_p\left(\left(T\tau_T\right)^{-1/2}\right)$ using the assumption regarding the rate of the numerical differentiation. Finally, using the structure of total expenditure and expected quantity of clicks, we can write:

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

Then if $\Omega = \operatorname{Var}\left(\frac{u_i\left(v_i,b_i+\tau_T;b_{-i},\overline{s}_i,\widehat{\varepsilon}_{it},C_t\right)-u_i\left(v_i,b_i-\tau_T;b_{-i},\overline{s}_i,\widehat{\varepsilon}_{it},C_t\right)}{\sqrt{\tau_T}}\right)$, it follows that the and i.i.d. Assumption 2, bootstrap is valid by [13] and

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

C 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 [9], while general results for identification in the auction settings are given in [3]. 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

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

(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 $4\,(N-1)$ 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:

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

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 [6]. As a result, the estimates of identified set of valuations will be described by Theorem 2.1 in [6].*

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 $\{C_{t^*,k,j-1}^i\}_{k=1}^{K(T)}$ 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

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

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

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

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 [6]. The confidence sets can be recovered using the tools developed in [12].

D 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 $U\left[0, 1\right]$. 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

Table 1: 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 2: 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

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.

E 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 $F_{\gamma,i}(\cdot \mid z)$ and the distribution of the estimated slot-specific click-through rate is denoted $F_{\alpha,j}(\cdot \mid z)$. The distribution of bidder valuations is also a common knowledge among bidders. The following proposition establishes the fact that we can recover distributions

Table 3: 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

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.

F Computing equilibria via numerical continuation

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

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

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

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

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

$$\sum_k a_k^i \, \dot{b}_k = c^i, \tag{F.27}$$

and

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

and

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

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

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

where the elements of matrix $A\left(\mathbf{b}, \tau\right)$ can be computed as $A_{ik}\left(\mathbf{b}, \tau\right) = a_k^i$. 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 $\lim\limits_{\delta \to +0} \delta^{-\alpha} \frac{\partial^2 EU_i\left(b_i, b_{-i}\right)}{\partial b_i \partial b_j}\Big|_{\|b\| = \delta} < \infty$ 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

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

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

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

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

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

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

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

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

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

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

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

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

G Tables

Table 4: Log-values recovered from the computational algorithm

Search phrase	Mean	25%	50%	75%
#1	-.7527981	-1.287738 3	-.7610942	-.1050781
#2	1.892349	1.271759	2.10132	2.575151
#3	-.6609135	-1.309133	-.8079711	-.1149696

We report the quantiles of the logarithm of valuations recovered from solving the first-order condition for each bidder in the SEU environment across bidders. The values are normalized by the highest observed bid for search phrase #1

Table 5: Means of valuations for different models across search phrases

Model; Search phrase	#1	#2	#3
NU-Envy Free	.4783782	3.223342	.4147022
NU-EOS	1.184797	5.95363	.9659285
SEU	1.093177	5.1296	.8188958

We report the means of logarithms of recovered valuations across search phrases and bidders. The valuations in the SEU environment are obtained by solving the bidder's first-order condition. The valuations in the NU environment are obtained by computing the weights for the ICC curves that make configurations monotone. The values are normalized by the highest observed bid for search phrase #1.

Table 6: Recovered normalized profit per click $\frac{\text{Avg.}(value-CPC)}{\text{Avg.}(CPC)}$ (aggregation at the advertisement level)

Model	Mean	25%	50%	75%
		Search phrase #1		
SEU	1.995002	.381061	1.030544	2.040322
NU-EOS	2.294112	.4381046	1.225826	2.641843
NU-Envy Free LB	1.1638212	.0402093	.2037002	.5073325
		Search phrase #2		
SEU	2.1402	.6332364	1.759544	3.485528
NU-EOS	3.113286	.774698	2.4309701	3.576068
NU-Envy Free LB	.1532421	.0377244	.0919325	.2888345
		Search phrase #3		
SEU	2.152566	.8692392	1.239167	1.843645
NU-EOS	2.371822	.7725573	1.777851	2.723807
NU-Envy Free LB	.4713152	.1273392	.3558505	.6382902

Reported profits per click are averaged at the bidder (advertisement) level. We use the values obtained from our computational algorithms. The valuations in the SEU environment are obtained by solving the bidder's first-order condition. The valuations in the NU environment are obtained by computing the weights for the ICC curves that make configurations monotone.

