To Score or Not to Score? Structural Estimation of Sponsored Search Auctions

Yu-Wei Hsieh* USC

Matthew Shum† Caltech

Sha Yang‡§ USC Marshall School of Business

Preliminary draft: November 3, 2014

Abstract

We estimate a structural model of a sponsored search auction model. The econometric model matches two pieces of data for each auction: (i) the *allocation* of sponsored search listing to the bidders; and (ii) the *per-click prices* paid by the winners. To accommodate the "position paradox", our model relaxes the assumption of decreasing click volumes with position ranks, which is often assumed in the literature. We estimate the model using data from the Chinese online marketplace Taobao.com. We find that sellers of different qualities adopt different bidding strategies: high quality sellers bid more aggressively for informative keywords, while low quality sellers are more likely to be sorted to the top positions for vague keywords. Counterfactual evaluations show that score weighting in sponsored search auctions would dramatically increase the per-click price for the top ad position, but leave prices for lower positions largely unaffected.

Keywords: Generalized Second Price Auction; Sponsored-search Advertising; Structural Estimation; Assignment Game; Two-sided Matching; MCMC.

JEL classification: D44; D47; C11; C15

*yuwei.hsieh@usc.edu

†mshum@caltech.edu

‡shayang@marshall.usc.edu

§We thank Arie Beresteanu, Baiyu Dong, Hashem Pesaran, Roger Moon, Geert Ridder and Guofu Tan for their helpful discussions. Thanks also go to the seminar participants at USC, UC-Riverside, UCLA, 2013 Annual Conference of Taiwan Econometric Society, 2014 California Econometrics Conference and 2014 North American Summer Meeting of the Econometric Society.

1 Introduction

Auctions have become the dominant mechanism whereby sponsored search results are sold by major search engines on the Internet (eg. Google, Yahoo, Microsoft Bing). In this paper, we estimate a structural model of these keyword auctions, and recover bidders' payoffs from each position. The data come from Taobao.com, one of the largest websites in the world by traffic. Taobao is a Chinese-language online market platform, where sellers post prices and sell directly to consumers. There is no direct analog of Taobao in the United States, but it shares features of both eBay and Craigslist.

We utilize a novel econometric setup and model which exploits an equivalence between the position auctions and the classical assignment game of Shapley-Shubik (1972), which has been used to model two-sided matching markets with transferable utility. In contrast with most sponsored-search auction models in the literature, we do not assume that click volumes decrease with position ranks; this accommodates the "position paradox" (Jerath, Ma, Park and Srinivasan, 2011), an empirical stylized fact that ads in high positions do not systematically obtain a higher click volume. Methodologically, we develop a nested "Metropolis-Hastings within Gibbs" algorithm for Bayesian inference of our model. This estimation procedure deals conveniently with two factors which complicate the structural estimation of the sponsored-search auction model: the large number of latent variables (namely, the valuations that each bidder places on an ad in each position), as well as the multiplicity of equilibrium bids/prices. Another key benefit of our Bayesian estimator is that it does not require resolving the game for each draw of latent variables.

Our estimation results show that sellers of different qualities adopt different bidding strategies: high quality sellers bid more aggressively for informative keywords, while low quality sellers are more likely to be sorted to the top positions for vague keywords. Thus we find evidence of both horizontal and vertical differentiation in these auctions; auction outcomes can be non-assortative, in that higher quality sellers do not obtain higher positions.

Building on this, we use counterfactual simulations to explore several policy questions in auction design. During our sample period, Taobao awarded positions purely based on the prices which were bid. However, at the present time, and following most other platforms running ad auctions, bids are multiplied by a "score" factor, which typically depends on a popularity index of each bidder. By "scoring" firms in this way, platforms try to award the higher positions to more popular firms, who will generate more web traffic (and hence raise ad revenues for the platform). Using our estimation results, we show that scoring would lead to higher per-click price for the 1st position, but the prices only decrease slightly for the rest of positions.

Existing Literature

At Taobao, as at most other websites utilizing sponsored search auctions, ad positions are allocated via a Generlized Second Price Auction (GSPA). Bidder who submit the j-th highest bid will be assigned to the j-th highest position to display her link, and the per-click price she will be charged is the $j+1$-th highest bid. The total payment to the search engines depends on the click volume, the number of clicks through the sponsored-link, times the per-click price determined in auction. Edelman, Ostrovsky, and Schwarz (EOS, 2007) and Varian (2007) pointed out that this is a multi-item auction (Demange, Gale and Sotomayor, 1986) between bidders under complete information and, formally, is identical to the classic assignment game in Shapley and Shubik (1972).[1]

Although sponsored-link auctions have received great attention in the theoretical literature, empirical research remains sparse. Ghose and Yang (2009) analyze the Google data, and Yang, Lu and Lu (2013) structurally estimate EOS's model. As a result, they are unable to accommodate the position paradox. 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 allow for bidders' bids to be multiplied by a "quality score", and also accommodate uncertainty in bidders' perceptions (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). These features complicate the underlying bidding model substantially. Due to these complexities, in practice many bidders rely on 3rd-party marketing firms that dynamically manage (possibly via automatic bidding algorithms) their positions and per-click prices.

In contrast, the auctions run by Taobao were simpler – during our sample period, there was no score weighting on the bid, and bidders needed to actively revise their bids, rather than relying on automatic bidding mechanism. Another distinct feature is that Taobao bidders usually revise their bids less frequently, rather than the real-time style in Yahoo! or Google auction. For this reason, in this paper, we estimate a model which closely follows the Shapley-Shubik assignment game (and hence EOS and Varian). Even though this game presumes complete information among the bidders regarding each other's valuations, so that auction outcomes (both allocations and prices) are deterministic from the bidders' point of view, the outcomes are stochastic for the econometrician, due to unobserved variables (such as characteristics of the bidders or the auctions).

At the same time, because our econometric approach to sponsored search auctions is based on the Shapley-Shubik model, this paper has a deep connection with the recent empirical literature using two-sided matching models. Most of the papers here (including Choo and Siow (2006), Fox (2013), Galichon and Salanie (2012), Graham (2011)) consider matching games with large (or continuum) number of players. In contrast, our approach in this paper uses the exact solution of the game, and is more suitable for markets with small number of players (as in sponsored search auctions, where typically only a small number of positions are available). Moreover, most of these papers consider the case when price data are unavailable, and the corresponding estimators essentially difference out these (unobserved) prices.

In contrast, a distinguishing feature of our econometric approach is that we model the equilibrium prices explicitly and then utilize the price variation to estimate the model parameters. This is appropriate as the main object of interest in studying auctions is to model the price process. In the empirical two-sided matching literature, Fox and Bajari (2013) and Akkus, Cookson and Hortacsu (2013) also discuss how to use the transfer data to estimate

[1] This model has been used to model two-sided matching markets with transfers, including marriage markets (Becker (1973)) and hedonic pricing equilibria in differentiated product markets (Chiappori, McCann, Nesheim (2009)).

matching models, but a distinguishing characteristic of the two-sided matching approach as applied to sponsored search auctions is that the generalized second-price auction mechanism imposes a *monotonicity* condition on the prices, which not only has ramifications for the existence of equilibrium, but also raises nontrivial difficulties with the econometric modeling of the prices.[2]

In the following section we provide some details on Taobao.com, our study company, and present the structural model for the generalized second price auction. Here, we also derive the crucial link between the auction model and two-sided matching models, which we exploit in estimation. In section 3 we derive the likelihood function and show how to estimate the structural parameters vis Metropolis-Hastings within Gibbs sampler. Section 4 describes the dataset. The main empirical results are presented in section 5. In section 6 we perform the counterfactual analysis for the weighted generalized second price auction. Section 7 concludes.

