DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

TO SCORE OR NOT TO SCORE? ESTIMATES OF A SPONSORED SEARCH AUCTION MODEL

Yu-Wei Hsieh
University of Southern California

Matthew Shum
Caltech

Sha Yang
University of Southern California

SOCIAL SCIENCE WORKING PAPER 1402

February 2015

To Score or Not to Score? Estimates of a Sponsored Search Auction Model

Yu-Wei Hsieh   Matthew Shum   Sha Yang

Abstract

We estimate a structural model of a sponsored search auction model. To accommodate the "position paradox", we relax the assumption of decreasing click volumes with position ranks, which is often assumed in the literature. Using data from "Website X", one of the largest online marketplaces in China, we find that merchants of different qualities adopt different bidding strategies: high quality merchants bid more aggressively for informative keywords, while low quality merchants are more likely to be sorted to the top positions for vague keywords. Counterfactual evaluations show that the price trend becomes steeper after moving to a score-weighted generalized second price auction, with much higher prices obtained for the top position but lower prices for the other positions. Overall there is only a very modest change in total revenue from introducing popularity scoring, despite the intent in bid scoring to reward popular merchants with price discounts.

JEL classification numbers: D44; D47; C11; C15

Key words: Sponsored-search advertising; Auctions; Market design; Two-sided Matching; Bayesian estimation

To Score or Not to Score? Estimates of a Sponsored Search Auction Model

Yu-Wei Hsieh* Matthew Shum† Sha Yang‡

1 Introduction

Presently, sponsored search is, by far, the most salient online advertising format. A report from the Internet Advertising Bureau in 2013 shows that sponsored-search advertising is now generating more than $18 billion per year, nearly half of the annual online advertising spending in US. Due to its effectiveness and other benefits, sponsored advertising, which was pioneered by search engines like Google and Bing, have now also been adopted on many shopping platforms (akin to Amazon, eBay, etc.) to help merchants promote their products. For example, when users type "Nike shoes" in Amazon's search box, sponsored ads are displayed at the bottom of the search result page, along with the "organic" results (those generated by Amazon's search algorithm).

Consumer behavior vis-à-vis sponsored ads is likely to be quite different depending on whether the ads are on a shopping platform, or on a page of general search results. Ads on shopping platforms contain very specific information – including price, shipping parameters, delivery time, new or used condition, etc. – and allow consumers to selectively click on those ads that meet their preferences. Users only need to click on links for options they are seriously considering, which can garner substantial traffic even for ads at lower positions. As such, this leads to a so-called "position paradox" (Jerath et al. (2011)) – an empirical finding that ads in higher positions may attract fewer clicks than ads in

*Department of Economics, USC; yuwei.hsieh@usc.edu
†HSS, Caltech; mshum@caltech.edu
‡Department of Marketing, USC Marshall School of Business; shayang@marshall.usc.edu

Acknowledgement: We thank Arie Beresteanu, Baiyu Dong, Hashem Pesaran, Joris Pinkse, Sergio Montero, Roger Moon, Geert Ridder and Guofu Tan for their helpful discussions. Thanks also go to the seminar participants at Colorado-Boulder, Penn State, Rice, Rochester, USC, UC-Riverside, UCLA, Yale-SOM, 2013 Annual Conference of Taiwan Econometric Society, 2014 California Econometrics Conference and 2014 North American Summer Meeting of the Econometric Society.

lower positions. On the other hand, with ads on search engines, which typically do not contain price or specific product information, a "top-down" search – in which users click firstly on the top ad and proceed downwards – is more likely. Such behavior leads to a decreasing trend in click-through rates (CTR) across ad positions.

Auctions are the dominant mechanism whereby sponsored ad positions are sold on the Internet, with competing merchants bidding for the available positions. Many existing models of sponsored search auctions simplify the characterization of equilibrium bidding by directly assuming that click-through rates for sponsored ads strictly decrease with ad position.[1] These studies have mainly focused on sponsored ad auctions run by search engines, for which this assumption may be appropriate, as discussed above; however, it may be unrealistic for sponsored search advertising on shopping platforms, which we consider here.

In this paper, we propose a general structural model of sponsored search auctions. We estimate the model using a unique dataset of sponsored search auctions at a large Chinese shopping platform. Since our study company is a shopping platform, our model specification is geared to accommodates the "position paradox". Methodologically, we develop a novel econometric setup which exploits the equivalence between the position auctions and the classical assignment game of Shapley-Shubik (1972).

Our study website (anonymously dubbed "WebsiteX") is one of the largest shopping platforms in China, and one of the largest websites in the world by traffic. For a number of years, the platform implemented a standard Generalized Second-Price (GSP) auction to determine ad positions and prices for each keyword. We focus on digital cameras, which is a very popular product category on this platform.

Our estimation results show that merchants of different qualities adopt different bidding strategies: high quality merchants bid more aggressively for informative keywords, while low quality merchants are more likely to be sorted to the top positions for vague keywords. One explanation is that, for digital camera-related keywords, informative keywords (including specific camera model numbers) are likely to be queried by serious buyers, and high quality merchants are more experienced and have learned that clicks from these buyers are more likely to lead to sales. On the other hand, users who query vague keywords (including brand names and promotional terms) may not be ready to buy, and so experienced high quality merchants are less interested in these auctions. Thus we find evidence of both horizontal and vertical differentiation in these auctions;

[1] Edelman et al. (2007); Varian (2007); Athey and Nekipelov (forthcoming). See Börgers et al. (2013) for a general discussion of equilibria in these auctions without imposing these restrictions on preferences.

auction outcomes can be non-assortative, in that higher quality merchants do not obtain higher positions. Such results would not be consistent with the assumption of decreasing click-through rates typically assumed in the literature (as discussed above).

To score or not to score? Using our estimates, we proceed to address an important auction design question, as reflected in our paper's title. After our sample period, our study website revised the standard GSP by incorporating *popularity scoring*, a practice originally developed by Google. In the score-based GSP auction, the ad position obtained by a bidder will depend on her bid times a popularity/quality score, which typically reflects the popularity of a bidder or her ads (as measured by past sales or click volumes). By introducing scoring, our study website intended to reward "popular" merchants – those whose previous ads generated many clicks, and sold many products in the past – with price discounts. Introducing scoring to the GSP auction essentially penalizes smaller and less prominent merchants and rewards larger merchants (i.e. a less popular bidder pays a higher price than a more popular bidder for the same position, holding everything else constant).[2]

Indeed, soon after our platform implemented such score-based GSP auction, a large protest organized by small merchants broke out. They blamed the new scoring rule for making them uncompetitive against big merchants. One small merchant, *Longzhi521* complained that "small merchants on WebsiteX are unable to do business anymore, now WebsiteX is just for the big merchants, its really unfair!" *ILoveRainyDays* lamented that he "recently quit my job to sell on WebsiteX fulltime, but after the change it is impossible, as a small merchant, to get my ads noticed." *HangzhouWuMing*, quoting an ancient proverb, warns ominously that WebsiteX's new rules which discourage small merchants may have bad long-run consequences: "the waters (large merchants) which carry the boat can also capsize it."[3]

Our simulations show that the price trend becomes steeper under scoring, with much higher prices obtained for the top position but lower prices for the other positions. This suggests that while the intention of scoring was to reward popular merchants with price discounts, these discounts were undone (in part) by more aggressive bidding by these merchants. On the one hand, this confirms the fears of the small merchants quoted above that they have to pay higher prices to get top positions. However, since CTRs in

[2]The use of bid scoring in sponsored search auctions has similar implications as bid preference policies in procurement auctions (see McAfee and MacMillan (1989), or Krasnokutskaya and Seim (2011).) Generally, a scoring rule which favors stronger (high quality) contractors may lead to higher quality work, while favoring weaker bidders enhances competition and lowers procurement costs.

[3]News release from sina.com.cn (date: Sept. 7, 2010)

our model are not necessarily decreasing in ad position (and indeed are quite random), the effect on the platform's total revenue is ambiguous. Our simulations show that, indeed, there is only a very modest change in total revenue from introducing popularity scoring. Although other long term benefits from scoring – such as encouraging merchants to improve ad quality, thus generating more clicks and enhancing the attractive of the selling platform – are beyond the scope of our analysis, it is remarkable that even the platform's short term profits may not be much affected by rewarding popular merchants with price discounts.

Existing literature. Although sponsored advertising auctions have received great attention in the theoretical literature (beginning with Edelman et al. (2007) and Varian (2007)), empirical research is still sparse. Börgers et al. (2013) utilize a revealed-preference approach to test whether bids in Yahoo search auctions satisfied the Nash Equilibrium inequalities for the sponsored search auction model. Yang et al. (2014) studied how competition affects sponsored search advertisers' bidding behavior, and they modeled bids as equilibrium outcomes using the specification in Edelman et al. (2007). Athey and Nekipelov (forthcoming) propose and estimate a structural model tailored to specific features of sponsored search auctions run by US search engines (such as Google or Microsoft). Specifically, they estimate a model characterized by score and entry uncertainty ("SEU"), in which bidders face uncertainty when choosing their bids, due to randomness in a bidder's quality score over time, as well as in the set of competitors bidding in the auction at any time.

Our paper is different from the aforementioned studies in several major ways. First, during our sample period, our study website ran un-scored auctions, which eliminates an important source of uncertainty in Athey and Nekipelov's SEU model. Indeed, the study website started scoring auctions only after our data sample, and the main question addressed here is how this affected bidder allocations and platform revenue. Second, we allow bidders to have preferences for positions which are not multiplicative in bidder- and position-specific effects. This is in line with empirical evidence (Jeziorski and Segal (forthcoming), Jeziorski and Moorthy (2014), Goldman and Rao (2014)) which contradicts the multiplicative hypothesis. Finally, our dataset, compared to Athey and Nekipelov's dataset, contains substantial cross-sectional variation (a large number of keyword queries), but no time series variation (but click volume and per-click prices averaged over a one-month period). We model the aggregated outcomes as arising from a static bidding model, thus abstracting away from changes in bidding behavior over the aggregation period. This is not an unreasonable approximation for WebsiteX's auctions because, during our sample period, all bids were submitted manually and, accordingly, participating merchants revised their keyword choices and bids relatively infrequently[4]

Specifically, in this paper, we estimate a model of sponsored search auctions which is closely related to the Shapley-Shubik assignment game.[5] This connection to the Shapley-Shubik model also ties our paper to the recent empirical literature using two-sided matching models (e.g. Choo and Siow (2006), Fox (2013), Galichon and Salanie (2012), Graham (2011)). One important difference vis-à-vis these papers, is that we explicitly model and estimate the price formation process in the sponsored search auctions we study, while these other papers focus on explaining the observed allocations.

In the following section we provide background information about the platform and present the structural model for ad positions and prices from the generalized second price auction. We derive the crucial link between the auction model and two-sided matching models, which we exploit in estimation. In section 3, we derive the "Metropolis-Hastings within Gibbs" Bayesian algorithm to estimate the structural parameters. In Sections 4 and 5, we describe the dataset and present the empirical findings. In section 6 we conduct the counterfactual analysis to address the "to score or not to score" question. Section 7 concludes.

2 A Structural Model of Sponsored Search Auctions

2.1 Background: sponsored search auctions at "WebsiteX"

WebsiteX is one of the largest online marketplaces in China and, hence, one of the most prominent websites in the world by traffic. Given the high costs and regulatory and bureaucratic hurdles associated with opening brick-and-mortar businesses in China, many small merchants market and sell their wares mainly using internet shopping platforms like our study website. As a marketplace, WebsiteX has no direct American counterpart, but shares features of both eBay and Craigslist. Unlike eBay, goods on WebsiteX are not sold via auction, but rather by merchants posting prices for their products. WebsiteX provides a platform whereby buyers can make secure money transactions to merchants. Sponsored search results typically appear as "tiles" on the right-hand side and bottom margins of each search page (see Figure 3 for an example).

[4]This differs from advertisers on search engines like Google or Yahoo, who use automated bidding mechanism which make it easy to revise bids or keywords.

[5]As such, our work echoes Demange, Gale and Sotomayor (1986) and Fox and Bajari's (2013) study of the FCC spectrum auctions, both of which apply an assignment game approach to multi-unit auctions. Other non-auction settings in which the assignment game has been applied include marriage markets (Becker (1973)), mergers (Akkus et al. (2013)), and hedonic pricing models (Chiappori et al. (2009)).

Since WebsiteX is a marketplace, the content and role of its sponsored ads differ substantially from the ads appearing on search engine result pages (like Google or Yahoo). As Figure 3 shows, WebsiteX's sponsored ads (as well as the non-sponsored "organic" search results) typically contain a picture of the product, price, merchant name and information, shipping details, and product specifications; in contrast, such details are typically absent from Google's sponsored search results, which contain only the URL along with some brief slogans.

Since WebsiteX is a shopping platform, most of its users have a serious intent of purchasing, and are using WebsiteX to find prices and product specifications to suit their needs.[6] Consequently, as we mentioned earlier, the top positions may not always receive more clicks (and indeed, they do not, as we will show below). For this reason, the common assumption in the existing sponsored search auction models (Edelman et al. (2007) and Varian (2007), Athey and Nekipelov (forthcoming)) that the surplus matrix is supermodular (being the product of a vector of bidder-specific constants and a nonincreasing vector of position-specific click-through rates) seems inadequate for WebsiteX. Accordingly, in our setup, we will allow the surplus to vary arbitrarily among merchants and across ad positions, and also allow click volumes to be non-monotonic in ad position.

2.2 Generalized Second-Price Auction (GSPA)

Next we describe the generalized second-price auction framework. There are N available positions and $M \geq N+1$ potential bidders (synonymously, merchants) for a generic keyword auction. If bidder i obtains the j-th position, he obtains valuation (or surplus) V_{ij}, for all bidders i and positions j. In what follows, without loss of generality we will index the positions from top to bottom by $i = 1, \ldots, N$, and similarly we will also label the $N+1$ highest bidders by $i = 1, \ldots, N+1$.

The rules of the generalized second-price auction are as follows: for N positions, the N-highest bidders will be winners, with the i-th ($1 \leq i \leq N$) highest bidder obtaining position i at the per-click price equal to the $i+1$-th bidder's bid. Using the terminology in Varian (2007) and Börgers et al. (2013), we focus on the so-called "symmetric" Nash equilibria in this complete-information bidding game.[7] The equilibrium conditions

[6]It has been discussed for a while that many retailers fear becoming Amazon's Showroom. This phenomena is even more radical in China as running a retail store would incur more tax and fee liability. By contrast, running a online store can sidestep these hidden costs. Consequently, the price gap between online and retail stores in China is even larger than the USA counterpart.