Table 7: Recovered weighted profits per click $\frac{\text{Avg.}(value-CPC)}{\text{Avg.}(CPC)}$(aggregation at the impression level)

Model	Mean	25%	50%	75%
		Search phrase #1		
SEU	1.226478	.956925	1.160303	1.388966
NU-EOS	2.264668	1.320135	1.659808	2.114972
NU-Envy-Free LB	.3648388	.249972	.3322766	.4268979
		Search phrase #2		
SEU	.603822	.2271468	.561978	.963982
NU-EOS	.8147091	.398568	.6503538	.9690881
NU-Envy-Free LB	.1015507	.0802903	.0233549	.0355081
		Search phrase #3		
SEU	2.132596	.8826663	1.09802	1.784544
NU-EOS	2.186924	1.791808	2.08713	2.386247
NU-Envy-Free LB	.365921	.2546077	.359719	.4714175

Reported profits per click are averaged at the impression level. We use the values obtained from our computational algorithms. The valuations in the SEU environment are obtained by solving the bidder's first-order condition. The valuations in the NU environment are obtained by computing the weights for the ICC curves that make configurations monotone.

Table 8: Mean deviations of the predicted revenues from the true revenues (normalized by actual mean revenues)

Model (values)	Mean	25%	50%	75%
		Search phrase #1		
SEU (SEU)	2.4893939	.0240292	.9483932	2.5784334
NU-EOS (SEU)	3.2292928	.2583932	1.3392231	3.721294
NU-Envy Free LB (SEU)	2.5782184	.0473928	.9132732	3.443231
NU-EOS (NU-EOS)	5.615954	.4347139	1.755482	5.611716
NU-Envy Free LB (NU-EF)	1.858746	-.095282	.4155265	1.860668
		Search phrase #2		
SEU (SEU)	3.247832	.2398351	.7783552	2.278349
NU-EOS (SEU)	11.58404	.2632192	1.237076	5.828762
NU-Envy Free LB (SEU)	3.655106	-.0301596	.646016	2.025008
NU-EOS (NU-EOS)	12.08087	1.309723	3.406704	7.418589
NU-Envy Free LB (NU-EF)	2.774772	-.1279501	.4839024	1.679651
		Search phrase #3		
SEU (SEU)	3.078249	.0142574	.7893584	2.139244
NU-EOS (SEU)	4.037667	.2345959	1.276185	2.180654
NU-Envy Free LB (SEU)	3.90874	.2132468	.7608118	1.858421
NU-EOS (NU-EOS)	6.833638	.287898	.8622311	2.685828
NU-Envy Free LB (NU-EF)	2.650613	-.0532067	.4336354	1.226672

To compute the numbers in this table we use the values obtained from solving bidder's first-order condition in the SEU environment. Then we compute equilibrium bids corresponding to SEU and NU environment. Reported numbers reflect mean-squared deviation of the revenue per impression predicted using equilibrium concepts in the NU and SEU environments and actual revenues. All numbers are normalized by the mean actual revenue per impression for each search phrase.