2 A Structural Model for Sponsored Search Auctions

2.1 Background: sponsored search auctions at Taobao.com

Our study company, Taobao.com, is the largest online marketplace 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 sellers promote their wares and do most of their business via Taobao. As a marketplace, it has no direct American counterpart, but shares features of both eBay and Craigslist. Unlike eBay, goods on Taobao are not sold via auction, but rather by sellers posting prices for their products. Taobao provides a platform whereby buyers can make secure money transactions to sellers. Sponsored search results typically appear as "tiles" on the right-hand side and bottom margins of each search page (See Figure 3).

Since Taobao is a marketplace, the content and role of its sponsored ads differs substantially from the ads appearing on search pages of general search engines such as Google or Yahoo!. Taobao's sponsored ads (as well as the non-sponsored "organic" search results) typically contain a picture of the product, price, seller's information, promotional terms, the number of recent transactions, product specification, etc; these are typically missing from Google's sponsored ads. The sponsored links on Google usually contain the URL with a short explanation only.

As most of the shoppers who use Taobao are, in fact, searching for a good price relative to the retail stores[3], they can simply click on the ad with the best price quote regardless of its position. Moreover, image based search ads (with price and quality attributes) can

[2]Similar issues arise in structural econometric models of pricing in markets with pure vertical differentiation (eg. Bresnahan (1981), Esteban and Shum (2007)) or ascending auctions with asymmetric bidders (eg. Hong and Shum (2003)).

[3]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 on Taobao can sidestep these hidden cost. Consequently, the price gap between online and retail store is even larger than the USA counterpart.

change users' search behavior significantly (the marketing literature suggests that attribute prominence can lead to users' selective processing of information). Consequently, the top positions may not always receive more clicks.[4] For this reason, the standard assumption in the existing sponsored search auction models (Edelman, Ostrovsky, and Schwarz (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 an nonincreasing vector of position-specific clickthrough rates) seems inadequate for the Taobao setting. In our econometric specification, we allow the surplus to vary arbitrarily among sellers and across ad positions. Furthermore, click volumes need not to be diminishing with position rank in our setup.

2.2 Generalized Second-Price Auction (GSPA)

Next we introduce our model of sponsored-search auctions which, while based on the existing literature, has special features which are tailored to the Taobao. The basic model is a generalized second-price auction. Suppose there are N available positions and $M \geq N+1$ potential bidders for a generic keyword auction.[5] 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. Following Varian (2007), we focus on the so-called "symmetric" Nash equilibria in this complete-information bidding game.[6] The equilibrium conditions 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.[7] Now 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),[8] we

[4]The click volume is more "irregular" in Taobao data, in the sense that click volume does not generally appear to be systematically lower for ads in lower positions. This fact has been documented in the marketing literature known as position paradox; e.g., Jerath, Ma, Park and Srinivasan (2011).

[5]It is straightforward to consider the case where the number of bidders exceeds the number of positions. However, for the empirical exercise, we assume that the number of bidders coincides with the number of positions. This is because the data at hand only contains information on the "winning bids" (ie. the price per click for each position) and hence the valuations for the non-winning bidders would not be identified.

[6]See also Börgers et. al. (2013). These are closely-related to the "locally envy-free" equilibria which EOS focus on. These equilibria are convenient to analyze, and easy to compute via linear programming; as noted in Börgers et. al. (2013), no such characterization is available for the asymmetric Nash equilibria. Despite this fact, our estimator can still be extended to the case of asymmetric Nash equilibria. We leave the details as a future research.

[7]For now, we will assume that these click volume parameters are non-stochastic, but in the empirical work below we allow them to be incompletely known to bidders at the time they bid.

[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.

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 $u_i \equiv V_{ii} - \alpha_i p_i$ the equilibrium payoff of bidder i, and $t_j \equiv \alpha_j p_j$ 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)). Moreover, introducing the binary indicators $\mu(i, j) = 1$ if bidder i obtains position j, and zero otherwise, and summing up across all bidders and positions, we have

$$\sum_{i,j} u_i + t_j = \sum_{i,j} \mu(i, j) V_{ij}. \tag{3}$$

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.[9] 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 get the and first position, then the resulting μ is

[9]Following the convention of estimating complete information game, 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 matchings and μ_ω as an generic element of Ω, where $\omega = 1, 2, \ldots, N!$. The total surplus under matching μ_ω is denoted by $S_{\mu_\omega}(\mathbf{V})$

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

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

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

Based on $\mathbf{V}$, the total surplus $S(\mathbf{V})$ for each allocation $\mu_\omega \in \Omega$ can be calculated. The linear sum assignment problem in Shapley and Shubik (1972) consider the 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)):

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

Since there are only finitely many allocation in Ω, such maximization problem always exist a solution (Roth and Sotomayor, 1990).[10]

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

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 $\{S_{\mu_\omega}(\mathbf{V})\}_{\omega=1,\ldots,N!}$ is strict almost surely. ■

This result stands in sharp contrast to the two-sided matching models without transfer in which the multiple equilibria is the main concern; e.g., Boyd et al. (2006), Logan, Hoff and Newton (2008), Menzel (2011), Hsieh (2011), and Echenique, Lee, Shum and Yenmez (2013), among others.

[10]For algorithm that solves (P), consult Dantzig (1968) and Burkard et al (2009).

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, $u_i \geq 0, t_j \geq 0$, is known as individual rationality condition: both bidders and search engine should have non-negative profit. The second set of constraints, $u_i + t_j \geq V_{ij}$, is also known as the no-blocking-pair condition: if bidder i does not match with the j-th position, it must be that the profit demanded by both parties exceed what the matching can support. The set of (u_i, t_j) that solves (DP) is defined as stable matching (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. (3) above, which provides the link between the GSPA and the assignment game, as we alluded to before.

In general, there exists multiple transfer $\mathbf{t} = (t_1, \ldots, t_N)$ that solve (DP), and there exist bidder-optimal $\underline{\mathbf{t}}$ and seller-optimal $\bar{\mathbf{t}}$.[11] By contrast, the corresponding optimal matching μ that solves (P) is unique almost surely if the error term ϵ_{ij} is a continuous random variable. Among the set of stable matching $\mathbf{t}$ we should restrict our attention to the subset

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

This is the set of stable matching that selected by the mechanism of generalized second price auction, and will be referred as the set of stable per-click prices.

Remark 1: (Non-)existence of equilibrium in GSPA. Although Shapley and Shubik proved that the set of stable matching is always nonempty for arbitrary V_{ij}, it is not the case once the monotonicity condition on prices, $p_1 > p_2 > \ldots p_N$, being imposed. Without extra assumptions on V_{ij} or the click volumes, the generalized second-price mechanism may not necessarily guarantee the existence of a Nash equilibrium (or equivalently, competitive price system). For example, if the valuation matrix is given by

	1	2
bidder1	3	4
bidder2	1	3

[11] The literature on multi-item auction concerns the design problem of selecting particular stable matching by auction mechanism. For example, Demange, Gale and Sotomayor (1986) propose an auction mechanism in which the bidder-optimal stable matching is the equilibrium outcome. Another example is EOS who consider the equilibrium refinement problem of Varian (2007).

and $\alpha_1 = \alpha_2 = 1$, it is easy to verify that intersection of $p_1 > p_2$ and the set of stable matching of Shapley-Shubik is empty. The idea is that when both players value the second position more than the first one, then there is no way to sell the first position with a higher market price. Because we implicitly allow for arbitrary V_{ij} according to assumption 1, such cases will arise with positive probability. Since we want to restrict attention only to the set of V_{ij} that lead to a nonempty set of stable per-click price, this introduces a complicated truncation problem, similar to that arising in Hong and Shum's (2003) econometric study of asymmetric ascending auctions. ■

Remark 2: Specification of bidders' preferences. Varian (2007), EOS, and Athey and Nekipelov (forthcoming) consider a multiplicative specification of bidder's valuations:

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

By assuming the decreasing pattern on the click volume with rank, $\alpha_1 > \alpha_2 \cdots > \alpha_N$, it is easy to show that the optimal allocation is perfect assortative matching: the k-th position assigns to the k-th highest valuation. Furthermore, Varian (2007) is able to analytically solve (DP). In particular, the lower bound of the stable matching is given by

$$b_k = v_{k:N} - \frac{\alpha_k}{\alpha_{k-1}}(v_{k:N} - b_{k+1}), \tag{6}$$

where b_k is the k-th highest bid and $v_{k:N}$ is the k-th highest valuation out of N bidders. This is also known as the locally-envy free equilibrium in Edelman, Ostrovsky and Schwarz (2007). The subsequent empirical research are mostly based on this characterization to infer valuation distribution from bid and click volume data; e.g., Yang, Lu and Lu (2013). The identification argument is straightforward as one can invert the bid to obtain the order statistics, and hence the valuation distribution is nonparametrically identified.