[7]These are closely-related to the "locally envy-free" equilibria in Edelman et al. (2007). These equilibria are convenient to analyze, and easy to compute via linear programming; as noted in Börgers

satisfied by a bid vector $(b_1, \ldots, b_M)$ for $M \geq N+1$ are

$$V_{ii} - \alpha_i b_{i+1} \geq V_{ij} - \alpha_j b_{j+1}, \ \forall i, j \tag{1}$$

where α_j denotes the click volume for position j. This inequality ensures that, at the equilibrium, the bidder who obtains positions i (who obtains valuation V_{ii} and makes a payment equal to the click volume α_i times b_{i+1}, the per-click price submitted by the bidder in position $i+1$), does not wish to deviate to position j, for which the surplus would be equal to the RHS of the inequality.[8] Making the substitution $p_i = b_{i+1}$ (that is, the per-click price for the i-th position equals the bid in the $i+1$-th position), we have

$$V_{ii} - \alpha_i p_i \geq V_{ij} - \alpha_j p_j, \ \forall i, j. \tag{2}$$

2.3 GSPA as Two-sided Matching

Our estimation approach relies critically on the reinterpretation of the GSPA as an assignment game of Shapley and Shubik (1972). To draw this connection, we consider a "matching" problem where bidders are matched to positions. We denote by

$$u_i \equiv V_{ii} - \alpha_i p_i \tag{3}$$

the equilibrium payoff of bidder i, and

$$t_j \equiv \alpha_j p_j \tag{4}$$

the equilibrium payoff for the platform from the j-th position. Now rewriting the equilibrium inequalities (2) above, we get

$$u_i + t_j \geq V_{ij}$$

with equality (by construction) iff $i = j$. These are the well-known "no-blocking" conditions from the matching problem with transfers (cf. Roth and Sotomayor (1990; chap. 8)). To see why, consider a bidder i and position j, and assume that $u_i + t_j < V_{ij}$ or, equivalently, $u_i < V_{ij} - t_j$. In this case, since bidder i's payoff u_i is lower than the net

et al. (2013), no such characterization is available for the asymmetric Nash equilibria. Nevertheless, it is possible to adapt our estimator to the case of asymmetric Nash equilibria; we leave the details for future research.

[8] In contrast, in asymmetric Nash equilibria, Eq. (1) holds only for $j > i$, but is $V_{ii} - \alpha_i b_{i+1} \geq V_{ij} - \alpha_i b_j$ for $j < i$. This recognizes an asymmetry that in order to switch to a lower position, bidder i only needs to beat the price of that position, but to switch to a higher position, bidder i must beat the *bid* of the winner of that position.

surplus that she would obtain from deviating to position j ($=V_{ij} - t_j$), she would not agree to the given allocation, and the equilibrium would break down; since the pair of bidder i and position j would "block" the proposed allocation, the payoffs (u_i, t_j) cannot support this allocation in equilibrium.

Moreover, introducing the binary indicators $\mu(i,j) = 1$ if bidder i obtains position j, and zero otherwise,[9] and summing up across all bidders and positions using Eqs. (3,4), we have

$$\sum_{i,j} u_i + t_j = \sum_{i,j} [\mu(i,j)V_{ij} - \alpha_i p_i + \alpha_i p_i] = \sum_{i,j} \mu(i,j)V_{ij}. \tag{5}$$

This "feasibility" condition is the link between the sponsored-search auction model and the assignment game, as it is implied by the duality theorem of linear programming for that latter model. In the remainder of this section, we flesh out this connection.

2.4 Optimal allocation in GSPA: matching Positions to Bidders

If bidder i obtains the j-th position, the valuation function is given by

$$V_{ij} = \delta(X_i, Z^j; \beta) + \epsilon_{ij} \tag{4}$$

where X_i is the vector of bidder i's characteristics and Z^j is the j-th position-specific characteristics. The $\delta(\cdot)$ is the deterministic component of the valuation function parametrized by a finite dimensional parameter β, and ϵ_{ij} is the unobservable match-and-auction-specific valuation.[10] We will refer to $\mathbf{V}$ the valuation matrix, where the (i,j) entry of $\mathbf{V}$ is V_{ij}. We further assume the unobserved valuation shocks satisfy the following assumption

Assumption 1. *ϵ_{ij} is a continuous random variable with mean zero and variance $\sigma^2 < \infty$ with unbounded support on $\mathcal{R}$. ϵ_{ij} is mutually independent across index i, j.*

An allocation (or matching) μ, is a binary matrix indicating which bidder acquires which position. We use 1 to indicate the assignment of positions to bidders, and zero otherwise. For example, if bidder 1 gets the second position, and bidder 2 gets the first position, the allocation μ is

[9] Using our indexing convention that bidder i is allocated position i, we have $\mu(i,j) = 1$ for $i = j$, and zero otherwise.

[10] This is a complete information game in which ϵ_{ij} is assumed to be observable to all players within the game but unobservable to the researchers.

	1	2
bidder1	0	1
bidder2	1	0

In the one-to-one assignment game, the row sum and column sum are equal to 1. We will write $\mu(i,j)$ as the (i,j) entry of μ. In a N-by-N assignment game, there are $N!$ allocations. We will refer to Ω as the set of all possible allocations and μ_ω as an generic element of Ω, where $\omega = 1, 2, \ldots, N!$. The total surplus under allocation μ_ω is denoted by $S_{\mu_\omega}(\mathbf{V})$

$$S_{\mu_\omega}(\mathbf{V}) = \textstyle\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\delta(X_i, Z^j;\beta) + \epsilon_{ij}\right]\cdot \mu_\omega(i,j) \equiv \Delta_{\mu_\omega} + \Xi_{\mu_\omega}, \text{ where}$$

$$\Delta_{\mu_\omega} = \textstyle\sum_{i=1}^{N}\sum_{j=1}^{N}\delta(X_i, Z^j;\beta)\cdot\mu_\omega(i,j) \text{ and}$$

$$\Xi_{\mu_\omega} = \textstyle\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ij}\cdot\mu_\omega(i,j)$$

Based on $\mathbf{V}$, the total surplus $S(\mathbf{V})$ for each allocation $\mu_\omega \in \Omega$ can be calculated. Shapley and Shubik (1972) consider the assignment problem of finding the optimal one-to-one allocation that maximizes the total (social) surplus, as well as the stable price systems to support/decentralize the optimal allocation. The social planner's problem, can be formulated as the following linear program (which we denote (P)):

$$\begin{aligned} &\max_{\mu(i,j)} \textstyle\sum_{i,j} V_{ij}\mu(i,j) \\ \text{s.t.} \quad & \textstyle\sum_i \mu(i,j) = 1, \forall i \\ & \textstyle\sum_j \mu(i,j) = 1, \forall j \\ & \mu(i,j) \in \{0,1\}, \forall (i,j) \end{aligned} \tag{P}$$

Since there are only finitely many allocations in Ω, a solution always exists in such maximization problem (Roth and Sotomayor, 1990).[11]

Lemma 1. *Under assumption 1, the optimal allocation (the solution to (P))is unique almost surely.*

[11] This result contrasts with two-sided matching models without transfers, in which the multiplicity of stable allocations becomes a major concern; e.g., Boyd et al. (2006), Logan et al. (2008), Menzel (2011), Hsieh (2011), and Echenique, et al. (2013), among others.

Proof. Take any two allocations $\mu_l \neq \mu_q$, $(\mu_l, \mu_q) \in \Omega$. The event $\{S_{\mu_l}(\mathbf{V}) = S_{\mu_q}(\mathbf{V})\} = \{\Xi_{\mu_l} - \Xi_{\mu_q} = \Delta_{\mu_q} - \Delta_{\mu_l}\}$ is a set of measure zero under assumption 1. It immediately follows that the ordering of $\{S_{\mu_\omega}(\mathbf{V})\}_{\omega=1,\dots,N!}$ is strict almost surely. ■

2.5 Equilibrium prices in the GSPA

The linear program (P) yields the optimal allocation of assigning positions to bidders. Since our goal is to analyze not only allocations, but also prices, we will turn to the dual linear program, which yields the prices supporting the optimal allocation in equilibrium.

Recall that we denote by u_i the equilibrium payoff of bidder i, and $t_j (\equiv \alpha_j p_j)$ the equilibrium payoff of the j-th position. By the duality theorem of linear programming, the dual problem of (P) is given by

$$
\begin{aligned}
& \min \textstyle\sum_{i=1}^{N} u_i + \sum_{j=1}^{N} t_j \\
\text{s.t.} \quad & u_i \geq 0, t_j \geq 0, \forall i, j \\
& u_i + t_j \geq V_{ij}, \forall i, j
\end{aligned}
\tag{DP}
$$

The first set of constraints, $u_i \geq 0, t_j \geq 0$, are individual rationality condition: both bidders and search engine should have non-negative profit. The second set of constraints, corresponds to the incentive compatibility, or no-blocking pair conditions. The set of (u_i, t_j) that solves (DP) is denoted the set of stable matchings (see Roth and Sotomayor, 1990). Shapley and Shubik (1972) further show that $u_i + t_j = V_{ij}$ iff $\mu_{ij} = 1$. By summing this across (i, j), we obtain Eq. (5) above, which provides the link between the GSPA and the assignment game, as we alluded to before.

In general, it is well-known that there exist multiple transfers $\mathbf{t} = (t_1, \dots, t_N)$ that solve (DP), and there exist bidder-optimal $\underline{\mathbf{t}}$ and platform-optimal $\overline{\mathbf{t}}$.[12] This multiplicity in equilibrium prices raises issues for estimation, as we will discuss below. By contrast, given Lemma 1, the corresponding optimal matching μ that solves (P) is unique almost surely.

[12] The literature on multi-item auction raises a design problem of how specific auction mechanism may select particular stable matchings. For example, Demange, Gale and Sotomayor (1986) propose an auction mechanism in which the bidder-optimal stable matching is the equilibrium outcome.

Within the set of stable matchings $\mathbf{t}$, the generalized second-price auction mechanism selects a subset in which the transfers are monotonically decreasing in ad position:

$$\{\mathbf{t}|\mathbf{t} \text{ solves (DP) \& } p_1 > p_2 > \dots p_N; t_i = \alpha_i p_i\}. \tag{6}$$

We will refer to this as the set of "stable per-click prices."

Example: (Non-)existence of equilibrium in GSPA. While Shapley and Shubik (1972) proved that the set of stable matchings is always nonempty for arbitrary V_{ij}, this is no longer true once the additional monotonicity condition (6) is imposed. Hence, without extra assumptions on V_{ij} or the click volumes, the generalized second-price mechanism may not necessarily guarantee the existence of a symmetric Nash equilibrium (or equivalently, competitive price system). Consider an example: the valuation matrix is given by

	1	2
bidder1	3	4
bidder2	1	3

and $\alpha_1 = \alpha_2 = 1$. The set of stable prices for the Shapley-Shubik assignment game is shaded in blue in Figure 4. We see that this region lies completely above the 45-degree line, where $p_2 > p_1$ (the price for the second position exceeds that of the top position); the GSPA, on the other hand, requires $p_1 > p_2$. The problem in this example is that both players value the second position more than the first one, making it impossible to sell the first position with a higher market price. In our empirical model, our assumptions on valuations (Assumption 1) allow cases similar to this example to arise with positive probability. For estimation, we wish to restrict attention only to the set of V_{ij} that are consistent with equilibrium in the GSPA, which introduces a complicated truncation problem (similar to that arising in Hong and Shum's (2003) econometric study of asymmetric ascending auctions). ■

3 Data

We obtained a one-month (in 2010) sponsored-link auction dataset from WebsiteX, which is the largest online marketplace in China. The dataset includes aggregate information on 487 keywords of digital camera/camcorder and related accessories. The number of ad positions ranges from 5 to 9. According to insiders at WebsiteX, merchants there usually review their keyword lists and make purchase decisions infrequently. As a result, the actual auction environment is not as complicated as Google or Yahoo!, in which bidders often apply some automatic bidding algorithm to dynamically manage their positions and per-click prices. During the study period, WebsiteX applies standard GSPA without score weighting. The bidding environment of WebsiteX is therefore more closely related to the static model described in Edelman et al. (2007) and Varian (2007).

For each keyword string, we observe several sets of variables. First, for each winning bidder (merchant), we observe their average (over one-month) per-click price, along with with the aggregate click volume over this month. We fit these data into our static auction framework by sorting the bidders by their average per-click price, and then placing the bidder paying the highest average price in the the top position, the bidder paying the second highest price in the second position, etc. Second, we observe very little about individual merchants except "quality ratings" which WebsiteX creates (using their proprietary algorithm) based on buyer feedback and merchants' sales volumes. Specifically, each bidder is rated on an increasing quality scale from 1 to 20, which is also broken into quality brackets: top (16-20), high (11-15), medium (6-10) and low (1-5). In our dataset, there are no top merchants, so in our analysis we will distinguish between three quality types: low, medium, and high. Third, we construct dummy variables to describe characteristics of each keyword string: 1. *Brand*, whether the keyword includes a brand name such as Nikon or Canon; 2. *Specific*, whether the keyword includes a specific model/series number, such as D300s or 500D; 3. *Promotional*, whether the keyword includes promotional terms such as "cheap" or "sale". Summary statistics for these variables are given in Table 1.

A first look at the data. We begin with some reduced form statistics and tabulations from the data to motivate details of our model specification. From the perspective of search engines it is important to know how advertisers' bidding strategies are related to their quality rankings and keyword characteristics. Do certain types of bidders bid more aggressively for certain type of keywords? What is the sorting pattern between quality and rank? Table 2 contains a contingency table summarizing bidders' quality versus the ad positions they obtained. Overall, the evidence for assortative matching in bidder quality across ad positions is mixed. On the one hand, from the top panel of the table, we see that about one-third of the high-quality bidders (=7.2% of total bidders) get the top ad position in the auctions in which they won a position, and this percentage falls across positions. However, for medium quality bidders, we see that most of them are sorted to position 5 (16.1%) followed by 3 (15.6%). The results are qualitatively stable after doing the analysis separately for different types of keyword queries, as is shown in the remaining panels in Table 2.

Although it seems that the allocation patterns do not vary with keyword characteristics, the per-click prices and click volumes do move dramatically. The boxplot of log click volumes[13] and per-click prices are depicted in Figure (1) and (2) respectively. In Figure (1) we compare the boxplot of log click volumes across the top 5 positions, conditional on different dummies. First, clearly the assumption of decreasing click volume with rank is violated since the middle of the boxplot does not decrease when rank decreases. In fact, only 5 out of 487 keywords have strictly decreasing click volumes. Second, we observe that keyword characteristics shift the click volume distributions. Keywords containing specific model number usually receive more click volumes across all ranks of position (top-left of Figure 1). Keywords containing brand name slightly decrease the click volumes (top-right of Figure 1). It is also interesting to note that keywords containing promotional terms in fact generate smaller click volumes (bottom-left of Figure 1).

Lastly, we turn our attention to the distribution of per-click prices (Figure 2). We find that per-click prices are generally higher (and have more extreme outliers) for keywords containing specific model number (top-left of Figure 2). Adding brand name on average does not change per-click price, but it does create more outliers (top-right of Figure 2). Adding promotional terms does not change per-click price (bottom-left of Figure 2).