Table 9: Predicted counterfactual revenues and welfare for the SEU generalized second price auction model versus the NU-EOS (equivalent to query-by-query Vickrey auctions) model, using SEU values and actual bidder configurations for both models

Model (values)	Mean	25%	50%	75%
Search phrase #1				
Revenue SEU (SEU)	.1423501	.0155821	.1319029	.2973121
Revenue NU-EOS (SEU)	.1540002	.0159291	.1322911	.3399102
Welfare SEU (SEU)	.4593142	.2892011	.4410259	.601212
Welfare NU-EOS (SEU)	.4619217	.2912914	.4452871	.6293282
Search phrase #2				
Revenue SEU (SEU)	1.216925	.3027593	1.212913	1.967921
Revenue NU-EOS (SEU)	1.285954	.4549824	1.297483	1.957212
Welfare SEU (SEU)	3.658921	2.127683	3.987584	4.729252
Welfare NU-EOS (SEU)	3.678215	2.173752	4.023745	4.752598
Search phrase #3				
Revenue SEU (SEU)	.1718925	.0001793	.1572834	.4982731
Revenue NU-EOS (SEU)	.1691242	.0001395	.1783529	.494372
Welfare SEU (SEU)	.2835921	.09231292	.2274856	.7672827
Welfare NU-EOS (SEU)	.2942763	.1024853	.2472872	.7874526

To compute the numbers in this table we use the values obtained from solving bidder's first-order condition in the SEU environment. Then we compute equilibrium bids corresponding to SEU and NU environment. Reported numbers correspond to mean per impression revenues and welfare for considered search phrases. All numbers are normalized by the mean actual revenue per impression for each search phrase.

Table 10: Characteristics of competition for search phrase #1

	Mean	Mean	Elasticity			
Avg. ranking	$\frac{\text{Value-Bid}}{\text{CPC}}$	$\frac{\text{Bid-CPC}}{\text{CPC}}$	Mean	25%	50%	75%
[1, 1.5)	1.265633	.2015842	1.506112	.762117	1.472858	2.250106
[1.5, 2.5)	1.224156	.3067305	1.598213	1.272535	1.565803	2.477033
[2.5, 4)	1.426913	.238303	1.568633	.987747	1.589449	2.371734
[4, 5.5)	1.38519	.3506651	1.632358	.992719	1.696646	2.063183
[5.5, 8)	1.874973	.2203497	2.023551	.941602	1.717833	2.348808

We report mean elasticities of the MC curve corresponding to bidders whose average position is in the displayed bracket. We also report mean per impression markup and revenue and their quantiles.

Table 11: Characteristics of competition for search phrase #2

Avg. ranking	Mean $\frac{\text{Value-Bid}}{\text{CPC}}$	Mean $\frac{\text{Bid-CPC}}{\text{CPC}}$	Elasticity Mean	Elasticity 25%	Elasticity 50%	Elasticity 75%
[1, 1.5)	1.88835	.8059352	2.64721	1.90698	2.430392	4.32612
[1.5, 2.5)	1.5066034	.344606	2.054345	1.710727	1.954463	2.235109
[2.5, 4)	1.076497	.2314135	1.191285	1.066815	1.665737	2.182481
[4, 5.5)	1.201434	.3027201	1.539007	1.16887	1.505857	1.941993
[5.5, 8)	1.154609	.2267382	1.263357	1.328631	1.391573	2.128564

We report mean elasticities of the MC curve corresponding to bidders whose average position is in the displayed bracket. We also report mean per impression markup and revenue and their quantiles.

Table 12: Characteristics of competition for search phrase #3

Avg. ranking	Mean $\frac{\text{Value-Bid}}{\text{CPC}}$	Mean $\frac{\text{Bid-CPC}}{\text{CPC}}$	Elasticity Mean	Elasticity 25%	Elasticity 50%	Elasticity 75%
[1, 1.5)	1.1103034	.1155944	1.2570988	.9980923	1.2198768	1.999362
[1.5, 2.5)	1.3237916	.4541549	1.561206	1.559298	1.559298	1.566089
[2.5, 4)	1.5599176	.2661763	1.934037	1.739429	1.769591	2.047367
[4, 5.5)	1.3842787	.1636832	1.675953	1.531604	2.031604	2.156488
[5.5, 8)	1.8510012	.217807	2.022899	1.2036649	2.0376654	3.104773

We report mean elasticities of the MC curve corresponding to bidders whose average position is in the displayed bracket. We also report mean per impression markup and revenue and their quantiles.