The specification in (5) together with decreasing click volumes, however, implies that every bidder has exactly the same preference ordering over positions: everyone prefers higher position. By contrast, the specification of V_{ij} here is more flexible that allow us to recover richer pattern from the data. Furthermore, the pattern of click volume in reality is more complicated than the monotonicity assumption implies. Therefore, to account for these stylized facts it is necessarily to relax Varian's assumptions to a more general version of Shapley-Shubik model. As perfect assortative matching no longer holds, the standard identification argument in empirical auction, relating the observed bid to order statistics, no longer applies. An alternative estimation strategies will be developed in the subsequent section. ■

3 Estimation

In this section we consider the estimation of the sponsored search auction detailed above using data on the observed allocation as well as the per-click prices, conditional on click volumes, bidder-specific characteristics and position-specific characteristics. The structural model implies that the variation of allocation is driven by the variation of underlying valuation

V_{ij} across different keyword auctions. On the other hand, the variation of per-click price is driven by both valuation and click volumes. As there exist multiple equilibrium per-click prices, this extra layer of uncertainty will also contribute to the likelihood of per-click prices.

We use a Bayesian approach to estimate this model by explicitly specifying an equilibrium selection. An important virtue of the Bayesian approach is the use of data augmentation to deal with the many latent variables in this model (ie. bidders' valuations V_{ij}), which is generally difficult to handle under frequentist framework. Another key benefit of our Bayesian estimator is that it does not require resolving the game for each draw of ϵ_{ij}. By contrast, the partial identification approach, would require solving the game for each draw of ϵ_{ij}.[12] Moreover, the estimation approach in Athey and Nekipelov (forthcoming), which is based on bidders' first-order optimality conditions, does not work here since the equilibrium bids are characterized by a set of inequalities rather than first-order conditions.

3.1 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 keyword auction to keyword auction, defining the random vector $\mathbf{p}$ and random matrix μ deserves additional care to avoid the 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 index system 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 per-click price $\mathbf{p}_t = (p_{1t}, \ldots, p_{N_t t})$. 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 $\mathbf{V}$ being 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, while here we focus on the form of the likelihood, which is given by

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

The likelihood for each keyword auction t can be further decomposed into

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

[12] We explain the detail in Appendix B

We derive the conditional likelihood of μ_t and $\mathbf{p}_t$ from, respectively, the primal and the dual assignment problem.

Conditional distribution of per-click price $\mathcal{L}_1$:

For the time being we will drop the index t for the exposition purpose. Conditional on μ, the set of stable per-click prices is a convex polyhedron defined by a set of linear inequalities derived from the 1. no-blocking-pair (NBP) condition, 2. reservation price, 3. individual rationality condition, and 4. the rules of generalized second price auction. Being explicitly conditional on (optimal) μ and $\mathbf{V}$, there is no need to find the optimal value for the objective function in (DP) anymore since it equals to $\sum_{i,j} \mu_{ij} V_{ij}$. One can then 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,

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

If the search engine has a policy of minimum per-click price $\underline{p}$, then the second set of inequalities are

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

where $A_2 = -\mathbf{I}_{N \times N}$ and $b_2 = -\underline{p}\mathbf{1}_{N \times 1}$. In this paper $\underline{p}$ will be set to zero.

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

$$\{\mathbf{p} | A_3 \mathbf{p} \leq b_3\},$$

where A_3 is the diagonal matrix with $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ on the main diagonal and $b_3 = (V_{11}, V_{22}, \ldots, V_{NN})'$. Finally, the fourth set of inequalities states that $p_1 > p_2 > \cdots > p_N$, corresponding to the 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,

$$A_4 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \text{ and } \quad b_4 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{10}$$

The set of stable per-click prices in the sponsored-search auction 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 $\mathbf{b} = (b_1'|b_2'|b_3'|b_4')'$. 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.

Clearly, the conditional distribution of per-click prices is a distribution defined on a convex polyhedron: $\mathcal{L}_1(\mathbf{p}_t|\mu_t,\theta,\mathbf{V}_t;\mathbf{X}_t,\mathbf{Z}_t,\alpha_t) = \mathcal{L}_1(\mathbf{p}_t|\mu_t,\mathbf{P}(\mathbf{V}_t,\alpha_t),\lambda)$. λ is the nuisance parameter that governs how $\mathbf{p}_t$ is selected from the set of stable per-click price $\mathbf{P}(\mathbf{V}_t,\alpha_t)$. In literature, there are two main streams to deal with models with multiple equilibria. One is the partial identification approach that does not require an explicit specification of equilibrium selection; e.g., Ciliberto and Tamer (2009). Another alternative, which we adopt here, is by specifying an equilibrium selection to complete the likelihood specification; e.g., Bajari, Hong and Ryan (2010). Specifying such distribution in our case is nontrivial, as the support $\mathbf{P}(\mathbf{V},\alpha)$ depends on the latent variables. We consider two parsimonious specifications: uniform distribution and truncated beta distribution. 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 $\{\mathbf{p} : A(\alpha_t)\mathbf{p} \leq \mathbf{b}(\mathbf{V}_t)\}$:

$$\mathcal{L}_1(\mathbf{p}_t|\mu_t,\mathbf{P}(\mathbf{V_t},\alpha_t),\lambda) = \begin{cases} \dfrac{1}{vol(\mathbf{P}(\mathbf{V}_t,\alpha_t))} & \text{if } A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t) \\ 0 & \text{otherwise.} \end{cases}$$

Consequently, there is no need to estimate the nuisance parameter λ under the uniformity assumption. 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 $\mathbf{P}(\mathbf{V}_t,\alpha_t)$:

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

where $B(a,b)$ is the Beta function. Under this specification the equilibrium selection parameter λ will be the shape parameters (a,b) in the beta distribution. We will assume uniform prior for λ: $\pi(\lambda) \propto \frac{1}{c}, 0 < c < \infty$.

Conditional distribution of allocation $\mathcal{L}_2$:

Regarding $\mathcal{L}_2$, once conditioning on $\mathbf{V}_t$, one can solve the unique optimal allocation in (P) and verify whether the resulting solution coincide with μ_t. This implies that either $\mathbf{V}_t$ can rationalize μ_t or not:

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

Note that, 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, there is no need to solve the primal problem either, and

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

3.2 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, Hoffman, and Newton (2008).[13]

4 Data

We obtained a one-month (2010 June.) sponsored-link auction data from Taobao.com, 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 the data source, advertisers in Taobao 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, Taobao applies standard GSPA without score weighting. The bidding environment of Taobao.com is therefore more closely related to the static model described in EOS and Varian (2006).

For each keyword, we observe four sets of variables: 1. endogenous variables, 2. bidders' characteristics (exogenous variables), 3. keyword-specific characteristics and 4. one-month aggregate click volume for each position. First, for the endogenous variables, we have the (one-month) average allocation data (the links between bidders' IDs and positions) as well as the (one-month) average per-click price. Second, regarding bidders' characteristics, in Taobao each bidder has a quality rating ranging from 1 to 20, with score 20 being the highest. In

[13]Logan et al consider the estimation problem of two-sided matching game *without* transfer. Their Gibbs sampler does not take into account multiple equilibria and hence cannot be directly applied here.