4 Estimation

Next we consider the estimation of the sponsored search auction using data on the observed allocation as well as the per-click prices. We use a Bayesian approach to estimate this model. There are a large number of latent variables in this model, relative to the observed variables: in each auction, for the N-dimensional vector of observed bids, there are N^2 corresponding unobservabled: namely, the full set of valuations $\{V_{ij}\}_{i,j=1\ldots N}$ that each merchant has for each position. An important virtue of the Bayesian approach is the use of "data augmentation" (Tanner and Wong (1987)), whereby these latent variables are treated as unknown parameters, and jointly inferred in the estimation procedure.[14]

However, the multiplicity of stable per-click prices, mentioned earlier, raises difficulties with a Bayesian estimation approach, as it leads to indeterminacy of the likelihood function for the prices. In this paper, we complete the model by assuming particular parametric equilibrium selection rules, as we describe below.[15]

[13] As there are some keywords that receive extremely large amount of click volumes, for graphical presentation purpose we depict the boxplots in the log scale.

[14] In contrast, in a frequentist framework, these latent valuations must be integrated out of the estimating equations; this is difficult using typical numerical integration methods (quadrature, simulation) due to the large dimensionality of the integration.

[15] In Appendix B, we consider an alternative approach, using bounds based upon the struture of the equilibrium price set, which is agnostic as to the equilibrium selection rule. Such an approach turns out

4.1 The likelihood of auction allocations and prices

The central component of our Bayesian estimation procedure is the derivation of the joint likelihood function for the auction outcomes (the matching of bidders to positions and the corresponding prices). This is the focus of this section.

As the set of bidders and the number of positions vary from auction to auction, defining the random vector $\mathbf{p}$ and random matrix μ requires additional care to avoid a labeling problem. Following the previous section we will sort the element of $\mathbf{p}$, p_i, in decreasing order, and therefore X_i corresponds to the bidder who pay p_i. Under this indexing system, the allocation μ will be the identity matrix. Suppose the econometrician observes T independent keyword auctions indexed by $t = 1, 2, \ldots, T$. In each keyword auction t one observes two sets of dependent variables, the allocation μ_t and the vector of per-click prices $\mathbf{p}_t = (p_{1t}, \ldots, p_{N_t t})$. The exogenous variables are $\mathbf{X}_t = (X_{1t}, \ldots, X_{N_t t}), \mathbf{Z}_t = (Z^{1t}, \ldots, Z^{N_t t}), \alpha_t = (\alpha_{1t}, \ldots, \alpha_{N_t t})$. Let $\mathbf{V}$ denote the collection of latent valuation matrices for all keyword auctions $(\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_T)$, and similarly we define $\mu = (\mu_1, \mu_2, \ldots, \mu_T)$, $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_T)$ and $(\mathbf{X}, \mathbf{Z}, \alpha)$. The posterior is given by

$$f(\theta, \mathbf{V}|\mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha) \propto \mathcal{L}(\mu, \mathbf{p}|\theta, \mathbf{V}; \mathbf{X}, \mathbf{Z}, \alpha) p_0(\theta, \mathbf{V}|\mathbf{X}, \mathbf{Z})$$

The specification of the priors is given in the Appendix. Here we focus on the form of the likelihood, given by

$$\mathcal{L}(\mu, \mathbf{p}|\theta, \mathbf{V}; \mathbf{X}, \mathbf{Z}, \alpha) = \Pi_{t=1}^{T} \mathcal{L}(\mu_t, \mathbf{p_t}|\theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t, \alpha_t). \tag{7}$$

Importantly, the likelihood above differs from the likelihood which would be optimized in a frequentist setting (ie. MLE), because the unobserved valuations $\mathbf{V}$ are treated as conditioning variables. This is due to "data augmentation", as discussed above. In the frequentist likelihood, in contrast, $\mathbf{V}$ cannot be conditioned on (since it is unobserved), and would need to be integrated out.[16] The likelihood (7) can be further decomposed into

$$\mathcal{L}(\mu_t, \mathbf{p}_t|\theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) = \mathcal{L}_1(\mathbf{p}_t|\mu_t, \theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)\mathcal{L}_2(\mu_t|\theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t) \tag{7}$$

the two terms of which are the conditional likelihoods of per-click prices (given the allocation and valuations), and the allocation (given valuations). The explicit forms

to be quite computationally challenging, compared to the approach we use in this paper.

[16]Specifically, the frequentist likelihood is $\mathcal{L}(\mu, \mathbf{p}|\theta, \mathbf{V}; \mathbf{X}, \mathbf{Z}, \alpha) = \int \cdots \int \mathcal{L}(\mu, \mathbf{p}|\theta, \mathbf{V}; \mathbf{X}, \mathbf{Z}, \alpha) dG(\mathbf{V}|\mathbf{X}, \mathbf{Z})$ which is difficult to evaluate due to the large dimensionality of $\mathbf{V}$.

of $\mathcal{L}_1$ and $\mathcal{L}_2$ are derived from, respectively, the dual and primal LP problems for the assignment game. We consider each component in turn.

The component $\mathcal{L}_1$. For expositional simplicity we will drop the auction index t for now. Conditional on μ, the set of stable per-click prices is a convex polyhedron defined by a set of linear inequalities derived from the dual LP problem of the assignment game, which comprise (i) the "no-blocking pair" inequalities (Eq. (1)); (ii) the individual rationality constraints, that $V_{ii} - t_i \geq 0$; and (iii) the monotonicity and non-negativity conditions on per-click prices: $p_1 > p_2 > \ldots > p_N \geq 0$. The details of these linear inequalities are given in the appendix; we let $\mathbf{P}(\mathbf{V}, \alpha)$ denote the polyhedron of equilibrium per-click prices in the sponsored-search auction, given a set of bidder valuations and position-specific click volumes $(\mathbf{V}, \alpha)$. As $\mathbf{P}$ depends on the realization of unobserved ϵ_{ij}, the set of stable per-click prices itself is a random closed convex polyhedron.

Correspondingly, the conditional likelihood for the per-click prices is a distribution supported on a convex polyhedron: $\mathcal{L}_1(\mathbf{p}_t|\mu_t, \theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) = \mathcal{L}_1(\mathbf{p}_t|\mu_t, \mathbf{P}(\mathbf{V_t}, \alpha_t), \lambda)$, where λ, a probability measure over $\mathbf{P}$, denotes an equilibrium selection rule whereby $\mathbf{p}_t$ is selected from $\mathbf{P}(\mathbf{V}_t, \alpha_t)$. To proceed, we consider parametric equilibrium selection rules to complete the likelihood specification, following Bajari et al. (2010).[17] Specifying such a distribution in our case is nontrivial, as the support $\mathbf{P}(\mathbf{V}, \alpha)$ depends on the latent variables. We consider two parsimonious specifications. First, we assume that all prices satisfying the no-blocking conditions are drawn with uniform probability, which is equal to the reciprocal of the volume of the polyhedron of values of prices which satisfy the linear inequalities $\{\mathbf{p} : A(\alpha_t)\mathbf{p} \leq \mathbf{b}(\mathbf{V}_t)\}$:

$$\mathcal{L}_1(\mathbf{p}_t|\mu_t, \mathbf{P}(\mathbf{V_t}, \alpha_t), \lambda) = \mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t))\frac{1}{vol(\mathbf{P}(\mathbf{V}_t, \alpha_t))}$$

In what follows we will call this the "Uniform" specification of equilibrium selection.

Second, we assume that each component of $\mathbf{p}_t$, p_{it}, is independently draw from beta distribution defined on $[\underline{p}_{it}, \bar{p}_{it}]$ truncated to $A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t)$, where $\Pi_{i=1}^{N_t}[\underline{p}_{it}, \bar{p}_{it}]$ is the smallest bounding box of $\mathbf{P}(\mathbf{V}_t, \alpha_t)$:

[17] In the existing literature, another common approach for dealing with multiple equilibria is to identify and compute bounds on the structural parameters of the model, thus avoiding the explicit specification of the equilibrium selection rule λ; e.g., Ciliberto and Tamer (2009); Beresteanu et al. (2011); Galichon and Henry (2011) among others. Applying this approach in our context is briefly discussed in Appendix B.

$$\mathcal{L}_1(\mathbf{p}_t|\mu_t, \mathbf{P}(\mathbf{V_t}, \alpha_t), \lambda) \propto \mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t))\Pi_{i=1}^{N_t}\Big\{\frac{(p_{it} - \underline{p}_{it})^{a-1}(\bar{p}_{it} - p_{it})^{b-1}}{B(a,b)(\bar{p}_{it} - \underline{p}_{it})^{a+b-1}}\Big\},$$

where $B(a,b)$ is the Beta function. Under this specification (which we will denote the "Beta" specification), the shape parameters (a,b) in the beta distribution will be additional parameters to be estimated.

The component $\mathcal{L}_2$. Conditional on $\mathbf{V}_t$, the likelihood $\mathcal{L}_2$ is binary (0-1) valued:

$$\mathcal{L}_2(\mu_t|\theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t) = \mathcal{L}_2(\mu_t|\mathbf{V}_t) = \begin{cases} 1 & \text{if } \mathbf{V}_t \text{ rationalizes } \mu_t \\ 0 & \text{otherwise} \end{cases}$$

Moreover, at the observed $(\mathbf{p}_t, \mu_t)$, all $\mathbf{V}_t$ which lead to a nonzero value for $\mathcal{L}_1$ automatically rationalize μ_t, and lead to a nonzero value of $\mathcal{L}_2$. As a result, $\mathcal{L}_2$ is redundant and we can simplify

$$\mathcal{L}(\mu_t, \mathbf{p}_t|\theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) \propto \mathcal{L}_1(\mathbf{p}_t|\mu_t, \theta, \mathbf{V}_t; \mathbf{X}_t, \mathbf{Z}_t, \alpha_t).$$

4.2 Estimation algorithm

We estimate the structural parameters via a Metropolis-Hastings within Gibbs sampler (Robert and Casella (2005)). We summarize the procedure here, but give complete details in the appendix. The "outer loop" is a Gibbs sampler which loops sequentially over three conditional densities: the conditional density of θ, the conditional density related to equilibrium selection of multiple equilibrium prices, and the conditional density of the latent valuations V_{ij} (the augmented component). Since these are difficult to sample from directly, within each Giibs step we use a Metropolis-Hastings approach to obtain draws from these three conditional densities. The main idea closely follows that of Albert and Chib (1994), and Logan et al. (2008).[18]

4.3 Specification details

We end this section with some details of our model specification. First, in our model we allow the click volumes (the α's) to vary non-monotonically across positions, in line with the evidence presented earlier in Figure (1). However, in our data, there are instances when the top position may receive zero clicks (cf. Table 1), but still attract bidders.

[18] Logan et al estimate a two-sided matching game *without* transfers. Their Gibbs sampler does not take into account multiple equilibria, and hence cannot be directly applied here.

Such an occurence would not be explicable in our model unless we allow for bidders to be uncertain about the click volume at the time they are submitting their bids, and hence bid based on expectations about the click volume (so that the realized click volume may not coincide with the ex ante expectation).[19] Hence, we allow for this by assuming that bidders' beliefs about α_{jt}, the click volume of the j-th position at auction t, satisfies the following shifted log-normal process (and are independent of their valuations V_{ij}):

$$\log(\alpha_{jt}+1) = \gamma_{j0} + \gamma_{j1}N_t + \gamma_{j2}\text{specific}_t + \gamma_{j3}\text{promotional}_t + \gamma_{j4}\text{brand}_t + \eta_{jt},$$

where N_t is the number of available positions and η_{it} follows $\mathcal{N}(0,\sigma^2)$.[20] The unknown parameters can then be estimated by OLS, and the expected click volume is given by

$$E[\alpha_{jt}] = \exp(\gamma_{j0} + \gamma_{j1}N_t + \gamma_{j2}\text{specific}_t + \gamma_{j3}\text{promotional}_t + \gamma_{j4}\text{branded}_t + \sigma_j^2/2) - 1.$$

We use the estimated $E[\alpha_{jt}]$ instead of α_{jt} in our estimation procedure.[21]

Second, for bidders' valuations V_{ij} we consider two specifications which contain interactions of bidder-specific and position-specific variables:

Model I

$$\begin{aligned}
V_{ijt} \; &= \beta_0 + \beta_1\text{specific}_t \\
&+\beta_2\text{promotional}_t + \beta_3\text{brand}_t \\
&+\beta_4\text{high}_i\tfrac{1}{j} \qquad\qquad +\beta_5\text{medium}_i\tfrac{1}{j} \\
&+\beta_6\text{high}_i\tfrac{1}{j}\times\text{specific}_t \qquad +\beta_7\text{medium}_i\tfrac{1}{j}\times\text{specific}_t \\
&+\beta_8\text{high}_i\tfrac{1}{j}\times\text{promotional}_t \qquad +\beta_9\text{medium}_i\tfrac{1}{j}\times\text{promotional}_t \\
&+\beta_{10}\text{high}_i\tfrac{1}{j}\times\text{branded}_t \qquad +\beta_{11}\text{medium}_i\tfrac{1}{j}\times\text{brand}_t + \sigma\epsilon_{ijt}
\end{aligned} \tag{11}$$

where ϵ_{ijt} is an i.i.d. standard normal random sequence. In this specification, we impose a hyperbolic decay $(1/j)$ in coefficients across positions. Note that while this specification imposes monotonicity in the coefficients across positions, the trend is allowed to be either decreasing or increasing (depending on the sign of the β's).

[19] Indeed, without this, zero click volume for a position would imply zero transfer for the top position, which would unreasonably imply zero transfer for all positions in the model.

[20] A similar specification is used in Yang et al. (2014), but here we do not impose that α_i decays from high to low positions.

[21] Even after regression smoothing, $E[\alpha_{jt}]$ is still nondecreasing in j in some cases. An important restriction of the assignment game framework is that we cannot allow the (expected) click volume to be bidder-specific. We will return to this point when discussing the counterfactual simulations below.