Table 13: Counterfactual behavior of top bidder for search phrase #1

Top bidder Bid	Top bidder Avg.Position	Top bidder ProfitPC	All bidders Bid	All bidders Avg.Position	All bidders Profit PC	SocialWelfare	Ad platform Revenue	Ad platform % Soc. welf. receivedby ad platform
Fact								
6.235762	1	.3129024	4.872921	3.782131	.1243753	.5472249	.1492331	.2727089
Increased competition: top bidder changes bid, welfare of new bidders included								
5.990234	1.673962	.3002127	4.976542	4.410475	.0943098	.5512012	.1504498	.2728435

To compute the numbers in this table we use the values obtained by solving first-order conditions for bidders in the SEU environment. Then we compute the bids by solving for a new equilibrium with the bids of new entrants fixed.

Table 14: Counterfactual behavior of top bidder for search phrase #2

Top bidder			All bidders				Ad platform	
Bid	Avg.Position	ProfitPC	Bid	Avg.Position	Profit PC	SocialWelfare	Revenue	% Soc. welf. receivedby ad platform
Fact								
52.343965	1	2.374845	35.637842	3.266937	1.129363	4.817875	1.324213	.2748551
Increased competition: top bidder changes bid, welfare of new bidders included								
49.987542	1.578332	2.350315	36.967433	3.848345	1.053244	5.137943	1.443214	.2808932

To compute the numbers in this table we use the values obtained by solving first-order conditions for bidders in the SEU environment. Then we compute the bids by solving for a new equilibrium with the bids of new entrants fixed.

Table 15: Counterfactual behavior of top bidder for search phrase #3

Top bidder			All bidders				Ad platform	
Bid	Avg.Position	ProfitPC	Bid	Avg.Position	Profit PC	SocialWelfare	Revenue	% Soc. welf. receivedby ad platform
Fact								
3.754932	1	.1129721	2.287534	3.043752	.0624725	.1786313	.0588345	.3293612
Increased competition: top bidder changes bid, welfare of new bidders included								
3.564321	1.594967	.1123745	2.457386	3.407652	.05502137	.1817662	.0596821	.3283457

To compute the numbers in this table we use the values obtained by solving first-order conditions for bidders in the SEU environment. Then we compute the bids by solving for a new equilibrium with the bids of new entrants fixed.