Taobao, score 16-20 is categorized as “Gold Crown”, 11-15 “Crown”, 6-10 “Diamond”, 1-5 “Heart”. The quality score of bidders in our data set ranges from score 3 to 15. Therefore, there is no Gold Crown seller in our data. For ease of interpretation, we will sometimes refer “Crown” as “High Quality”, “Diamond” as “Medium Quality”, and “Heart” as “Low Quality”. Third, for each keyword, we construct 3 dummies to describe its characteristics: 1. “Brand”, whether the keyword includes a brand name such as Nikon or Canon. 2. “Specific”, whether the keyword includes a specific model/series that exclusively refers to a specific product such as D300s or 500D. “Promotional”, whether the keyword includes promotional terms such as “cheap” or “sale”. Summary statistics for our data are given in Table (1).

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 type of bidders bid more aggressively for certain type of keywords? What is the sorting pattern between quality and rank? The contingency table of bidders’ quality and position ranks are summarized in Table (2). Overall, we observe positive assortative matching between quality and rank. However, it is a interesting fact that even medium quality bidders are highly likely to be sorted into the 5-th position. We further conditional on the dummies of keyword characteristics, and the resulting contingency tables remains similar.

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[14] 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 fewer click volumes (bottom-left of Figure 1). We also conditional on the crown dummy (bottom-right of Figure 1) but it is difficult to explain the resulting difference in boxplots across ranks. Next 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). Finally, crown type bidders tend to bid more aggressively (bottom-right of Figure 2).

4.1 Specification details

In the model presented before, the click volume for each position (α_{it}) is treated as a known number. In reality, bidders have to take into account uncertainty in the click volume. In

[14] 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.

our data, there are instances when the top positions may receive zero clicks, but still attract bidders. A simple and likely explanation is that bidders are uncertainty about the click volume at the time they are submitting their bids, and hence bid based on expectations about the click volume. To accommodate this, we estimate bidders' expectation using a simple regression model. We assume that the click volume of the i-th position at auction t, α_{it}, satisfies the following shifted log-normal process[15]:

$$\log(\alpha_{it} + 1) = \gamma_{i0} + \gamma_{i1} N_t + \gamma_{i2}\text{specific}_t + \gamma_{i3}\text{promotional}_t + \gamma_{i4}\text{brand}_t + \eta_{it},$$

where N_t is the number of available positions and η_{it} follows $\mathcal{N}(0, \sigma^2)$.[16] The unknown parameters can then be estimated by OLS, and the expected click volume is given by

$$E[\alpha_{it}] = \exp(\gamma_{i0} + \gamma_{i1} N_t + \gamma_{i2}\text{specific}_t + \gamma_{i3}\text{promotional}_t + \gamma_{i4}\text{branded}_t + \sigma_i^2/2) - 1.$$

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

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

The first specification of the valuation of bidder i obtaining the j-the position in auction t is given by

Model I

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

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. In the second specification, we impose no assumption on how valuations decay across positions:

[15] An important restriction of the assignment game framework is that we cannot allow the click volume to be bidder-specific, otherwise the price system will become bidder-specific. We will return to this point when discussing the counterfactual simulations below.

[16] Similar specification is also adopted by Yang, Lu and Lu (2013), but here we do not impose that α_i decays from high to low positions.

Model II

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

For the top 5 positions, the regression coefficient is now position-specific. We set the deterministic component to 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.

The appendix contains a discussion of parameter identification for this specification of preferences.

5 Estimation Results

Table (3) summarize the results of our first model specification. Since the results are similar using both the uniform and beta equilibrium equilibrium selection rules, we will focus on the uniform selection results in the discussion. First, we find positive assortative matching pattern between bidders' quality and ad positions. The posterior mean of coefficients of crown and diamond are all positive (β_4 and β_5), and distinct from zero (larger than 3 times of standard deviation). Furthermore, the magnitude of the coefficient of crown is larger than that of diamon, implying that the top quality bidders have strongest incentive to bid aggressively. Second, this positive assortative matching pattern is further amplified in auctions containing product-specific keywords (see β_6 and β_7). 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. Third, the coefficients of assortative matching (β_8, β_9) for promotional keywords are negative, and sometimes are insignificant. This estimation result suggest that adding marketing terms into keywords does not change bidders' preference over positional rankings. On the other hand, the coefficients of assortative matching (β_{10}, β_{11}) for keywords containing brand name are strongly negative. Their posterior mean are far below zero (more than 3 standard deviations). Their magnitude would offset the positive assortative matching found in β_4 and β_5. This finding suggests that high quality online (camera) sellers are less interested in bidding keywords containing brand names, and therefore low quality sellers have a higher chance to win such auctions.

These estimation results may be explained by heterogeneity across consumers who search using these different types of keywords. For cameras, using product-specific keywords essentially result in the most narrow range of search results. Major camera manufacturers usually use unique product names (or model numbers) to distinguish their products from the others. For example, Nikon's DSLR (digital single-lens reflexive) camera models typically start will "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.[17] As different brand names are unambiguously linked to unique product names, entering the product-specific search string "D300s" will yield the same search result as appending a brand name ("Nikon D300s"). Hence, shoppers who use product-specific keywords are those who already 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 sellers are willing to engage in aggressive bidding to get exposed to these more valuable customers.

In contrast, shoppers who use brand name to search may be more interested in browsing and collecting information about different camera models; searching for, say, "Nikon" will return a wide varieties of models and accessories across many price points. These consumers have a more muted intention of purchasing, and our results imply that high quality online sellers are less interested in targeting this group of consumers. Finally, the design of Taobao 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 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 sellers.

Table (4) and (5) summarize the estimation results of model II. Again, we focus on the case of uniform selection rule. We find qualitatively similar assortative matching pattern compared with model I. The decreasing β_{1j} and β_{2j} implies that the valuations of both high and medium quality bidders decrease with position ranks. Again, the product-specific keywords reinforce such preference. The estimation results for keywords containing brand names and promotional terms also echo with the results of model I. Quantitatively, Table (4) does uncover some interesting findings. In some cases, it is the second position that generates the largest value for high quality sellers, not the first position; for specific keywords, as an example, crown and diamond sellers have the highest coefficient for position 2. This phenomena may be related to the empirical fact we found in the click volume data. Quite often it is the second position that generates the largest click volume, even after the regression smoothing. From sellers' 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), the sponsored-link auction also plays the role of information intermediary. If the links are sorted according to sellers' quality, then it allows the consumer to efficiently search for sellers who fit their quality needs. If the valuation matrix is positive assortative between quality and ranking, then high quality sellers will bid aggressively and hence GSP is an efficient way to convey information to online shoppers.

On the other hand, as we noted before, Taobao sellers typically post their prices on the sponsored ads, thus providing consumers with the most important piece of purchase-related

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

information without requiring further clicking behavior. In this setting, consumer search may be less relevant and, hence, the click volume in Taobao is less regular than described in the literature motivated from Yaoo!, Google and Microsoft; e.g., Varian (2006), Athey and Nekipelov (forthcoming).[18]

Finally, it is also interesting to note that the posterior mean of the shape parameters of beta equilibrium selection in both models implies that the selection density is left-skewed. This suggests that the observed prices are better explained as optimal for Taobao, rather than for the sellers; that is, the prices are on the higher end within the spectrum of equilibrium prices.