Model II

$$
\begin{aligned}
V_{ijt} &= \beta_0 + \beta_1 \text{specific}_t \\
&+\beta_2 \text{promotional}_t + \beta_3 \text{brand}_t \\
&+\beta_{1j} \text{high}_i \qquad\qquad\qquad +\beta_{2j} \text{medium}_i \\
&+\beta_{3j} \text{high}_i \times \text{specific}_t \qquad +\beta_{4j} \text{medium}_i \times \text{specific}_t \\
&+\beta_{5j} \text{high}_i \times \text{promotional}_t \qquad +\beta_{6j} \text{medium}_i \times \text{promotional}_t \\
&+\beta_{7j} \text{high}_i \times \text{brand}_t \qquad +\beta_{8j} \text{medium}_i \times \text{brand}_t + \sigma\epsilon_{ijt} \\
&\text{if } j \leq 5 \text{ and } = \sigma\epsilon_{ijt} \text{ if } j > 5
\end{aligned} \tag{12}
$$

This is a more flexible specification which allows each of the β coefficients to be position specific, and no longer imposes a monotonic trend as in Model I. (For tractability, we assume these coefficients to be zero for positions lower than 5, as click volumes and per-click prices for such lower positions are generally so small that they can be ignored.)[22]

Comparison: model specification in existing papers. Our specification details here contrast with those in Varian (2007), Edelman et al. (2007), and Athey and Nekipelov (forthcoming), who consider a multiplicative specification of bidder's valuations:

$$V_{ij} = v_i \alpha_j. \tag{5}$$

In tandem with the assumption that click volumes decrease in ad position ($\alpha_1 > \alpha_2 \cdots > \alpha_N$), this multiplicative specification (5) implies that every bidder has exactly the same preference ordering over positions: everyone prefers a higher position. Subsequently, it is easy to show that the optimal allocation is perfectly assortative matching: the k-th position is assigned to the bidder having the k-th highest valuation. In contrast, our specifications of valuations and click volume described earlier allow us to accommodate the richer patterns in allocations and click volumes which were evidenced in the data. ■

5 Estimation Results

Table (3) contains the results of the Model I specification. Since the results are similar using both the uniform and beta equilibrium selection rules, we will focus on the uniform rule (results in third column of Table 3) in the discussion. We find evidence for positive assortative matching between bidders' quality and ad positions. The posterior mean of coefficients of high and medium quality dummies are all positive (688.96 and 479.75,

[22]See the appendix for a discussion of parameter identification for these specifications of preferences.

respectively), and distinct from zero (larger than 3 times the standard deviation). Furthermore, the magnitude of the coefficient of high quality is larger than that of medium, implying that the top quality bidders have the highest valuations, all else equal.

However, there is substantial heterogeneity. We see that this positive assortative matching pattern is further amplified in auctions containing product-specific keywords (the interactions with "specific" are 539.45 and 516.41 for high and medium bidders, respectively). Since the positive assortative matching is extremely strong in product-specific keyword auctions, it is less likely that low quality bidder would win top positions here. In contrast, the interactions with "brand" are strongly negative (-341.05 for high-quality bidders, -395.04 for medium quality), and offset the positive coefficients described earlier. This finding suggests that high quality online (camera) merchants have a relatively low assessment of keyword strings containing brand names, so that lower quality merchants stand a higher chance of winning such auctions. (The interactions are likewise negative for the promotional keywords, but not as large in magnitude.)

These results imply horizontal differentiation across different types of keywords, as high quality merchants have relatively higher valuations for keywords including specific model names, and relatively lower valuations for other types of keywords. This difference may be explained by heterogeneity in the consumers who use the different keyword queries. For cameras, keyword queries with model-specific keywords result in the most narrow range of search results. Major camera manufacturers usually use unique model numbers to distinguish their products from others. For example, Nikon's DSLR (digital single-lens reflexive) camera models typically start with a "D" followed by numbers; e.g., D3, D90, D300s, etc. Similarly, Canon models typically begin with numbers followed by "D"; e.g., 550D and 5D.[23] Hence, shoppers querying with model-specific keywords are probably well-informed consumers who have a clear idea which specific products they are interested in, and are searching with a strong intention of purchasing. Our results indicate, then, that high quality online merchants, who are typically also more experienced, gravitate towards more narrowly defined keyword queries which are likely to be made by serious buyers.

In contrast, shoppers who use brand names to search may be more interested in browsing and collecting information about different camera models; searching for, say, "Nikon" will return a wide variety of models and accessories across many price points. These consumers may have a more muted intention of purchasing, and our results imply that high quality online merchants – again, those who are more experienced – have correspond-

[23]Fujifilm uses the combination of "X" and numbers, Sony uses "A" and numbers, Pentax use "K" and numbers, and Olympus usually starts with "E".

ingly lower assessments of these keywords. Finally, the design of WebsiteX ads partially neutralizes the effect of additional promotional terms. Most of the sponsored links directly contain price information (Figure 3), making it extremely easy for consumers to compare different prices on the internet. Hence, their purchase probabilities are unlikely to be swayed by purely marketing terms such as "big sale", rendering these promotional keywords of little value to the more-experienced high-quality merchants.

Table (4) and (5) summarize the estimation results of model II. Again, we focus on the results for the uniform equilibrium selection rule (in Table 4). The main differences between the Model I and Model II specifications is that the latter allows the parameters in bidder's valuations to be completely flexible vis-a-vis position rank. But even after allowing this flexibility, we find qualitatively similar patterns in valuations compared with the more restrictive model I results. As before, we find that high-quality bidders have relatively higher valuations for top positions in keyword queries involving product-specific keywords, but they have relatively lower valuations for top positions in queries involving promotional of brand-specific keywords. Thus our finding of horizontal differentiation in preferences appears robust to different specifications of bidders' valuations.

Quantitatively, the more flexible Model II specification does yield some additional findings. In some cases, it is the second position that generates the largest value for high quality merchants, not the first position; for specific keywords, as an example, high quality merchants value position 2 most highly: the coefficient for position 2 is 203.12=230.99-27.87, while for position 1 it is only 112.22=16.39+95.83. Similarly, medium-quality merchants also value the second position most highly in specific keyword queries. This phenomena may be related to the empirical fact we found in the click volume data, that often it is the second position that generates the largest click volume, even after the regression smoothing. From the merchants' perspective, the second position is almost as good as the first slot, because the click volume is comparable with the top but the per-click price is lower.

However, when the valuation matrix is not positively assortative, which is the case in our results, then it is unclear whether the GSP mechanism is socially optimal. As pointed out by Athey and Ellison (2011) and Chen and He (2011), the sponsored-link auction also plays the role of information intermediary. If the links are sorted according to merchants' quality, then it allows the consumer to efficiently search for merchants who fit their quality needs. If the valuation matrix is positively assortative between quality and ranking, then high quality merchants will bid aggressively and hence GSP is an efficient way to convey information to online shoppers.

On the other hand, as we noted before, WebsiteX merchants typically post their prices on the sponsored ads, thus providing consumers with the most important piece of purchase-related information without requiring further clicking behavior. In this setting, consumer search may be less relevant and, hence, the click volume in WebsiteX is less regular than described in the literature motivated from Yaoo!, Google and Microsoft; e.g., Varian (2006), Athey and Nekipelov (forthcoming).[24]

Finally, the estimated coefficients of the Beta distribution parameters a and b are positive, with $a >> b$, for both the Model I and model II estimates (bottom of tables 3 and 5). This implies that the equilibrium selection density function is left-skewed, so that higher prices are much more likely to be chosen. This suggests that the observed prices are better explained as optimal for WebsiteX, rather than for the merchants.

6 Counterfactual: The Effects of Bid Scoring

Using our estimation results, we now turn to the main policy question of this paper, which is assessing the effects of the bid-scoring policy which WebsiteX enacted only shortly after our sample period. This bid scoring was implemented using a score-weighted version of the GSPA (which we will call WGSPA hereinafter); that is, the positions are allocated by ranking the product of each bid times a "popularity score" for this bidder. Specifically, letting $\kappa_i, i = 1, \ldots, N$ denote the bidder-specific popularity scores, the positions are assigned according to the weighted bids: $\kappa_1 b_1 > \kappa_2 b_2 > \cdots > \kappa_N b_N$. Furthermore, the bidder winning position i pays a per-click price p_{ii} such that bidder i's score $\kappa_i p_{ii}$ is exactly equal to the score of the bidder in the $i+1$-th position:

$$\kappa_i p_{ii} = \kappa_{i+1} b_{i+1} \quad \text{or} \quad p_{ii} = \frac{\kappa_{i+1}}{\kappa_i} b_{i+1}. \tag{13}$$

The total payment of a bidder in the i-th position is $\alpha_i p_{ii}$. This mechanism essentially rewards the high quality advertisers with price discounts (if $\kappa_i > \kappa_{i+1}$, then $p_{ii} < b_{i+1}$, while $p_{ii} = b_{i+1}$ under the unscored GSPA rule); at the same time, this also incentivizes online merchants to improve their quality. Intuitively, offering price discounts may reduce the platform's revenue.[25] We show, however, it is possible that the per-click price can be even higher under WGSPA.

Importantly, in the WGSPA, the implicit "price" that the bidder in position i must pay to obtain position j differs from bidder to bidder (depending on their score κ_i), and hence

[24]While it is difficult to collect price data from WebsiteX, we collected a limited sample of screen captures and found no noticeably trends in product prices across ad positions.

[25]Indeed, Myerson's (1981) classic work suggests that, for standard auctions, a platform may be able to raise expected revenues by discriminating *against* high-valuation bidders.

the resulting game cannot be formulated as an assignment game à la Shapley-Shubik (in which agents are essentially price-takers). Suppose a generic bidder is indexed by i, and a generic position and the bid paid for that position is indexed by j. An allocation μ is a one-to-one function that maps each bidder's index to the corresponding position index. $\mu(i) = j$ means that bidder i is assigned to the j-th position, and the inverse mapping, $\mu^{-1}(j) = i$, would identify who is assigned to the j-th position.

Since the WGSPA cannot be formulated as an assignment game, the equilibrium allocation may not be unique. An allocation and a sequence of bids $(\mu; b_1, \ldots, b_N, b_{N+1})$ constitutes a symmetric Nash equilibrium if the following inequalities are satisfied[26]:

1. $\kappa_{\mu^{-1}(1)} b_1 > \kappa_{\mu^{-1}(2)} b_2 > \cdots > \kappa_{\mu^{-1}(N+1)} b_{N+1}$ (Allocation Rule)
2. $\alpha_i p_{\mu^{-1}(i)i} \leq V_{\mu^{-1}(i)i}$ for all i (Individual Rationality)
3. Incentive compatibility:

$$V_{\mu^{-1}(i)i} - \alpha_i p_{\mu^{-1}(i)i} \geq V_{\mu^{-1}(i)j} - \alpha_j p_{\mu^{-1}(i)j}, \quad \text{for all } (i, j)$$

where the counterfactual deviation prices are defined by: $p_{\mu^{-1}(i)j} = \dfrac{\kappa_{\mu^{-1}(j+1)}}{\kappa_{\mu^{-1}(i)}} b_{j+1}$

Sources at WebsiteX tell us that one key element in determining their score index κ_i is the historical performance and click volume of the advertiser i, which is highly correlated with WebsiteX's own quality rating system (see section 4). Therefore, we perform the counterfactual analysis under two alternative scoring systems (1) a *coarser* one with $\kappa_i \in \{1, 2, 3\} \equiv \{\text{low, medium, high quality}\}$; and (2) a *finer* one with $\kappa_i \in \{3, 4, \ldots, 15\}$.[27]

We simulate the per-click prices $\tilde{\mathbf{p}}_t$ under the new mechanism. We also compute the corresponding componentwise upper bound $\bar{\mathbf{p}}_t$, and the componentwise lower bound $\underline{\mathbf{p}}_t$. Due to the computational burden of computing SNE of WGSPA, we only consider auctions with no more than 7 positions. Complete implementation details are given in the appendix. The summary statistics of the simulated (cross-sectional) per-click price distribution under different model specifications are summarized in Tables 6 and 7.

Results. For simplicity, we focus on the results for the coarser scoring rule, in Table 6, in our discussion. First, bid scoring appears to "steepen" the price gradient across positions, with the top position increasing in price but lower positions decreasing in price. Specifically, in the Model I-Uniform results, the price per click for the top position increases, on average, by 12 RMB (around 2 USD) relative to the baseline unscored

[26] See Varian (2007) for more details.

[27] As we do not have the data for the losers, we assume κ_{N+1} to be the lowest score.

scenario, while price for the second position decreases by 3 RMB (around \$0.50). Results are similar for Model II: there, the price for the top position increases by 6 RMB (around \$1) and the price for the second position also increases (by 2 RMB = \$0.30), but the prices for all lower positions decrease relative to the baseline scenario.

The price increases for the top positions are striking, especially as the bid-scoring rules were intended to reward popular merchants (those with high scores) with price discounts. To assess the extent of price discounting under the bid-scoring system, in Table 6 we also provide summary statistics for the "bid-discount ratios" $\frac{\kappa_{j+1}}{\kappa_j}$ for position j, which are a measure of the price discount (cf. Eq. (13)). When this ratio is less than one, then the bidder winning position j was given a price discount, while if it exceeds one, then the bidder paid a price premium. We see that, across all four specifications, and across the top four positions, this ratio was less than one in over 80% of the simulations. This implies that the prices changed in response to bid scoring *despite* the winning bidders being given price discounts. Apparently the price increases for the top positions were triggered by more aggressive bidding in response to the introduction of bid scoring.

In Table 6, we also provide summary statistics for the bid-discount ratios $\frac{\kappa_{j+1}}{\kappa_j}$ for the baseline unscored scenario. There, we see that these statistics are not much different from those generated in the bid scoring scenarios. This implies that bid-scoring does not increase the degree of assortative matching in the allocation positions – it does not appear that larger, more popular merchants were systematically more likely to end up in top positions after the move to bid scoring. Thus while the fears of the small merchants quoted in this paper's introduction – that they have to pay much higher prices to get their ads in top positions – seem justified to some extent by our simulation results, at the same time popular merchants are not more likely (relative to the baseline scenario) to get these top positions.

Finally, the right-hand side of Table 6 shows that total platform revenue from the auctions remains unchanged from bid scoring. At the least, the finding that revenue does not decrease upon the introduction of bid scoring may justify WebsiteX's move toward WGSPA: it can promote long term benefits beyond the scope of our analysis (such as improvements in ad quality, and buyer responsiveness to ads) without sacrificing short term revenue.

7 Conclusion

We conclude with a brief summary of some main points from our study. Empirically, we uncover some horizontal differentiation between different types of keywords, as high quality merchants have relatively higher valuations for informative keywords, and relatively lower valuations for vague keywords. This suggests that high quality bidders, who are typically also more experienced merchants, gravitate towards more narrowly defined keyword queries which are likely to be made by serious buyers, and leave the top positions in broader keyword queries for less experienced (lower quality) merchants. Counterfactual evaluations show that the price trend becomes steeper under the score weighted generalized second price auction, with much higher prices obtained for the top position but lower prices for the other positions. Apparently, while scoring the auction grants price discounts to popular merchants, our findings suggest that scoring also heightens the bid competition, thus leading to higher prices for top positions. Overall, we do not find large effects on platform revenue and sorting patterns from shifting to a scoring rule.

Our methodological approach is also novel as it is motivated by the equivalence between symmetric Nash equilibria in GSPA, and stable outcomes in the classic assignment game of Shapley and Shubik (1972). To accommodate some stylized empirical facts in sponsored search auctions, our specification generalizes previous work by allowing bidders to have preferences for positions which are not multiplicative in bidder- and position-specific effects, and click volumes to be nondecreasing with position ranks. For estimation, we utilize a Bayesian procedure and develop a Metropolis-Hastings within Gibbs sampler.

Inevitably, some strong assumptions underlie our analytical framework in this paper. It is possible, albeit at the cost of computational expense, to extend our Bayesian estimation procedure to a more general setup, in order to accommodate asymmetric Nash equilibria (as studied in Börgers et al (2013)), or bidder-specific click volumes (as in Jeziorski and Moorthy (2014)). We leave these for future inquiry.