H Graphs

Figure 8: Mean absolute deviation of ICC weights from 1 for search phrase #1

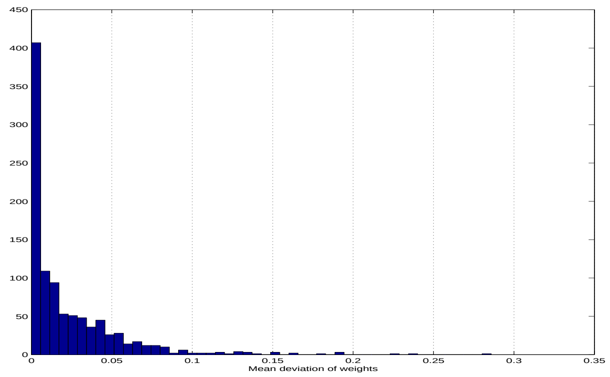


Figure 9: Mean absolute deviation of ICC weights from 1 for search phrase #2

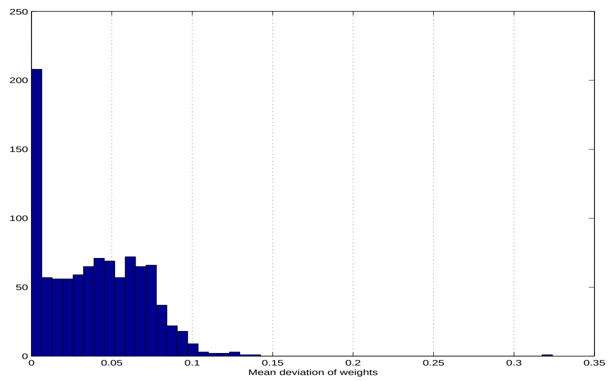


Figure 10: Mean absolute deviation of ICC weights from 1 for search phrase #3

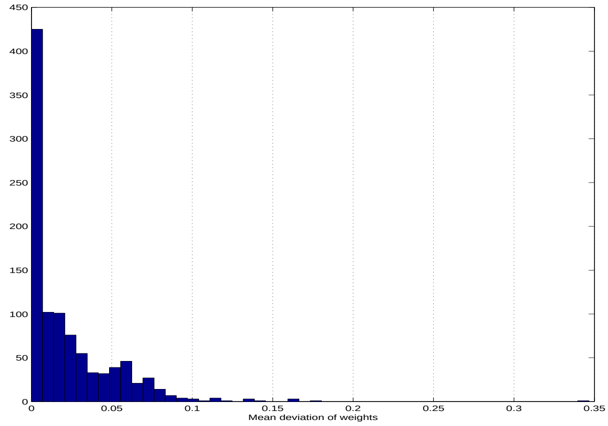


Figure 11: Valuations in NU-models plotted against the values in SEU for keyword #1

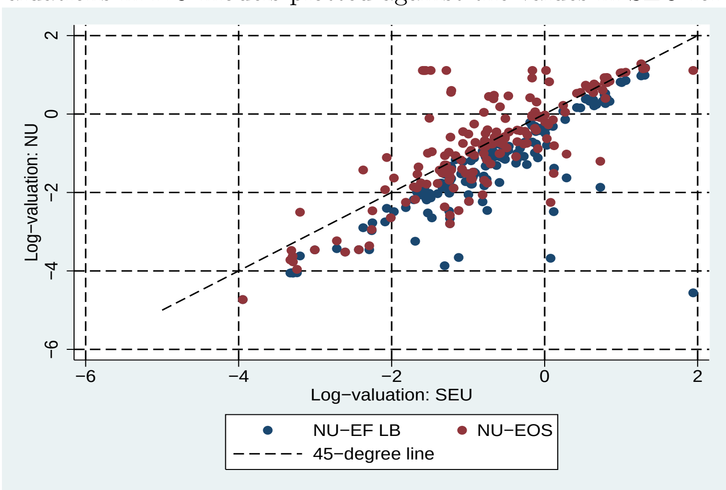


Figure 12: Valuations in NU-models plotted against the values in SEU for keyword #2

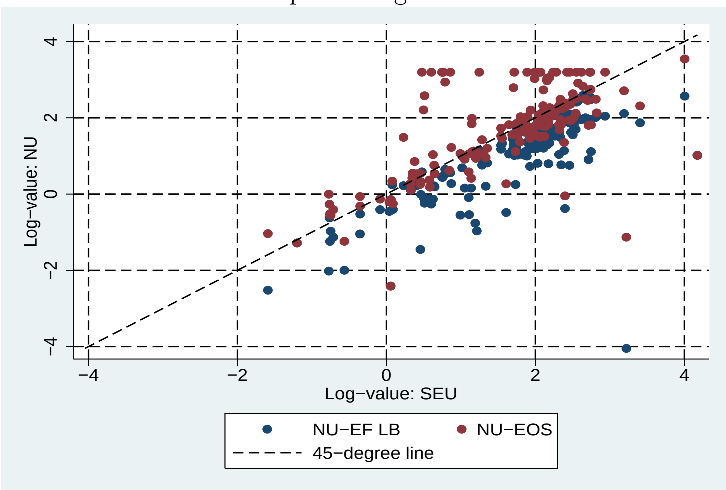


Figure 13: Valuations in NU-models plotted against the values in SEU for keyword #3