6 Counterfactuals: Scoring and Platform Revenues

We consider a counterfactual whereby Taobao implements 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 bidder bid times a "popularity score" for this bidder. Specifically, letting $\kappa_i, i = 1, \ldots, N$ denote the bidder-specific popularity scores, and 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 (note that 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 sellers to improve their quality. Intuitively, offering price discount may reduce platform's revenue. 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 the resulting game cannot be formulated as an assignment game a là 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 the 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[19]:

1. $\kappa_{\mu^{-1}(1)} b_1 > \kappa_{\mu^{-1}(2)} b_2 > \cdots > \kappa_{\mu^{-1}(N+1)} b_{N+1}$ (Allocation Rule)

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

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

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: $p_{\mu^{-1}(i)j} = \frac{\kappa_{\mu^{-1}(j+1)}}{\kappa_{\mu^{-1}(i)}} b_{j+1}$

To solve for the set of SNE in the WGSPA, we consider each possible allocation in turn; for each candidate allocation, we use the above inequalities 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 $B_\mu \equiv \{\mathbf{b} | D_\mu(\alpha, \kappa)\mathbf{b} \leq c_\mu(\mathbf{V})\}$.

Sources at Taobao tell us that one of the key element in determining κ_i in Taobao is the historical performance and click volume of the advertiser i, which is highly correlated with Taobao's own quality rating system (see section 4). Therefore, we construct the counterfactual analysis by using two quality score system in Taobao: (1) the coarser one with $\kappa_i \in \{1, 2, 3\} \equiv \{\text{heart,diamond,crown}\}$; and (2) the finer one with $\kappa_i \in \{3, 4, \ldots, 15\}$.[20]

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[21] and simulate b_i from the equilibrium polyhedron of bids B_μ[22] according to uniform distribution or the estimated beta distribution. Finally, we apply the scored pricing rule in Eq. (13) to obtain the per-click price $\tilde{\mathbf{p}}_t$. 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. The summary statistics of the simulated (cross-sectional) per-click price distribution is summarized in table (6) and (7).

The two scoring systems deliver almost identical results. The per-click price for the 1st position increases dramatically, while the rest of the positions only decrease slightly.[23] For example, under Model I-uniform, on average the per-click price for the 1st position increases by 12 dollars, while the price for the second position only drops by 3 dollars. As the first two positions generates most of the clicks, WGSPA may even boost platform's revenue, thus justifying Taobao's move towards such a rule.

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

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

[22]There is no need to simulate the maximum bid b_1 for the following two reasons: First, the per-click prices are irrelevant of b_1. Second, if $(b_2, \ldots, b_{N+1})$ satisfy all the inequalities, one can always choose b_1 large enough to meet the allocation rule.

[23]The price of the 5th position drops most. It is because we assume the 6th position is assigned to the lowest type, and hence the decreased price is more or less assumed.

These results suggest that WGSPA essentially makes the high quality bidders compete more aggressively, and hence increase the degree of assortative matching between quality and position ranks. In ongoing work, we are exploring the simulation results to better understand the mechanism underlying our results.

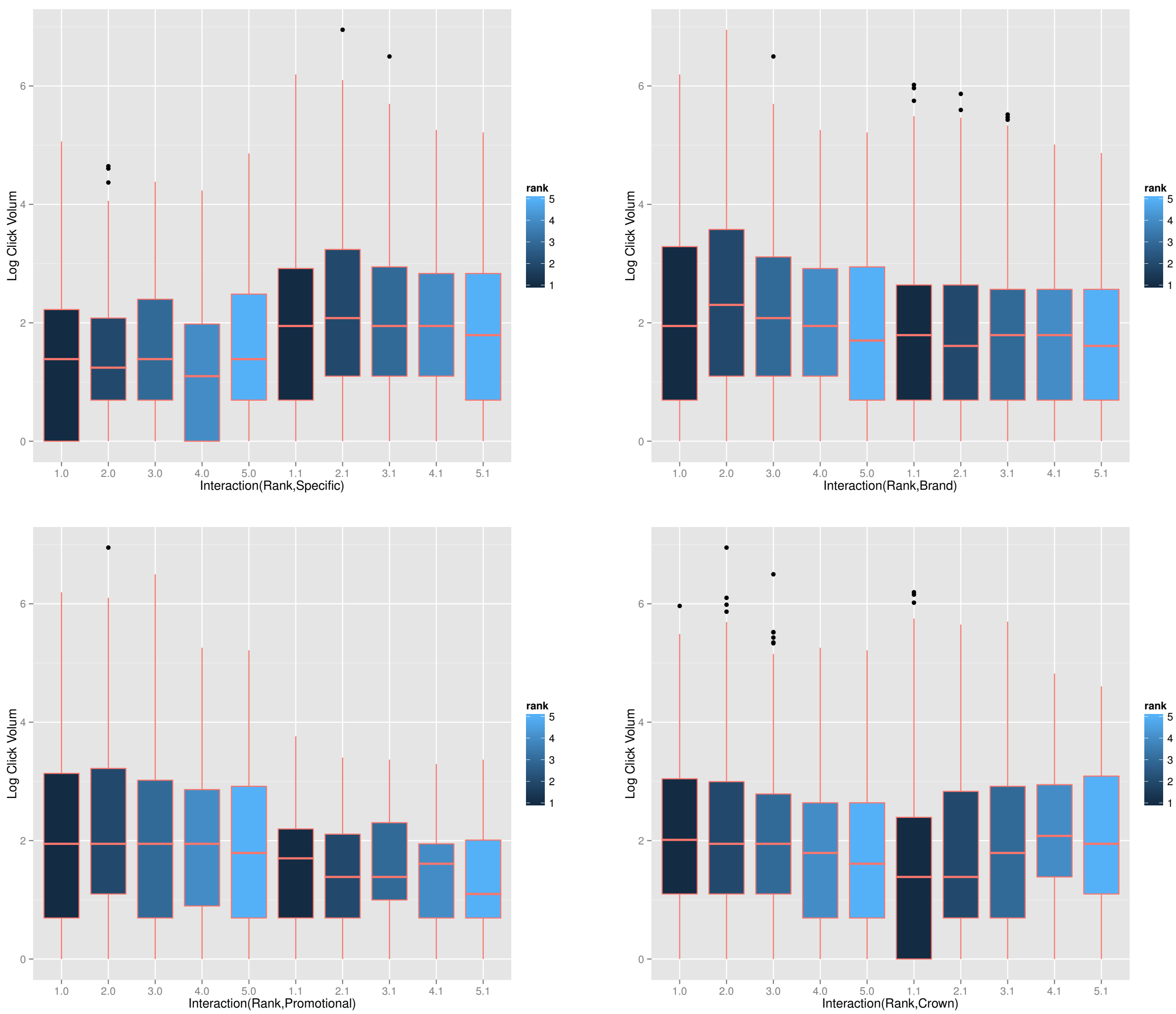
7 Conclusion

We contribute to the literature of empirical GSPA both methodologically and empirically. Our structural model is motivated by the equivalence result between the symmetric Nash equilibrium in GSPA, and the stability condition in the assignment game. The later model is easier to analyze and derive the likelihood function thanks to the duality theorem. For estimation, we develop a Metropolis-Hastings within Gibbs sampler, which does not require solving the aucton game for each draw of latent valuation. Our model is also more general than the previous literature, as we allow more general specification for valuation and click volumes. Empirically, we find that even though the WGSPA offers price discounts to high quality sellers, moving from the GSPA to the WGSPA has increased per-click prices for the top positions, which may result in higher revenue for the auction platform.

It is possible to extend our estimation procedure to a more general setup, such as asymmetric Nash equilibria studied in Börgers et al (2012), or bidder-specific click volumes. We leave these for future inquiry.