References

Akkus, O., J. A. Cookson and A. Hortacsu (2013). The Determinants of Bank Mergers: A Revealed Preference Analysis, working paper.

Albert, J. H. and Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association*, **88**, 669–679.

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

Athey, S. and D. Nekipelov (forthcoming). A Structural Model of Sponsored Search Advertising Auctions, *Journal of Political Economy*, to appear.

Bajari, P., H. Hong, and S. Ryan (2010). Identification and Estimation of Discrete Games of Complete Information, *Econometrica*, **78**, 1529–1568.

Becker, G. S. (1973). A theory of marriage: part I, *Journal of Political Economy*, **81**, 813–846.

Beresteanu, A. and I. Molchanov and F. Molinari (2011). Sharp Identification Regions in Models with Convex Moment Predictions, *Econometrica*, **79**, 1785-1821.

Börgers, T. and I. Cox and M. Pesendorfer and V. Petricek (2013). Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence, *American Economic Journal: Microeconomics*, **5**, 61-71.

Boyd, D., H. Lankford, S. Loeb and J. Wyckoff (2006). Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets, NBER working paper.

Choo, E. and A. Siow (2006). Who Marries Whom and Why, *Journal of Political Economy*, **114**, 175–201.

Chen, Y., and C. He (2011). Paid Placement: Advertising and Search on the Internet, *Economic Journal*, **121**, 309–328.

Chernozhukov, V., H. Hong, and E. Tamer (2007). Estimation and Inference on Identified Parameter Sets, *Econometrica*, **75**, 1243–1284.

Chiappori, P.-A., R. J. McCann and L. P. Nesheim (2009). Hedonic Price Wquilibria, Stable matching, and Optimal Transport: Equivalence, Topology, and Uniqueness, *Economic Theory*, **42**, 317–354.

Ciliberto, F. and E. Tamer (2009). Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, **77**, 1791–1828.

Dantzig, G. B. (1968). *Linear Programming and Extensions*, Princeton University Press, Princeton, NJ.

Demange, D., D. Gale and M. Sotomayor (1986). Multi-Item Auctions, *Journal of Political Economy*, **94**, 863–872.

Echenique, F., S. Lee, M. Shum, and M. Yenmez (2013). The Revealed Preference Theory of Stable and Extremal Stable Matchings, *Econometrica*, **81**, 153–171.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords, *American Economic Review*, **97**, 242–259.

Fox, J. T. (2013). Estimating Matching Games with Transfers, working paper.

Fox, J. T. and P. Bajari (2013). Measuring the Efficiency of an FCC Spectrum Auction, *American Economic Journal: Microeconomics*, **5**, 100–146.

Galichon, A. and M. Henry (2011). Set Identification in Models with Multiple Equilibria, *Review of Economic Studies*, **78**, 1264-1298.

Galichon, A. and B. Salanié (2012). Cupids Invisible Hand: Social Surplus and Identification in Matching Models. *Working Paper.*

Goldman, M. and J. Rao (2014). Experiments as Instruments: Heterogeneous Position Effects in Sponsored Search Auctions, working paper, UCSD.

Graham, B. S. (2011). Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers, *Handbook of Social Economics*, **1B**, 965–1052, Amsterdam, North-Holland.

Haile, P. A. and E. Tamer (2003). Inference with an Incomplete Model of English Auctions, *Journal of Political Economy*, **111**, 1–51.

Hong, H. and M. Shum. (2003). Econometric Models of Asymmetric Ascending Auctions, *Journal of Econometrics*, **112**, 327-358.

Hsieh, Y.-W. (2011). Understanding Mate Preferences from Two-Sided Matching Markets: Identification, Estimation and Policy Analysis, manuscript, New York University.

Jerath, K., L. Ma, Y.-H. Park and K. Srinivasan (2011). A Position Paradox in Sponsored Search Auctions, *Marketing Science*, **30**, 612–627.

Jeziorski, P. and S. Moorthy (2014). Advertiser Prominence Effects in Search Advertising, working paper, UC Berkeley, Haas School of Business.

Jeziorski, P. and I. Segal (forthcoming). What makes them Click: Empirical Analysis of Consumer Demand for Search Advertising, *American Economic Journal: Microeconomics.*

Krasnokutskaya, E. and K. Seim (2011). Bid Preference Programs and Participation in Procurement, *American Economic Review*, **101**, 2653-2686.

Logan, J. A., P. D. Hoff, and M. A. Newton (2008). Two-Sided Estimation of Mate Preferences for Similarities in Age, Education, and Religion, *Journal of the American Statistical Association*, **103**, 559–569.

McAfee, P. and J. McMillan (1989). Government Procurement and International Trade, *Journal of International Economics*, **26**, 291-308.

Menzel, K. (2011). MCMC Methods for Incomplete Structural Models of Social Interactions, *Working Paper.*

Myerson, R. (1981). Optimal Auction Design. *Mathematics of Operations Research*, **6**, 58-73.

Roth, A. E. and M. A. O. Sotomayor (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monographs No. 18, Cambridge University Press.

Shapley, L. S. and M. Shubik (1972). The Assignment Game, I: The Core *International Journal of Game Theory*, **1**, 111–130.

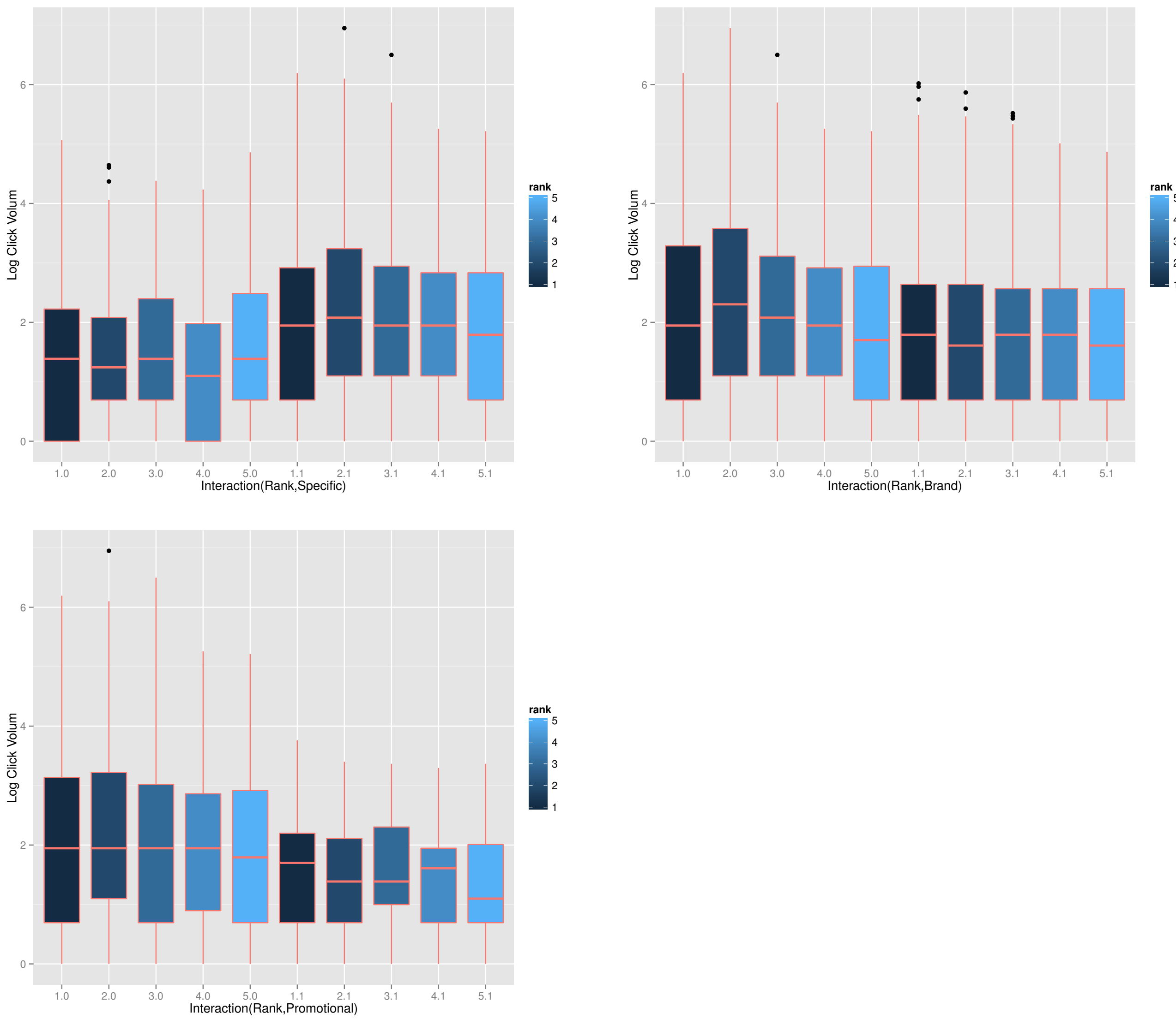
Tanner, M. and W. Wong (1987). The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association*, **82** 528-540.

Uetake, K. and Y. Watanabe (2012). Entry by merger: Estimates from a Two-sided Matching Model with Externalities, working paper, Yale School of Management.

Varian, H. R. (2007). Position Auction, *International Journal of Industrial Organization*, **25**, 1163–1178.

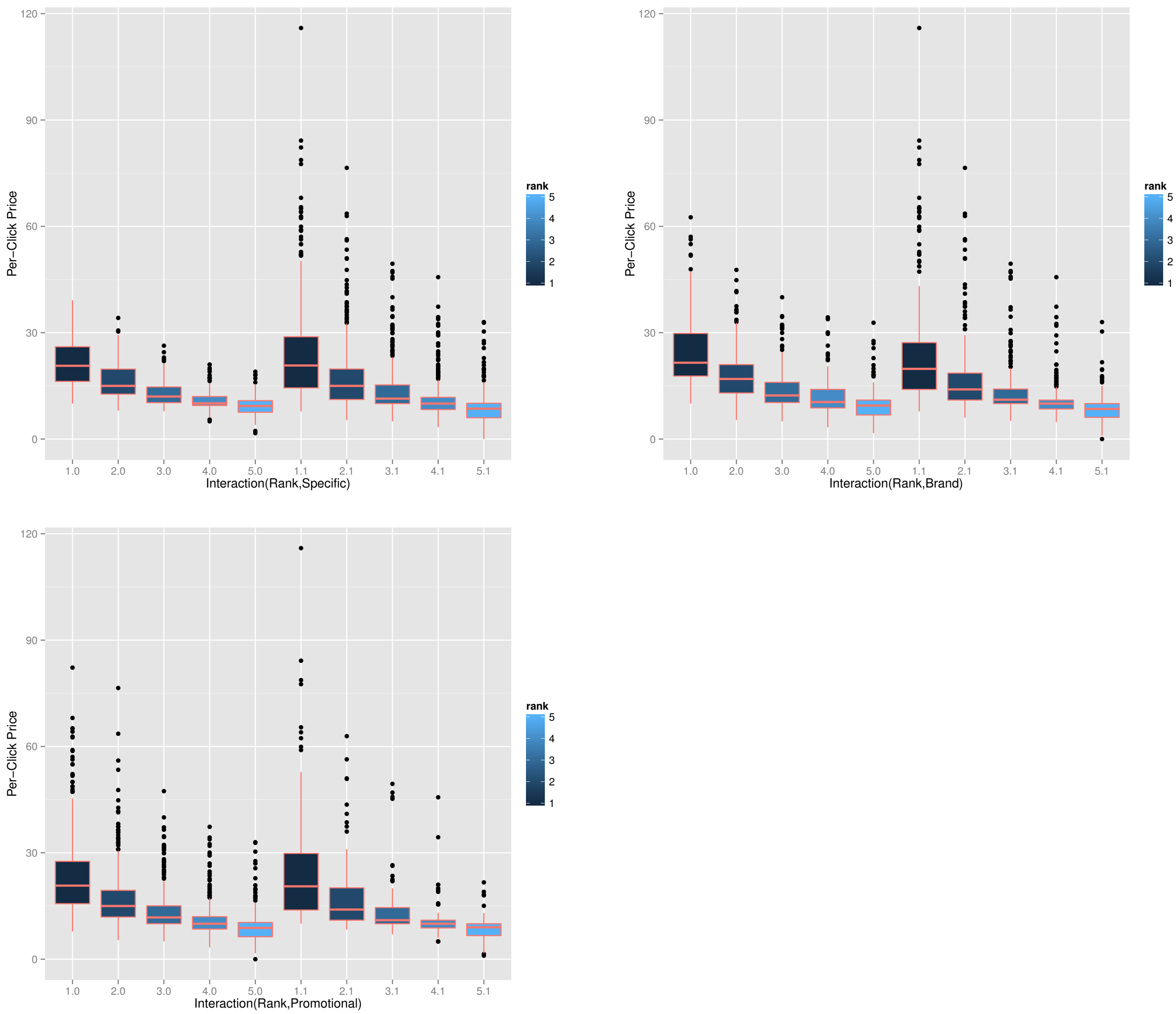
Yang, S., S. Lu and X. Lu (2014). Modeling Competition and Its Impact in Paid-Search Advertising, *Marketing Science*, **33**, 134-153.