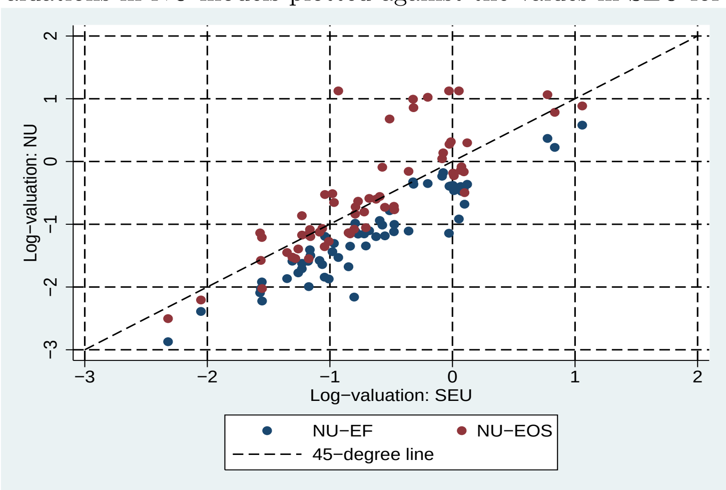


Figure 14: Logarithm of profit per click in NU-models plotted against the profits per click in SEU for search phrase #1

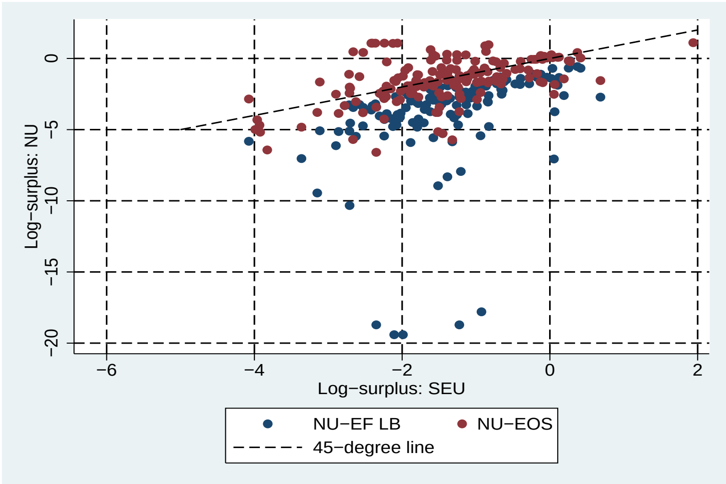


Figure 15: Logarithm of profit per click in NU-models plotted against the profits per click in SEU for search phrase #2

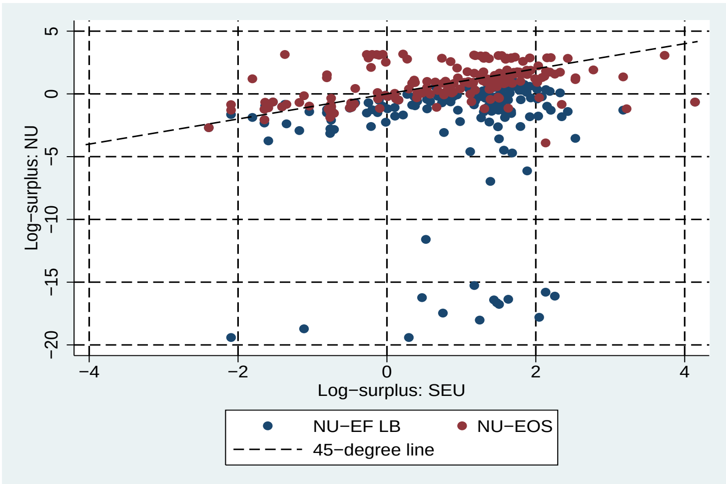


Figure 16: Logarithm of profit per click in NU-models plotted against the profits per click in SEU for search phrase #3