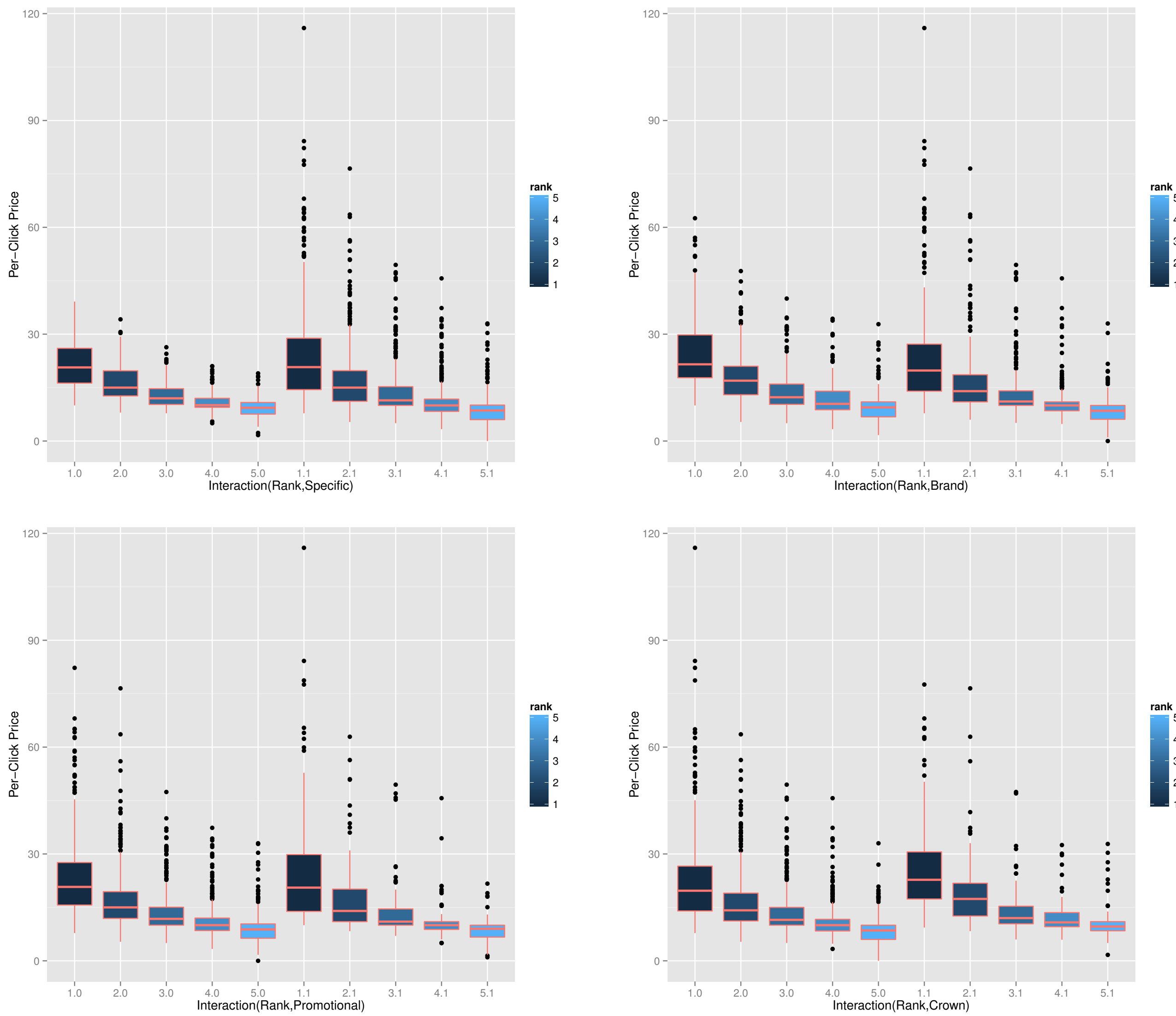
Graphs

Figure 1: Boxplot of Log Click Volume



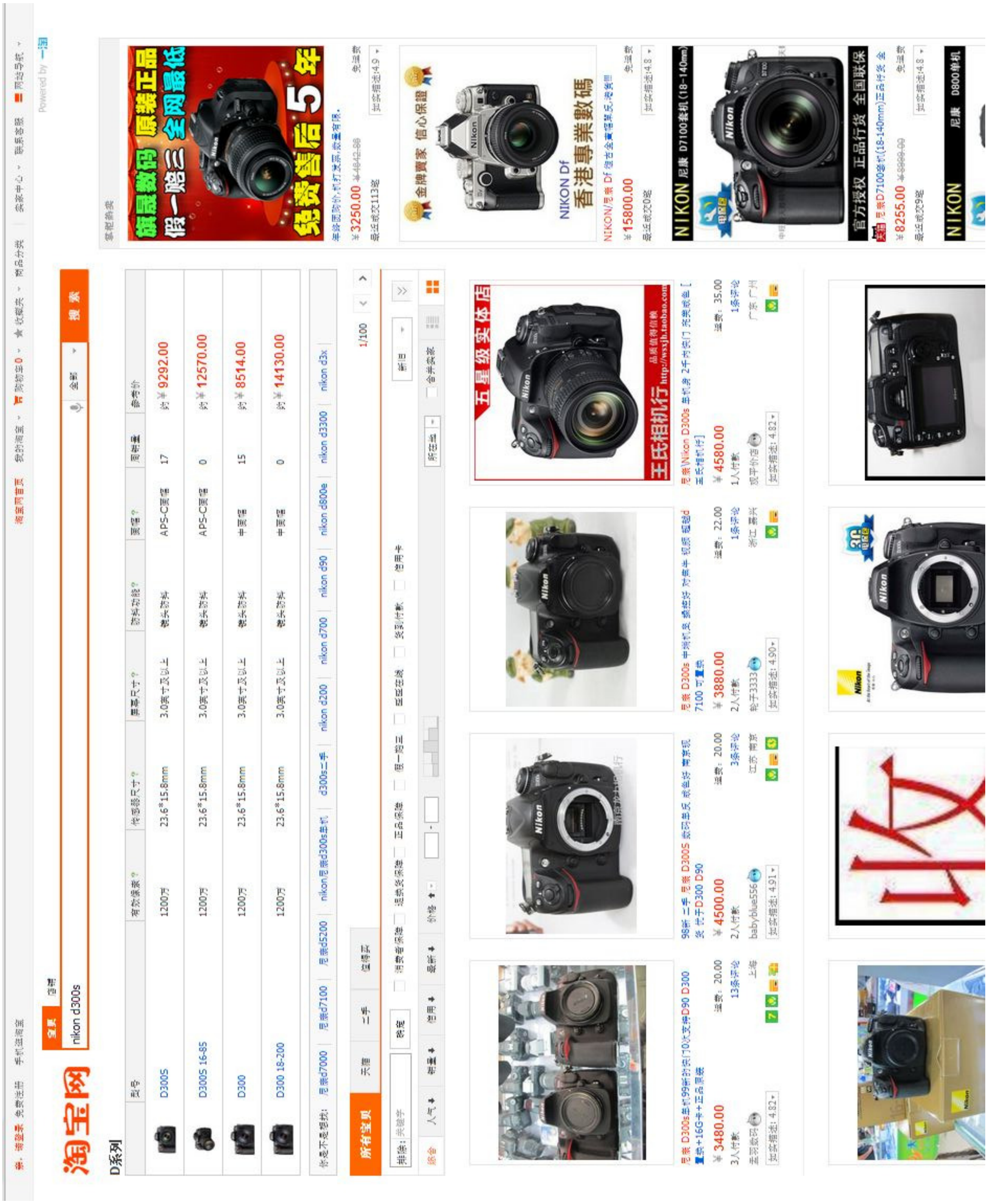
*The pair of numbers on the horizontal axis represents the values of two variables. For example, Interaction(Rank,Specific)= 3.1 on the x-axis of the top-left graph means “3rd position, Specific dummy = 1”

Figure 2: Boxplot of Per-Click Price



*The pair of numbers on the horizontal axis represents the values of two variables. 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 Taobao.com



Tables

Table 1: Variable List and Summary Statistics

Variable	Definition	Sample Mean
Keyword Characteristics		
Brand	= 1 if keyword includes the brand name	0.63 (0.48)
Specific	= 1 if keyword includes a specific model number	0.77 (0.42)
Promotional	= 1 if keyword includes promotional terms	0.18 (0.39)
Bidder Characteristics		
Crown	= 1 for high quality seller	0.2 (0.03)
Diamond	= 1 for medium quality seller	0.74 (0.03)
Auction Characteristics		
Per-Click Price		13.25 (9.25)
Click Volume		16.37 (42.98)

*Standard error in parentheses.