Figure 1: Boxplot of Log Click Volume



*The pair of numbers on the horizontal axis represents the values of two variables. For example, Interaction(Rank,Specific)= 3.1 on the x-axis of the top-left graph means “3rd position, Specific dummy = 1”

Figure 2: Boxplot of Per-Click Price



*The pair of numbers on the horizontal axis represents the values of two variables. Per-click prices are given in Chinese Renminbi (RMB, or "yuan"), with exchange rate roughly 6RMB=1USD. For example, Interaction(Rank,Specific)= 3.1 on the x-axis of the top-left graph means "3rd position, Specific dummy = 1"

Figure 3: Sample Search Results in WebsiteX

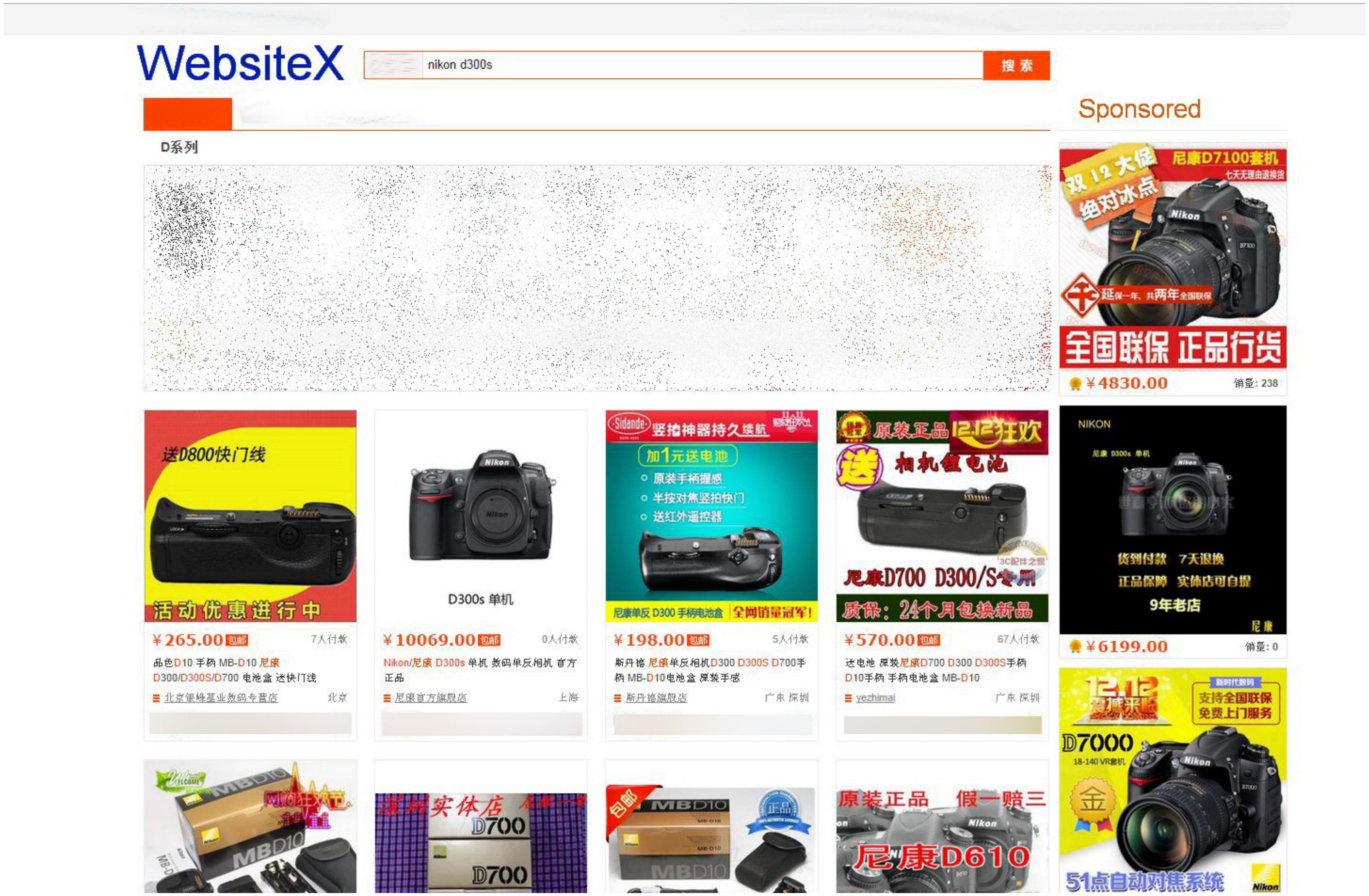
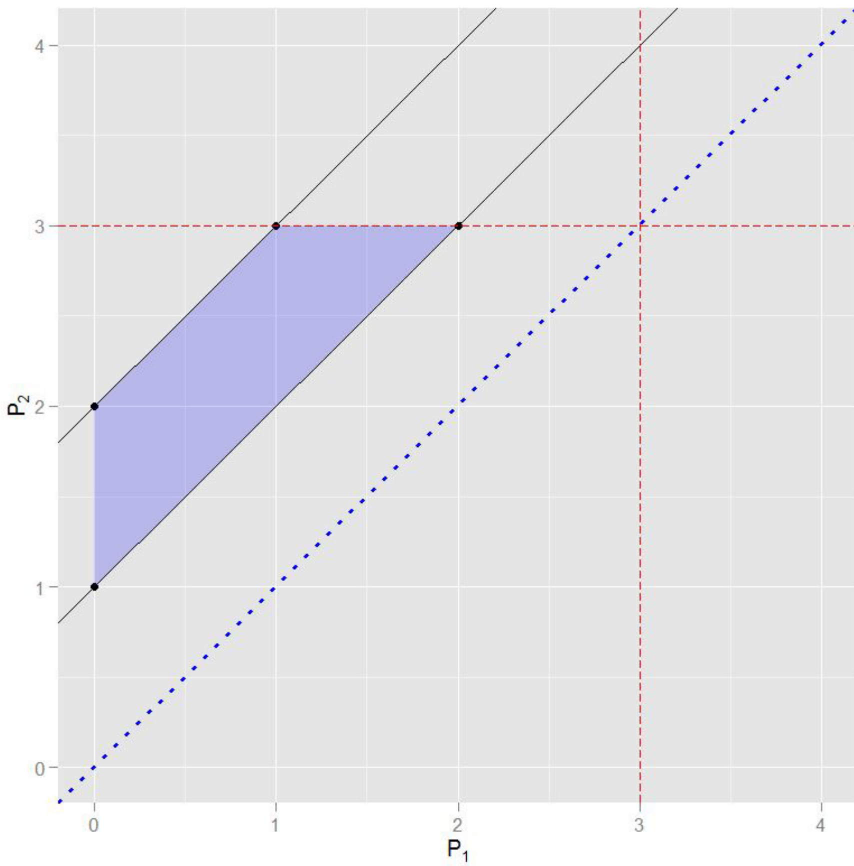


Figure 4: Non-Existence of Equilibrium Price System of GSPA



solid black line: no-blocking pair condition; dashed red line: individual rationality condition; blue dotted line: GSPA condition $p_1 > p_2$; shaded area: set of stable matching without imposing $p_1 > p_2$.

Table 1: Variable List and Summary Statistics

Variable	Definition	mean	min	max
Auction Characteristics				
Per-Click Price (RMB)		13.25	0	115.95
		(9.25)		
Click Volume		16.37	0	1040
		(42.98)		
Keyword Characteristics				
Brand	= 1 if keyword includes the brand name	0.63	0	1
		(0.48)		
Specific	= 1 if keyword includes a specific model number	0.77	0	1
		(0.42)		
Promotional	= 1 if keyword includes promotional terms	0.18	0	1
		(0.39)		
Bidder Characteristics				
High Quality	= 1 for high quality merchant	0.2	0	1
		(0.03)		
Medium Quality	= 1 for medium quality merchant	0.74	0	1
		(0.03)		

*Standard error in parentheses.

Table 2: Contingency Table of Bidders' Quality versus Position Ranks

All Keywords					
	Position 1	2	3	4	5
High	7.2*	4.9	3.7	3.8	2.6
Medium	12.6	14.6	15.6	14.9	16.1
Low	0.2	0.5	0.7	1.3	1.3
Keywords with Specific= 1					
	Position 1	2	3	4	5
High	7.1	4.7	3.7	4.1	2.9
Medium	12.7	14.7	15.8	15.0	16.2
Low	0.2	0.6	0.5	0.9	0.9
Keywords with Brand= 1					
	Position 1	2	3	4	5
High	7.5	4.6	3.7	3.7	2.2
Medium	12.3	14.7	15.5	15.0	16.6
Low	0.2	0.7	0.7	1.4	1.2
Keywords with Promotional= 1					
	Position 1	2	3	4	5
High	6.6	3.4	3.2	2.7	2.0
Medium	13.2	16.1	16.6	15.9	17.0
Low	0.2	0.5	0.2	1.4	0.9

*All numbers in the table are given as percentage of all bidders.

Table 3: MCMC Estimation: Model I

Dummy Regressor	Posterior	Model I Uniform Eq. Sel.	Model I Beta Eq. Sel.
constant	mean	-151.57	-214.89
	s.d.	(5.92)	(8.00)
specific	mean	-63.42	-21.96
	s.d.	(6.82)	(7.25)
promotional	mean	-14.58	-35.65
	s.d.	(10.10)	(12.69)
brand	mean	39.65	40.89
	s.d.	(6.27)	(9.98)
high	mean	688.96	704.04
	s.d.	(19.41)	(26.86)
medium	mean	479.75	563.92
	s.d.	(12.75)	(16.26)
high×specific	mean	539.45	382.34
	s.d.	(21.65)	(21.84)
medium×specific	mean	516.41	422.27
	s.d.	(16.60)	(15.21)
high×promotional	mean	-42.53	-139.99
	s.d.	(23.43)	(38.72)
medium×promotional	mean	-161.82	-86.42
	s.d.	(20.68)	(26.22)
high×brand	mean	-341.05	-307.86
	s.d.	(18.08)	(25.68)
medium×brand	mean	-395.04	-405.03
	s.d.	(14.67)	(18.03)
σ^2	mode	40484	49429
	mean	40385	49513
	s.d.	(793)	(366)
left parameter of beta distr. a	mean	N/A	3.27
	s.d.	N/A	(0.59)
right parameter of beta distr. b	mean	N/A	0.83
	s.d.	N/A	(0.09)

Table 4: MCMC Estimation: Model II-Uniform Equilibrium Selection

		Rank of Positions				
Dummy Interaction Terms	Posterior	1	2	3	4	5
high	mean	95.83	-27.87	-360.55	-431.62	-463.62
	s.d.	(15.34)	(26.93)	(16.00)	(22.21)	(24.45)
medium	mean	-175.62	-140.20	-281.94	-424.62	-359.59
	s.d.	(26.16)	(23.98)	(19.24)	(31.72)	(25.66)
high×specific	mean	16.39	230.99	-125.74	-157.83	-294.49
	s.d.	(34.01)	(28.67)	(38.07)	(44.10)	(52.92)
medium×specific	mean	-15.96	124.32	-190.90	-236.87	-332.36
	s.d.	(24.92)	(24.99)	(28.15)	(13.95)	(23.43)
high×promotional	mean	-356.05	-336.30	-153.56	-241.04	-78.22
	s.d.	(22.89)	(26.42)	(25.57)	(33.76)	(21.71)
medium×promotional	mean	-288.32	-324.80	-264.19	-305.36	-313.11
	s.d.	(21.04)	(14.11)	(18.97)	(31.80)	(21.93)
high×brand	mean	-109.84	-454.61	-47.33	-5.62	-16.25
	s.d.	(18.72)	(27.74)	(30.65)	(27.37)	(37.19)
medium×brand	mean	-114.13	-385.46	-50.34	96.47	67.63
	s.d.	(35.35)	(22.66)	(19.39)	(26.89)	(23.22)
Non-Interaction Terms		Constant	Specific	Promotional	Brand	
	mean	366.29	306.39	208.95	-70.1	
	s.d.	(20.82)	(21.94)	(14.89)	(19.79)	
σ^2	mode	36579				
	mean	36849				
	s.d.	(330)				

Table 5: MCMC Estimation: Model II-Beta Equilibrium Selection

		Rank of Positions				
Dummy Interaction Terms	Posterior	1	2	3	4	5
high	mean	19.82	-102.09	-400.17	-483.75	-522.48
	s.d.	(19.67)	(28.49)	(17.07)	(25.14)	(24.15)
medium	mean	-229.56	-186.11	-338.50	-479.67	-416.46
	s.d.	(15.04)	(15.60)	(20.58)	(14.98)	(20.22)
high×specific	mean	-217.75	-15.15	-361.31	-401.81	-490.52
	s.d.	(24.10)	(54.45)	(52.33)	(31.95)	(25.76)
medium×specific	mean	-247.19	-125.35	-426.39	-471.28	-565.03
	s.d.	(44.74)	(35.69)	(50.40)	(40.62)	(29.00)
high×promotional	mean	-334.95	-363.30	-179.40	-174.20	-124.02
	s.d.	(34.52)	(40.65)	(30.01)	(26.91)	(57.01)
medium×promotional	mean	-323.63	-351.31	-304.38	-320.22	-345.52
	s.d.	(31.75)	(31.43)	(27.12)	(27.11)	(19.50)
high×brand	mean	-189.93	-495.68	-134.79	-71.69	-101.03
	s.d.	(20.33)	(20.66)	(16.70)	(24.47)	(23.43)
medium×brand	mean	-202.43	-459.60	-117.65	19.30	-3.05
	s.d.	(17.07)	(17.93)	(26.27)	(24.02)	(22.43)
Non-Interaction Terms		Constant	Specific	Promotional	Brand	
	mean	397.12	549.03	240.05	10.14	
	s.d.	(13)	(34.03)	(26.28)	(17.96)	
σ^2	mode	32767				
	mean	32911				
	s.d.	(668)				
left parameter of beta distr. a	mean	2.62				
	s.d.	(0.43)				
right parameter of beta distr. b	mean	0.73				
	s.d.	(0.06)				

Table 6: Counterfactual per-click prices for Symmetric Nash Equilibrium of WGSPA: Finer Score

(Units of prices are Chinese RMB.)

		Rank of Positions					Total Revenue
Price-Raw Data	Statistics	1	2	3	4	5	
	mean	22.32	16.13	12.68	10.27	8.12	1133
	median	19.6	13.67	11	9.74	7.73	910
	variance	139.62	62.8	32.15	18.95	12.16	616476
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.70	0.61	0.63	0.65	0.80	
Price-Model I-Uniform	mean	35.17	13.02	8.12	5.10	2.74	1120
	median	34.55	11.78	7.45	4.34	2.01	1107
	variance	118.22	46.53	24.59	15.02	6.71	174682
	$[E(\underline{p}), E(\bar{p})]$	$[27.51, 41.02]$	$[7.18, 16.81]$	$[3.14, 11.56]$	$[1.15, 8.72]$	$[0.28, 6.68]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.67	0.63	0.61	0.60	0.83	
Price-Model I-Beta	mean	36.75	14.72	9.84	7.13	4.88	1267
	median	36.01	13.30	8.98	6.28	3.92	1251
	variance	139.84	62.22	35.26	25.25	16.64	196707
	$[E(\underline{p}), E(\bar{p})]$	$[25.69, 39.42]$	$[6.55, 16.37]$	$[2.73, 11.21]$	$[1.02, 8.49]$	$[0.22, 6.43]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.68	0.62	0.6	0.59	0.84	
Price-Model II-Uniform	mean	28.32	17.89	10.37	6.00	2.92	1137
	median	27.32	17.65	9.80	5.08	2.12	1088
	variance	103.90	50.03	33.68	21.97	8.75	219406
	$[E(\underline{p}), E(\bar{p})]$	$[19.38, 34.60]$	$[11.60, 22.25]$	$[4.18, 14.64]$	$[1.10, 10.49]$	$[0.10, 7.86]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.59	0.65	0.62	0.59	0.9	
Price-Model II-Beta	mean	30.29	19.42	12.14	8.03	5.19	1290
	median	29.11	19.06	11.56	7.01	3.97	1250
	variance	113.44	50.93	39.81	34.88	23.73	243226
	$[E(\underline{p}), E(\bar{p})]$	$[18.56, 33.09]$	$[11.15, 21.28]$	$[3.9, 13.79]$	$[0.95, 9.61]$	$[0.07, 7.10]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.59	0.66	0.61	0.59	0.91	

Table 7: Counterfactual per-click prices for Symmetric Nash Equilibrium of WGSPA: Coarser Score

(Units of prices are Chinese RMB.)