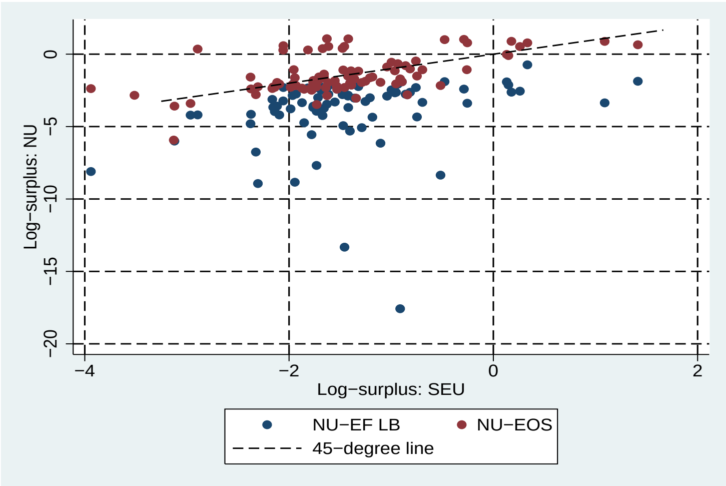


Figure 17: Logarithm of predicted revenues in SEU for keyword #1

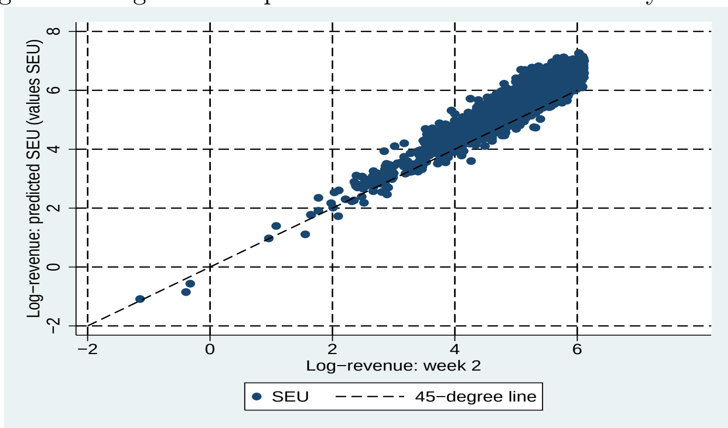


Figure 18: Logarithm of predicted revenues in NU for keyword #1

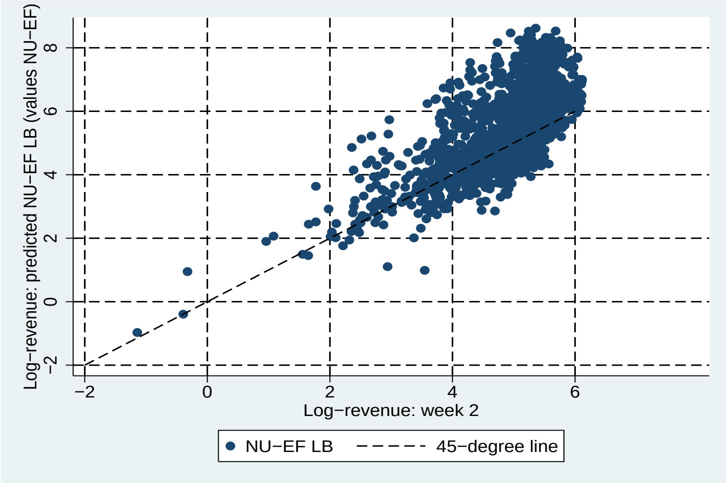


Figure 19: Logarithm of predicted revenues in NU for keyword #1