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

All Keywords					
	Position 1	2	3	4	5
Crown	0.072	0.049	0.037	0.038	0.026
Diamond	0.126	0.146	0.156	0.149	0.161
Heart	0.002	0.005	0.007	0.013	0.013
Keywords with Specific= 1					
	Position 1	2	3	4	5
Crown	0.071	0.047	0.037	0.041	0.029
Diamond	0.127	0.147	0.158	0.150	0.162
Heart	0.002	0.006	0.005	0.009	0.009
Keywords with Brand= 1					
	Position 1	2	3	4	5
Crown	0.075	0.046	0.037	0.037	0.022
Diamond	0.123	0.147	0.155	0.150	0.166
Heart	0.002	0.007	0.007	0.014	0.012
Keywords with Promotional= 1					
	Position 1	2	3	4	5
Crown	0.066	0.034	0.032	0.027	0.020
Diamond	0.132	0.161	0.166	0.159	0.170
Heart	0.002	0.005	0.002	0.014	0.009

*Crown, Diamond and Heart are respectively high, medium and low quality ranking in Taobao.com

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)
crown	mean	688.96	704.04
	s.d.	(19.41)	(26.86)
diamond	mean	479.75	563.92
	s.d.	(12.75)	(16.26)
crown×specific	mean	539.45	382.34
	s.d.	(21.65)	(21.84)
diamond×specific	mean	516.41	422.27
	s.d.	(16.60)	(15.21)
crown×promotional	mean	-42.53	-139.99
	s.d.	(23.43)	(38.72)
diamond×promotional	mean	-161.82	-86.42
	s.d.	(20.68)	(26.22)
crown×brand	mean	-341.05	-307.86
	s.d.	(18.08)	(25.68)
diamond×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
crown	mean	95.83	-27.87	-360.55	-431.62	-463.62
	s.d.	(15.34)	(26.93)	(16.00)	(22.21)	(24.45)
diamond	mean	-175.62	-140.20	-281.94	-424.62	-359.59
	s.d.	(26.16)	(23.98)	(19.24)	(31.72)	(25.66)
crown×specific	mean	16.39	230.99	-125.74	-157.83	-294.49
	s.d.	(34.01)	(28.67)	(38.07)	(44.10)	(52.92)
diamond×specific	mean	-15.96	124.32	-190.90	-236.87	-332.36
	s.d.	(24.92)	(24.99)	(28.15)	(13.95)	(23.43)
crown×promotional	mean	-356.05	-336.30	-153.56	-241.04	-78.22
	s.d.	(22.89)	(26.42)	(25.57)	(33.76)	(21.71)
diamond×promotional	mean	-288.32	-324.80	-264.19	-305.36	-313.11
	s.d.	(21.04)	(14.11)	(18.97)	(31.80)	(21.93)
crown×brand	mean	-109.84	-454.61	-47.33	-5.62	-16.25
	s.d.	(18.72)	(27.74)	(30.65)	(27.37)	(37.19)
diamond×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
crown	mean	19.82	-102.09	-400.17	-483.75	-522.48
	s.d.	(19.67)	(28.49)	(17.07)	(25.14)	(24.15)
diamond	mean	-229.56	-186.11	-338.50	-479.67	-416.46
	s.d.	(15.04)	(15.60)	(20.58)	(14.98)	(20.22)
crown×specific	mean	-217.75	-15.15	-361.31	-401.81	-490.52
	s.d.	(24.10)	(54.45)	(52.33)	(31.95)	(25.76)
diamond×specific	mean	-247.19	-125.35	-426.39	-471.28	-565.03
	s.d.	(44.74)	(35.69)	(50.40)	(40.62)	(29.00)
crown×promotional	mean	-334.95	-363.30	-179.40	-174.20	-124.02
	s.d.	(34.52)	(40.65)	(30.01)	(26.91)	(57.01)
diamond×promotional	mean	-323.63	-351.31	-304.38	-320.22	-345.52
	s.d.	(31.75)	(31.43)	(27.12)	(27.11)	(19.50)
crown×brand	mean	-189.93	-495.68	-134.79	-71.69	-101.03
	s.d.	(20.33)	(20.66)	(16.70)	(24.47)	(23.43)
diamond×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: The counterfactual per-click price of Symmetric Nash Equilibrium of WGSPA: The Case of Finer Score

		Rank of Positions				
Data	Statistics	1	2	3	4	5
	mean	22.32	16.13	12.68	10.27	8.12
	median	19.6	13.67	11	9.74	7.73
	variance	139.62	62.8	32.15	18.95	12.16
Model I-Uniform	mean	35.17	13.02	8.12	5.10	2.74
	median	34.55	11.78	7.45	4.34	2.01
	variance	118.22	46.53	24.59	15.02	6.71
	$[E(\underline{p}), E(\bar{p})]$	$[27.51, 41.02]$	$[7.18, 16.81]$	$[3.14, 11.56]$	$[1.15, 8.72]$	$[0.28, 6.68]$
Model I-Beta	mean	36.75	14.72	9.84	7.13	4.88
	median	36.01	13.30	8.98	6.28	3.92
	variance	139.84	62.22	35.26	25.25	16.64
	$[E(\underline{p}), E(\bar{p})]$	$[25.69, 39.42]$	$[6.55, 16.37]$	$[2.73, 11.21]$	$[1.02, 8.49]$	$[0.22, 6.43]$
Model II-Uniform	mean	28.32	17.89	10.37	6.00	2.92
	median	27.32	17.65	9.80	5.08	2.12
	variance	103.90	50.03	33.68	21.97	8.75
	$[E(\underline{p}), E(\bar{p})]$	$[19.38, 34.60]$	$[11.60, 22.25]$	$[4.18, 14.64]$	$[1.10, 10.49]$	$[0.10, 7.86]$
Model II-Beta	mean	30.29	19.42	12.14	8.03	5.19
	median	29.11	19.06	11.56	7.01	3.97
	variance	113.44	50.93	39.81	34.88	23.73
	$[E(\underline{p}), E(\bar{p})]$	$[18.56, 33.09]$	$[11.15, 21.28]$	$[3.9, 13.79]$	$[0.95, 9.61]$	$[0.07, 7.10]$

Table 7: The counterfactual per-click price of Symmetric Nash Equilibrium of WGSPA: The Case of Coarser Score

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

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.

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.

Bresnahan, T. (1981). Departures from Marginal Cost Pricing in the American Automobile Industry, *Journal of Econometrics*, **17**, 201-227.

Burkard, R., M. Dell'Amico and S. Martello (2009). *Assignment Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

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

Chiappoir, 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.

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

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

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.

Esteban, S. and M. Shum (2007). Durable Goods Oligopoly with Secondary Markets: the Case of Automobiles, *RAND Journal of Economics*, **38**, 332–354.

Fox, J. T. (2010). Identification in Matching Games, *Quantitative Economics*, **1**, 203–254.

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 B. Salanié (2010). Matching with Trade-offs: Revealed Preference over Competing Characteristics. *Working Paper.*

Ghose, A. and S. Yang (2009). An Empirical Analysis of Sponsored Search in Online Advertising, *Management Science*, **55**, 1605–1622.

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 Economics*, **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, manual script, 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.

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.

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

Robert, C. and G. Casella (2005). *Monte Carlo Statistical Methods*, Springer-Verlag.

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.