		Rank of Positions					Total Revenue
Price-Raw Data	Statistics	1	2	3	4	5	
	mean	22.32	16.13	12.68	10.27	8.12	1133
	median	19.6	13.67	11	9.74	7.73	910
	variance	139.62	62.8	32.15	18.95	12.16	616476
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.86	0.84	0.86	0.84	0.91	
Price-Model I-	mean	34.98	13.02	8.13	5.10	2.73	1119
Uniform	median	34.28	11.84	7.45	4.38	1.99	1101
	variance	118.54	46.76	24.09	14.37	6.55	170717
	$[E(\underline{p}), E(\bar{p})]$	$[27.23, 40.88]$	$[7.15, 16.82]$	$[3.08, 11.6]$	$[1.13, 8.77]$	$[0.27, 6.28]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.87	0.86	0.85	0.85	0.93	
Price-Model I-	mean	37.09	14.88	10.03	7.16	4.99	1277
Beta	median	36.07	13.39	9.16	6.34	4.09	1262
	variance	147.48	64.55	36.24	25.55	17.01	197976
	$[E(\underline{p}), E(\bar{p})]$	$[25.74, 39.78]$	$[6.54, 16.60]$	$[2.76, 11.41]$	$[0.98, 8.53]$	$[0.22, 6.59]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.86	0.84	0.86	0.86	0.93	
Price-Model II-	mean	28.36	17.98	10.45	6.05	2.92	1142
Uniform	median	27.39	17.66	9.93	5.09	2.13	1091
	variance	105.26	49.80	32.91	22.00	8.26	220235
	$[E(\underline{p}), E(\bar{p})]$	$[19.34, 34.63]$	$[11.65, 22.39]$	$[4.19, 14.77]$	$[1.09, 10.61]$	$[0.11, 7.88]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.80	0.86	0.85	0.83	0.95	
Price-Model II-	mean	30.22	19.52	12.12	8.13	5.21	1286
Beta	median	29.12	19.15	11.61	7.15	4.07	1243
	variance	110.84	52.13	37.75	33.41	21.10	240496
	$[E(\underline{p}), E(\bar{p})]$	$[18.57, 33.05]$	$[11.23, 21.39]$	$[3.85, 13.79]$	$[0.99, 9.77]$	$[0.08, 7.16]$	
$Pr(\frac{\kappa_{i+1}}{\kappa_i} \leq 1)$		0.80	0.85	0.84	0.83	0.95	

Appendix A Complete details of estimation procedure

A.1 Specifying the Priors

We assume the prior distribution of β follows normal distribution $\mathcal{N}(\beta_0, B_0)$ and σ^2 follows inverted Gamma distribution $\mathcal{IG}(\alpha_0/2, \delta_0/2)$. β and σ^2 are independent and their joint distribution will be denoted by $\pi(\theta)$. We further assume that $V_{ij} = \delta(X_i, Z^j; \beta) + \epsilon_{ij}$ with ϵ_{ij} follows i.i.d. $\mathcal{N}(0, \sigma^2)$[28] across the index (i, j) and different auction t. Although there always exists a stable matching for an arbitrary draw of V_{ij}, it is not the case once the GSP restriction $p_1 > p_2 > \cdots > p_N$ being imposed. The intersection of $p_1 > p_2 > \cdots > p_N$ and the set of stable matchings of Shapley and Shubik may be empty. When estimating the model we shall restrict our attention to the set of $\mathbf{V}_t$ that can guarantee the existence of an equilibrium[29]:

$$\{\mathbf{V}_t \in \text{GSP}\} \equiv \{\exists(t_1, \ldots, t_N) | (t_1, \ldots, t_N) \text{ solves (DP) \& } p_1 > p_2 > \ldots p_N; t_i = \alpha_i p_i\}.$$

The probability of this region is $Pr\{\mathbf{V}_t \in \text{GSP}\} = c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$ and can be approximated by simulation.[30] This specification, together with the restriction of the existence of equilibrium, implies that the joint distribution of $\mathbf{V}_t$ is a multivariate truncated normal distribution

$$f(\mathbf{V}_t | \theta; \mathbf{X}_t, \mathbf{Z}_t) = \mathbb{1}(\mathbf{V}_t \in \text{GSP}) \frac{\Pi_{i=1}^N \Pi_{j=1}^N \frac{1}{\sigma} \phi\left(\frac{V_{ijt} - \delta(X_{it}, Z^{jt}; \beta)}{\sigma}\right)}{c(\theta; \mathbf{X}_t, \mathbf{Z}_t)},$$

where $\phi(\cdot)$ is the pdf of standard normal distribution. The specification of $p_0(\theta, \mathbf{V}|\mathbf{X}, \mathbf{Z})$ is now completed after specifying $\pi(\theta)$ and $f(\mathbf{V}|\theta; \mathbf{X}, \mathbf{Z})$ since $p_0(\theta, \mathbf{V}|\mathbf{X}, \mathbf{Z}) = f(\mathbf{V}|\theta; \mathbf{X}, \mathbf{Z})\pi(\theta)$.[31] Finally, for the Beta equilibrium selection model, we assume uniform priors over the shape paraeters (a, b): $\pi(a, b) \propto \frac{1}{c}^2, 0 < c < \infty$.

A.2 Algorithm

We propose a Metropolis-Hastings within Gibbs sampler to draw $(\theta, \lambda, \mathbf{V})$ from the posterior. Gibb's sampling loops over the three conditional distributions $f_1(\mathbf{V}|\theta, \lambda, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha) =$

[28] While we use normality assumption in the empirical study, the proposed algorithm here can be easily applied to other distributions.

[29] Here the indices $1, \ldots, N$ refer only to the positions, and not to specific bidders.

[30] See A.3 for the implementation detail.

[31] By the independence assumption $f(\mathbf{V}|\theta; \mathbf{X}, \mathbf{Z}) = \Pi_{t=1}^T f(\mathbf{V}_t|\theta; \mathbf{X}_t, \mathbf{Z}_t)$

$\Pi_{t=1}^{T} f_1(\mathbf{V}_t|\theta,\lambda,\mu_t,\mathbf{p}_t,\mathbf{X}_t,\mathbf{Z}_t,\alpha_t)$, $f_2(\theta|\mathbf{V},\lambda,\mu,\mathbf{p},\mathbf{X},\mathbf{Z},\alpha)$, and $f_3(\lambda|\theta,\mathbf{V},\mu,\mathbf{p},\mathbf{X},\mathbf{Z},\alpha)$. As it is difficult to draw directly from the above conditional densities, one can instead using Metropolis-Hastings sampler. For f_1, we simulate $\mathbf{V}_t$ for each t from normal distribution, truncated to $A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t)$. This step amounts to impose equilibrium restriction, but without explicitly adjusting for the effect of multiple equilibria. We then use an independent M-H step to correct for it later, essentially weighted by the equilibrium selection probability. For f_2, it is a truncated normal likelihood function and hence can be simulated using standard M-H procedure too. Below is the implementation detail:

1. Conditional on $(\theta^{(\tau)},\lambda^{(\tau)},\mathbf{V}^{(\tau)})$, update $\mathbf{V}^{(\tau+1)}$ via independence M-H chain

1.1 Simulate

$$\begin{aligned}\tilde{\mathbf{V}}_t &\sim q_t(\tilde{\mathbf{V}}_t|\mathbf{V}_1^{(\tau+1)},\ldots,\mathbf{V}_{t-1}^{(\tau+1)},\mathbf{V}_t^{(\tau)},\mathbf{V}_{t+1}^{(\tau)},\ldots,\mathbf{V}_T^{(\tau)},\theta^{(\tau)},\lambda^{(\tau)}) \\ &= q_t(\tilde{\mathbf{V}}_t|\theta^{(\tau)}) \propto \mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\tilde{\mathbf{V}}_t)) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)};\mathbf{X}_t,\mathbf{Z}_t).\end{aligned}$$

Following Logan, et al. (2008) we suggest the following steps to simulate $\tilde{V}_{ijt}$[32]: 1. simulate the diagonal elements $\tilde{V}_{iit}$ from normal distribution with mean $\delta(X_{it},Z^{it};\beta^{(\tau)})$ and variance $\sigma^2_{(\tau)}$, left-truncated at $\alpha_{it}p_{it}$ (individual rationality). 2. Given the simulated $\tilde{V}_{iit}$, one then proceed to simulate the off-diagonal elements $\tilde{V}_{ijt}, i \neq j$. The NBP condition implies that $\tilde{V}_{ijt}$ follows normal distribution with mean $\delta(X_{it},Z^{jt};\beta^{(\tau)})$ and variance $\sigma^2_{(\tau)}$, right-truncated at $\tilde{V}_{iit} - (\alpha_{it}p_{it} - \alpha_{jt}p_{jt})$. 3. Notice that the monotonicity condition of per-click price does not directly affect the simulation of V_{ij}. Because the observed data already satisfies the monotonicity condition, the simulated V_{ij} will automatically lead to an equilibrium polytope of per-click prices that intersects with the set $p_1 > p_2 > \ldots, p_N$.

1.2 Take

$$\mathbf{V}_t^{(\tau+1)} = \begin{cases} \mathbf{V}_t^{(\tau)} & \text{with probability } 1-\rho, \\ \tilde{\mathbf{V}}_t & \text{with probability } \rho, \end{cases}$$

[32]The standard GHK simulator does not apply in this case.

where

$$\rho = 1 \wedge \Big[\frac{f_1(\tilde{\mathbf{V}}_t|\theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)}{f_1(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)}\Big]\Big[\frac{q_t(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)})}{q_t(\tilde{\mathbf{V}}_t|\theta^{(\tau)})}\Big]$$

$$= 1 \wedge \Big[\frac{\mathcal{L}_1(\mathbf{p}_t|\tilde{\mathbf{V}}_t, \theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}{\mathcal{L}_1(\mathbf{p}_t|\mathbf{V}_t^{(\tau)}, \theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}\Big]\Big[\frac{q_t(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)})}{q_t(\tilde{\mathbf{V}}_t|\theta^{(\tau)})}\Big]$$

$$= 1 \wedge \Big[\frac{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\tilde{\mathbf{V}}_t, \alpha_t), \lambda^{(\tau)}, \mu_t) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\mathbf{V}_t^{(\tau)}, \alpha_t), \lambda^{(\tau)}, \mu_t) f(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}\Big]\Big[\frac{\mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t^{(\tau)})) f(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}; \mathbf{X}_t, \mathbf{Z}_t)}{\mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\tilde{\mathbf{V}}_t)) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)}; \mathbf{X}_t, \mathbf{Z}_t)}\Big]$$

$$= 1 \wedge \frac{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\tilde{\mathbf{V}}_t, \alpha_t), \lambda^{(\tau)}, \mu_t)}{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\mathbf{V}_t^{(\tau)}, \alpha_t), \lambda^{(\tau)}, \mu_t)}.$$

In particular, if the equilibrium selection rule is uniform then $\rho = 1 \wedge \frac{vol(\mathbf{P}(\mathbf{V}_t^{(\tau)}, \alpha_t))}{vol(\mathbf{P}(\tilde{\mathbf{V}}_t, \alpha_t))}$. In this case, $\tilde{\mathbf{V}}_t$ will be accepted with probability 1 if the resulting polyhedron has smaller volume relative to $\mathbf{V}_t^{(\tau)}$. To sum up, one would need to independently simulate T valuation matrices for T keyword auctions, and then run T independent M-H steps to decide whether to accept the new draws or not.

2. Conditional on $(\theta^{(\tau)}, \lambda^{(\tau)}, \mathbf{V}^{(\tau+1)})$, update $\theta^{(\tau+1)}$ via random walk M-H chain

2.1 Simulate

$$\tilde{\theta} \sim q(\tilde{\theta}|\theta^{(\tau)}) = \begin{cases} A^{-1} & \text{for } \tilde{\theta} \in \theta^{(\tau)} \pm a \\ 0 & \text{otherwise,} \end{cases}$$

where a is a vector of the same dimension as θ and A is the volume of the box spanned by $\theta_i^{(\tau)} \pm a_i$. As $q(\cdot|\cdot)$ is a random-walk proposal density, it is also symmetric: $q(\tilde{\theta}|\theta^{(\tau)}) = q(\theta^{(\tau)}|\tilde{\theta})$.

2.2 Take

$$\theta^{(\tau+1)} = \begin{cases} \theta^{(\tau)} & \text{with probability } 1-\rho, \\ \tilde{\theta} & \text{with probability } \rho, \end{cases}$$

where

$$\rho = 1 \wedge \frac{f_2(\tilde{\theta}|\mathbf{V}^{(\tau+1)}, \lambda^{(\tau)}, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)}{f_2(\theta^{(\tau)}|\mathbf{V}^{(\tau+1)}, \lambda^{(\tau)}, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)}$$

$$= 1 \wedge \frac{\Pi_{t=1}^{T} f_1(\mathbf{V}_t^{(\tau+1)}|\tilde{\theta}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)\pi(\tilde{\theta})}{\Pi_{t=1}^{T} f_1(\mathbf{V}_t^{(\tau+1)}|\theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)\pi(\theta^{(\tau)})}$$

$$= 1 \wedge \frac{\Pi_{t=1}^{T} \mathcal{L}_1(\mathbf{p}_t|\mathbf{V}_t^{(\tau+1)}, \tilde{\theta}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\mathbf{V}_t^{(\tau+1)}|\tilde{\theta}, \mathbf{X}_t, \mathbf{Z}_t)}{\Pi_{t=1}^{T} \mathcal{L}_1(\mathbf{p}_t|\mathbf{V}_t^{(\tau+1)}, \theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\mathbf{V}_t^{(\tau+1)}|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)} \frac{\pi(\tilde{\theta})}{\pi(\theta^{(\tau)})}$$

Because $\mathcal{L}_1$ only depends on $(\mathbf{V}_t, \alpha_t, \lambda)$ and by construction $\mathbf{V}_t^{(\tau+1)} \in$ GSP, the above equation can be further simplified:

$$= 1 \wedge \frac{\Pi_{t=1}^{T}\left[\Pi_{i=1}^{N}\Pi_{j=1}^{N}\frac{1}{\tilde{\sigma}}\phi\left(\frac{V_{ijt}^{(\tau+1)}-\delta(X_{it}, Z^{jt}; \tilde{\beta})}{\tilde{\sigma}}\right)\right]}{\Pi_{t=1}^{T}\left[\Pi_{i=1}^{N}\Pi_{j=1}^{N}\frac{1}{\sigma^{(\tau)}}\phi\left(\frac{V_{ijt}^{(\tau+1)}-\delta(X_{it}, Z^{jt}; \beta^{(\tau)})}{\sigma^{(\tau)}}\right)\right]} \frac{\Pi_{t=1}^{T} c(\theta^{(\tau)}; \mathbf{X}_t, \mathbf{Z}_t)}{\Pi_{t=1}^{T} c(\tilde{\theta}; \mathbf{X}_t, \mathbf{Z}_t)} \frac{\pi(\tilde{\theta})}{\pi(\theta^{(\tau)})}$$

This step is nothing but treating $\mathbf{V}^{(\tau+1)}$ as the data, and then evaluate the likelihood ratio of the truncated normal density.