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

Yang, S., S. Lu and X. Lu (2013). Modeling Competition and Its Impact in Paid-Search Advertising, *Marketing Science*, forthcoming.

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. $\mathcal{N}(0, \sigma^2)$[24] 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[25]:

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

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.[26] 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 $p_0(\theta, \mathbf{V} | \mathbf{X}, \mathbf{Z}) = f(\mathbf{V} | \theta; \mathbf{X}, \mathbf{Z}) \pi(\theta)$.[27]

A.2 Algorithm

We propose a Metropolis-Hastings within Gibbs sampler (or Hybrid Gibbs sampler; see Robert and Casella, 2005) to draw $(\theta, \lambda, \mathbf{V})$ from the posterior. Gibb's sampling loops over the three conditional distributions $f_1(\mathbf{V} | \theta, \lambda, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha) = \Pi_{t=1}^T f_1(\mathbf{V}_t | \theta, \lambda, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)$, $f_2(\theta | \mathbf{V}, \lambda, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)$, 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 MH procedure too. Below is the implementation detail:

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

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

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

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

1. Conditional on $(\theta^{(\tau)}, \lambda^{(\tau)}, \mathbf{V}^{(\tau)})$, update $\mathbf{V}^{(\tau+1)}$

1.1 Simulate

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

We suggest the following steps to simulate $\tilde{V}_{ijt}$[28]: 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

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

where

$$\begin{aligned}\rho &= 1 \wedge \left[\frac{f_1(\tilde{\mathbf{V}}_t|\theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)}{f_1(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)}\right]\left[\frac{q_t(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)})}{q_t(\tilde{\mathbf{V}}_t|\theta^{(\tau)})}\right] \\ &= 1 \wedge \left[\frac{\mathcal{L}_1(\mathbf{p}_t|\tilde{\mathbf{V}}_t, \theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}{\mathcal{L}_1(\mathbf{p}_t|\mathbf{V}_t^{(\tau)}, \theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}\right]\left[\frac{q_t(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)})}{q_t(\tilde{\mathbf{V}}_t|\theta^{(\tau)})}\right] \\ &= 1 \wedge \left[\frac{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\tilde{\mathbf{V}}_t, \alpha_t), \lambda^{(\tau)}, \mu_t) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\mathbf{V}_t^{(\tau)}, \alpha_t), \lambda^{(\tau)}, \mu_t) f(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)}\right]\left[\frac{\mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\mathbf{V}_t^{(\tau)})) f(\mathbf{V}_t^{(\tau)}|\theta^{(\tau)}; \mathbf{X}_t, \mathbf{Z}_t)}{\mathbb{1}(A(\alpha_t)\mathbf{p}_t \leq \mathbf{b}(\tilde{\mathbf{V}}_t)) f(\tilde{\mathbf{V}}_t|\theta^{(\tau)}; \mathbf{X}_t, \mathbf{Z}_t)}\right] \\ &= 1 \wedge \frac{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\tilde{\mathbf{V}}_t, \alpha_t), \lambda^{(\tau)}, \mu_t)}{\mathcal{L}_1(\mathbf{p}_t|\mathbf{P}(\mathbf{V}_t^{(\tau)}, \alpha_t), \lambda^{(\tau)}, \mu_t)}.\end{aligned}$$

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.

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

2. Conditional on $(\theta^{(\tau)}, \lambda^{(\tau)}, \mathbf{V}^{(\tau+1)})$, update $\theta^{(\tau+1)}$

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 $\theta_i^{(\tau)} \pm a_i$. This is also know as the random-walk proposal density, which is 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

$$\begin{aligned} \rho &= 1 \wedge \frac{f_2(\tilde{\theta}|\mathbf{V}^{(\tau+1)}, \lambda^{(\tau)}, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)}{f_2(\theta^{(\tau)}|\mathbf{V}^{(\tau+1)}, \lambda^{(\tau)}, \mu, \mathbf{p}, \mathbf{X}, \mathbf{Z}, \alpha)} \\ &= 1 \wedge \frac{\Pi_{t=1}^T f_1(\mathbf{V}_t^{(\tau+1)}|\tilde{\theta}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)\pi(\tilde{\theta})}{\Pi_{t=1}^T f_1(\mathbf{V}_t^{(\tau+1)}|\theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{p}_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t)\pi(\theta^{(\tau)})} \\ &= 1 \wedge \frac{\Pi_{t=1}^T \mathcal{L}_1(\mathbf{p}_t|\mathbf{V}_t^{(\tau+1)}, \tilde{\theta}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\mathbf{V}_t^{(\tau+1)}|\tilde{\theta}, \mathbf{X}_t, \mathbf{Z}_t)}{\Pi_{t=1}^T \mathcal{L}_1(\mathbf{p}_t|\mathbf{V}_t^{(\tau+1)}, \theta^{(\tau)}, \lambda^{(\tau)}, \mu_t, \mathbf{X}_t, \mathbf{Z}_t, \alpha_t) f(\mathbf{V}_t^{(\tau+1)}|\theta^{(\tau)}, \mathbf{X}_t, \mathbf{Z}_t)} \frac{\pi(\tilde{\theta})}{\pi(\theta^{(\tau)})} \end{aligned}$$

Because $\mathcal{L}_1$ only depends on $(\mathbf{V}_t, \alpha_t, \lambda)$ and by construction $\mathbf{V}_t^{(\tau+1)} \in \text{GSP}$, 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 $\lambda^{(\tau+1)}$

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 Parameter Setup

For the first specification of bidder valuations, 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),[29]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 the second specification, 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.3.2 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)

[29]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}.

uniform random numbers from the smallest bounding box of $\mathbf{P}(\mathbf{V}, \alpha)$. Second, 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.3.3 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[30], 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.

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, 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

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

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.

It is possible to derive a bound for the structural parameters.

Theorem 1. *Suppose the econometrician observes T independent auctions indexed by $t = 1, 2, \dots, T$. 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 $\mathbf{P}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta)$. $\mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta) = \Pi_{i=1}^{N}[p_{il}, p_{iu}]$, where (p_{il}, p_{iu}) are respectively the smallest and largest elements of the i-th coordinate of $\mathbf{P}$. Define two random vectors $\bar{\mathbf{p}} = \max \mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta) = (p_{1u}, p_{2u}, \dots, p_{Nu})$ and $\underline{\mathbf{p}} = \min \mathbf{B}(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta) = (p_{1l}, p_{2l}, \dots, p_{Nl})$. The identified set of θ is given by*

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

Proof. Conditional on $(\mathbf{X}, \mathbf{Z})$, the (joint) distribution function of $\mathbf{p_o}$, $F(\mathbf{p_o}|\mathbf{X}, \mathbf{Z})$, is identified by the sampling process. Given $(\mathbf{X}, \mathbf{Z}, \alpha, \epsilon_{ij}; \theta)$, $\mathbf{p_o}$ is a measurable selection from the set of stable prices $\mathbf{P}$. By construction, $\underline{\mathbf{p}} \leq \mathbf{p_o} \leq \bar{\mathbf{p}}$, for almost-all ϵ_{ij}, where the inequality are defined componentwise. This implies that the distributions of $\underline{p}$, p_0, 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 assume ad hoc equilibrium selection, it is extremely computationally demanding to compute such bound since for each draw of latent valuation matrix, one has to solve (DP). If one follow the inference method of Chernozhukov, Hong and Tamer (2007), the total number of game solving is at least equal to the number of grid points of θ times the number of keyword auctions (= 487) times the number of $\mathbf{V}$ draws. What even worse is that, depending on $c(\theta; \mathbf{X}_t, \mathbf{Z}_t)$, one has to discard many $\mathbf{V}$ draws that do not satisfy the GSP restriction. Consequently, one would need to draw more latent variables to achieve the desired accuracy. By contrast, the Bayesian approach does not require repeated game solving step, and there is no waste of the simulated latent variables.