3. Conditional on $(\theta^{(\tau+1)}, \lambda^{(\tau)}, \mathbf{V}^{(\tau+1)})$, update $\lambda^{(\tau+1)}$ via random walk M-H chain

3.1 Simulate

$$\tilde{\lambda} \sim q(\tilde{\lambda}|\lambda^{(\tau)}) = \begin{cases} A^{-1} & \text{for } \tilde{\lambda} \in \lambda^{(\tau)} \pm a \\ 0 & \text{otherwise,} \end{cases}$$

where a is a vector of the same dimension as λ and A is the volume of the box spanned by $\lambda_i^{(\tau)} \pm\ a_i$.

3.2 Take

$$\lambda^{(\tau+1)} = \begin{cases} \lambda^{(\tau)} & \text{with probability } 1-\rho, \\ \tilde{\lambda} & \text{with probability } \rho, \end{cases}$$

where

$$\rho = 1 \wedge \frac{f_3(\tilde{\lambda}|\theta^{\tau+1}, \mathbf{V}^{\tau+1}, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)}{f_3(\lambda^{\tau}|\theta^{\tau+1}, \mathbf{V}^{\tau+1}, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)}$$

$$= 1 \wedge \frac{\Pi_{t=1}^{T} \mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\mathbf{V}_t^{\tau+1}, \alpha_t), \tilde{\lambda}, \mu_t)}{\Pi_{t=1}^{T} \mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\mathbf{V}_t^{\tau+1}, \alpha_t), \lambda^{\tau}, \mu_t)}.$$

Under uniform equilibrium selection, there is no need to perform step 3.

A.3 Other Implementation Details

A.3.1 Characterizing the set of equilibrium prices

As discussed in section 3.1 above, the set of equilibrium per-click prices is a convex polyhedron defined by various linear inequalities. First, we consider the no-blocking pair conditions in Eq. (1). one can substitute $u_i + t_i = V_{ii} \forall i$ into the rest of NBP conditions conditions: The resulting system of inequalities are

$$V_{ii} - V_{ij} \geq t_i - t_j \geq V_{ji} - V_{jj}. \tag{8}$$

Equivalently, equation (8) can be re-written using the matrix notation

$$\{\mathbf{p} | A_1 \mathbf{p} \leq b_1\}.$$

Take a 3-by-3 case as an example,

$$A_1 = \begin{bmatrix} \alpha_1 & -\alpha_2 & 0 \\ -\alpha_1 & \alpha_2 & 0 \\ \alpha_1 & 0 & -\alpha_3 \\ -\alpha_1 & 0 & \alpha_3 \\ 0 & \alpha_2 & -\alpha_3 \\ 0 & -\alpha_2 & \alpha_3 \end{bmatrix}, \text{ and } \quad b_1 = \begin{bmatrix} V_{11} - V_{12} \\ V_{22} - V_{21} \\ V_{11} - V_{13} \\ V_{33} - V_{31} \\ V_{22} - V_{23} \\ V_{33} - V_{32} \end{bmatrix}. \tag{9}$$

The second set of inequalities are the individual rationality conditions; i.e., bidders' payoff should be positive

$$\{\mathbf{p} | A_2 \mathbf{p} \leq b_2\},$$

where A_2 is the diagonal matrix with $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ on the main diagonal and $b_2 = (V_{11}, V_{22}, \ldots, V_{NN})'$. Finally, the last set of inequalities states that $p_1 > p_2 > \cdots > p_N \leq 0$, corresponding to the non-negative price systems that are consistent with the generalized second price auction.

$$\{\mathbf{p} | A_4 \mathbf{p} \leq b_4\},$$

For a 3-by-3 case,

$$A_3 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \text{ and } \quad b_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{10}$$

along with the nonegativity constraints

$$\{\mathbf{p}|A_4\mathbf{p} \leq b_4\},$$

where $A_4 = -\mathbf{I}_{N\times N}$ and $b_4 = -\underline{0}\mathbf{1}_{N\times 1}$.

vThe set of equilibrium prices is given by $\mathbf{P}(\mathbf{V},\alpha) = \{\mathbf{p}|A(\alpha)\mathbf{p} \leq \mathbf{b}(\mathbf{V})\}$, where $A = (A_1'|A_2'|A_3'|A_4')'$ and $\mathbf{b} = (b_1'|b_2'|b_3'|b_4')'$.

A.3.2 Parameter Setup

For the first specification of bidder valuations (Model I), the prior for β is assume to be joint normal distribution. The mean vector β_0 equals to the zero vector, and the covariance matrix B_0 equals to $100000I$, where I is an 12-by-12 identify matrix. The prior for σ^2 is assumed to be inverted Gamma distribution with shape parameter $\alpha_0 = 2$ and scale parameter $\delta_0 = 5$. We choose the uniform distribution on the half-line as the prior for the shape parameters (γ_1, γ_2) of the beta distribution when estimating the equilibrium selection density. This set of prior values only impose minimum prior information on the parameters. For example, when the shape parameter is 2, the variance of inverted Gamma distribution does not exist. Moreover, given the "large sample" feature in the Bayesian updating step for (β, σ),[33]the prior specification only plays a negligible role in determining the posterior. We make 30,000 draws and the first 10,000 draws are treated as the burn-in. We then keep every 50th draw of the remaining 20,000 draws to estimate the posterior mean and standard deviations. The radius for the random walk proposal for β, σ^2, and (γ_1, γ_2) are respectively 100, 1000 and 0.2.

For Model II, the prior for β is assume to be joint normal distribution. The mean vector β_0 equals to the zero vector, and the covariance matrix B_0 equals to $100000I$, where I is an 44-by-44 identify matrix. We make 50,000 draws and the first 30,000 draws are treated as the burn-in. We then keep every 50th draw of the remaining 20,000 draws to estimate the posterior mean and standard deviations. The radius for the random walk proposal for β, σ^2, and (γ_1, γ_2) are respectively 5, 100 and 0.2.

A.3.3 Calculating the Volume of Equilibrium Polytope of Prices

In order to evaluate the likelihood, one has to compute the volume of $\mathbf{P}(\mathbf{V},\alpha)$ if uniform selection is imposed. We do it via simulation: first, we draw 1000 independent (multivariate) uniform random numbers from the smallest bounding box of $\mathbf{P}(\mathbf{V},\alpha)$. Second,

[33]In a 5-player-5-position game, 25 V_{ij} will be drawn in MCMC. If there are 100 keyword auctions, (β, σ^2) will be estimated by 2500 simulated V_{ij}.

the volume of $\mathbf{P}(\mathbf{V}, \alpha)$ is approximately the volume of the bounding box times the proportion of the previous draws that belong to $\mathbf{P}(\mathbf{V}, \alpha)$. If beta selection is imposed, one instead draw beta random numbers from the bounding box. We also try 5000 draws and the accuracy seems to be similar.

A.3.4 Calculating the Truncation Probability $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$

We use a simulation-extrapolation strategy to compute the truncation probability, as we find that $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$ is a relatively smooth function of θ. Given θ, and for each keyword auction t we can simulate N valuation matrices $\mathbf{V}_i$ to approximate $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$ by $\frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(\mathbf{V}_i \in \text{GSP})$. This step would require solving (DP). We first make 10,000 MCMC draws by ignoring this truncation probability. The first 1000 draws are discarded, and we keep every 10th draw of the remaining chain. We simulate $c(\theta_j; \mathbf{X}_t, \mathbf{Z}_t)$ under these simulated θ_j, $j = 1, 2, \ldots, 900$. Finally, we regress $c(\theta_j; \mathbf{X}_t, \mathbf{Z}_t)$ on θ_j using beta regression[34], with the logistic function being the link function. The estimated regression coefficients are then used to calculate $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$ in MCMC. One important fact is that if $\mathbf{V}_i \notin \text{GSP}$, then $a + b\mathbf{V}_i \notin \text{GSP}$, where (a, b) are some scalar constants. It is the relative size of each component within $\mathbf{V}$ that leads to nonexistence of equilibrium, not because of its scale. As a result, the scale parameter and the non-interacted location parameters (see Appendix B) does not affect $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$. When running the regression, we discard the non-interacted location parameters, and normalized the interacted location parameters by the scale parameter σ.

Appendix B Identification

The specification of the latent valuation V_{ij} implicitly assumes that the distribution of V_{ij} belongs to the location-scale family. However, the meaning of *location* and *scale* should be carefully interpreted in matching models, as location parameter for V_{ij} may actually possess scale effect on the dependent variable $\mathbf{p}$.

There are three types of parameters in the specification of V_{ij}: interacted location parameter $(\beta_4, \ldots, \beta_{11})$, non-interacted location parameter $(\beta_0, \ldots, \beta_3)$ and scale parameter σ. First, the interacted location parameters characterize the preference over position ranks. Or equivalently, they characterize the complementarity between quality and ranks, and hence will determine the (cross-sectional) distribution of allocation. They are also related to the price distribution through the channel of the no-blocking-pair conditions,

[34] Beta regression is a flexible regression model to handel the cases when the dependent variable is proportion.

because they essentially determine the shape of the equilibrium polytope of price. The point identification results for the interacted location parameters (up to scale normalization) from the allocation data μ have been established in Choo-Siow model and Fox (2010) under different model assumptions.

The price data essentially provides identification power for parameters with scale effect on the valuation matrix. There are two types of such parameters: First, the non-interacted location parameters cannot be identified through the no-blocking-pair conditions, as they are differenced out. Instead, the individual rationality condition can be used to learn information about them. Such parameters do not affect the preference over ranks, and consequently they have no effect on the allocation. However, as they would shift the scale of the valuation matrix, they will also shift the size of the equilibrium polytope of prices. Similarly, the scale parameter σ have no effect on the allocation, but it will also affect the size of the equilibrium polytope of price. Both non-interacted location and scale parameters have scale effect on the price distribution. The difference is, the non-interacted location parameters only shift the size of the equilibrium polytope in *certain* directions (through individual rationality), while the scale parameter shift the size of the equilibrium polytope in *all* directions. By looking at the price distribution, and the shape/size of the support of prices one can then learn information about these parameters. The intuition is simple: if on average the price is \$5, the scale of $\mathbf{V}$ cannot be around \$1, as it would violate the individual rationality. On the other hand, it cannot be around \$100, as \$5 would be too cheap under competitive bidding.

Remark: a partial identification approach. It is possible to derive a bound for the structural parameters in terms of bounds of CDF (similarly to Haile and Tamer, 2003).

Theorem 1. *Suppose the econometrician observes T independent auctions indexed by $t = 1, 2, \ldots, T$. Let $F(\mathbf{p}|\mathbf{X}, \mathbf{Z}, \alpha)$ represents the conditional distribution of observed per-click prices. We denote by $\mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta)$, the smallest (random) bounding box that covers $\mathbf{P}(\mathbf{V}, \alpha) = \mathbf{P}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta)$. $\mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta) = \Pi_{i=1}^{N}[p_{il}, p_{iu}]$, where (p_{il}, p_{iu}) are respectively the smallest and largest elements of the i-th coordinate of $\mathbf{P}$. Define two random vectors*
$\bar{\mathbf{p}} = \max \mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta) = (p_{1u}, p_{2u}, \ldots, p_{Nu})$ *and* $\underline{\mathbf{p}} = \min \mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta) = (p_{1l}, p_{2l}, \ldots, p_{Nl})$.
The identified set of θ is given by

$$\Theta_0 = \left\{\theta | F(\underline{\mathbf{p}}|\mathbf{X}, \mathbf{Z}, \alpha; \theta) \geq F(\mathbf{p}|\mathbf{X}, \mathbf{Z}, \alpha) \geq F(\bar{\mathbf{p}}|\mathbf{X}, \mathbf{Z}, \alpha; \theta)\right\},$$

Proof. Conditional on $(\mathbf{X}, \mathbf{Z})$, the (joint) distribution function of $\mathbf{p}$, $F(\mathbf{p}|\mathbf{X}, \mathbf{Z})$, is identified by the sampling process. Given $(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta)$, $\mathbf{p}$ is a measurable selection from the set of stable prices $\mathbf{P}$. By construction, $\underline{\mathbf{p}} \leq \mathbf{p} \leq \bar{\mathbf{p}}$, for almost-all ϵ_{ij}, where the inequality are defined componentwise. This implies that the distributions of $\underline{p}$, p, and $\bar{p}$ are ordered in the sense of first-order stochastic dominance, which immediately implies the bound on CDF. ■

Although the bound approach is appealing as it does not require an equilibrium selection assumption, it is extremely computationally demanding to compute such bound since for each draw of latent valuation matrix, one has to solve (DP).[35] By contrast, as we discussed earlier, the Bayesian approach does not require solving for game for each draw of the latent valuations, and hence is computatinoally more appealing. See also Uetake and Watanabe (2012) for a bounds approach for estimating a two-sided matching model applied to bank mergers.

Appendix C Complete details of counterfactual simulation

To solve for the set of SNE in the WGSPA, we consider each possible allocation in turn; for each candidate allocation, we use the inequalities characterizing WGSPA to determine the set of equilibrium bids (which may be empty if this candidate allocation is not an equilibrium), and then repeat this routine for all possible allocations. Consequently, the problem of solving SNE of WGSPA is combinatorial. Parallel to the standard GSPA, the above inequalities can be written as matrix form. The set of equilibrium bid under allocation μ will be referred as $B_\mu \equiv \{\mathbf{b} | D_\mu(\alpha, \kappa)\mathbf{b} \leq c_\mu(\mathbf{V})\}$.

We then simulate the structural parameters θ from the posterior distribution (by directly using the MCMC output), and for each θ we draw the utility shocks ϵ_{ijt} to obtain the valuation matrix $\mathbf{V}_t$. Given $\mathbf{V}_t$ one can solve the game to obtain equilibrium per-click prices. As the number of SNE may be huge, we do not attempt to solve all SNE. Instead, we employ a simple routine to perform the counterfactual analysis. As long as an equilibrium allocation is found, we then stop searching for another equilibrium allocation[36] and simulate b_i from the equilibrium polyhedron of bids B_μ[37] according to

[35] Using the approach of Chernozhukov et al. (2007), the total number of times which the game must be solvedgame solving is at least equal to the numberθ times the number of keyword auctions (= 487) times the number of $\mathbf{V}$ draws. Moreover, depending on $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$, one has to discard many $\mathbf{V}$ draws that do not satisfy the GSPA restriction.

[36] To avoid the concentration on a particular type of allocation, we perform random search.

[37] There is no need to simulate the maximum bid b_1 for the following two reasons: First, the per-click

uniform distribution or the estimated beta distribution. Finally, we apply the scored pricing rule in Eq. (13) to obtain the per-click price $\tilde{\mathbf{p}}_t$.

prices are irrelevant of b_1. Second, if $(b_2, \dots, b_{N+1})$ satisfy all the inequalities, one can always choose b_1 large enough to meet the allocation rule.