NONPARAMETRIC ESTIMATION OF SPONSORED SEARCH AUCTIONS AND IMPACTS OF AD QUALITY ON SEARCH REVENUE

DONGWOO KIM

Department of Economics, Simon Fraser University

PALLAVI PAL

School of Business, Stevens Institute of Technology

This paper presents an empirical model of sponsored search auctions in which advertisers are ranked by bid and ad quality. We introduce a new nonparametric estimator for the advertiser's valuation and its distribution under the *'incomplete information'* assumption. The ad value is characterized by a tractable analytical solution given observed auction parameters. Using Yahoo! search auction data, we estimate value distributions and study the bidding behavior across product categories. We find that advertisers shade their bids more when facing less competition. We also conduct counterfactual analysis to evaluate the impact of score squashing (ad quality raised to power $\theta < 1$) on the auctioneer's revenue. Our results show that product-specific score squashing can enhance auctioneer revenue at the expense of advertiser profit and consumer welfare.

KEYWORDS: Sponsored search links, generalized second price auction, incomplete information, nonparametric estimation, bid shading, score squashing.

Dongwoo Kim: dongwook@sfu.ca

Pallavi Pal: ppal2@stevens.edu

We are grateful to Jeitschko D. Thomas, Kyoo il Kim, Hanzhe Zhang, Roger J. Calantone, Michael Conlin, Jay Pil Choi, Chenhui Guo, Ashish Agarwal, David Croson, Vivek Bhattacharya, Arthur Lewbel, and seminar participants at Michigan State University, Midwest Econometrics Group Meeting 2019, American Marketing Association conference 2019, Boston College, Grand Valley State University, and Cemmap/SNU Advances in Econometrics Workshop 2022 for useful discussions and suggestions. We would also like to thank Yahoo! Labs for providing the dataset. Kim gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada under the Insight Development Grant (430-2022-00841).

1. INTRODUCTION

Online advertisements, the primary source of revenue for the majority of digital platforms, have become a ubiquitous aspect of daily life. A vast amount of digital ads are processed daily on various online platforms, including search engines, social media, and video-sharing services. According to Statista, the projected global spending on digital advertisements in 2022 was $615.9 billion.[1] The largest segment of online advertising is represented by search advertisements, which are sponsored search links that appear among the standard search results on platforms such as Google, Yahoo!, Bing and Amazon.

Sponsored search links are priced and sold through an auction mechanism. To maximize auction revenue, search engines must comprehend how advertisers respond to changes in the auction design. Counterfactual analysis is a widely used approach to study advertiser behavior. However, it is imperative to first estimate the advertiser's unobserved valuation to conduct such an analysis. The auction literature has traditionally used the equilibrium bid to estimate ad value, with a primary focus on information assumptions that impact the robustness of ad value estimates and counterfactual predictions. This paper proposes a novel methodology to estimate advertisers' values under realistic information assumptions, accommodating features such as a large number of advertisers, uncertainty about value, and quality score. We also evaluate finite sample performances of our estimator through simulations, apply it to Yahoo! historical data, and investigate the impact of changes in the auction mechanism using counterfactual analysis.

The search engine uses a weighted Generalized Second Price (GSP^w) auction to price the sponsored search links. The ad positions are assigned in descending order of the weighted bids, where the weight captures the advertiser's click probability referred to as the "quality score". We begin by setting up an empirical model of GSP^w auctions, which is based on the assumption of *incomplete information*. Under this assumption, the advertiser only possesses knowledge of the distributions of the values and quality scores of other advertisers. This i.i.d. private value framework, an important and canonical setting in auction theory, allows for the handling of both value and score uncertainty. Given the high number

[1] https://www.statista.com/outlook/dmo/digital-advertising/worldwide#ad-spending

of competitors in this market, it is unrealistic for an advertiser to have exact knowledge of the bid and score for each one of his competitors. Thus, we assume that the advertiser does not know who is participating in the auction but is aware of the distributions of the opponents' values and scores. Alternatively, another realistic scenario can be the case of advertisers forming beliefs about the weighted bid distribution through repeated interactions, rather than knowing the primitive value and score distributions. We demonstrate that our main results are valid under either of these assumptions, making the method applicable under more realistic assumptions on information available to the advertisers.

We derive the unique symmetric efficient Bayes Nash equilibrium in the GSP^w auction by incorporating the weights (quality scores) in the simpler GSP auction model proposed in Gomes and Sweeney (2014) (GS14 hereafter). This extension proposes a way to deal with the advertiser's multi-dimensional type which occurs due to the introduction of weights in the GSP auction. As highlighted in GS14, the extension of private information to a multi-dimensional setting is a crucial and challenging step forward. The existing literature in auction theory has centered on deriving the equilibrium bid based on the advertiser's valuation and its distribution, which are unknown to the econometrician. The equilibrium bid function in the GSP^w auction not only requires knowledge of the valuation and its distribution but also is analytically challenging to invert. We overcome this challenge by establishing nonparametric identification of the advertiser's valuation and its distribution given observed bids and auction characteristics. Our identification result is crucial for empirical analysis as it allows us to determine the latent values based only on the observables.

We also propose nonparametric methods to estimate the valuation and its distribution and density functions. Our methodology aligns with that of Guerre et al. (2000) as it relies on the observed bid distribution, which can lead to possible extensions of our methods to asymmetric bidders and random reserve prices. In contrast to the estimator proposed by Guerre et al. (2000), our estimator does not involve nonparametric density estimation of the observed bids and thus does not face the boundary problem. Therefore, our estimator enjoys favorable convergence rates.

Through simulation studies, we validate the efficacy of the proposed framework in accurately recovering the latent valuation and its distribution. The estimator demonstrates desirable finite sample properties, ensuring precise estimation of the underlying value distribution with appropriate sample size. Our findings indicate that bidders tend to reduce their bids when the competition is lower and that this bid shading behavior decreases with an increase in the number of participants. Additionally, our results suggest that bidders with higher weighted values possess a greater capacity to shade their bids.

The proposed method is applied to online search auction data obtained from Yahoo! Labs. The dataset encompasses approximately 78 million observations, covering all advertisements displayed on Yahoo! search result pages over four months, across five main categories: *cruise travel, car insurance, laptops, cable TV*, and *collectible coins*. The data does not contain information on the position-specific click rates and quality scores of advertisers. Thus, our estimation process involves two phases. In the first phase, we estimate the position-specific click-through rates and quality scores using a linear probability model (LPM). In the second phase, we estimate the latent valuations given the estimates from the first phase and the observed bids. We then compute the empirical distribution of estimated values.

Our findings suggest that the advertisers with higher weighted values benefit more from the GSP^w auction as they can shade their bids more effectively. Since the bids in the data were re-scaled using an unknown factor to conceal the actual amounts, we focus more on the bid shading amount as a percentage of the advertiser's value. In all categories, the median bid shading percentage was close to 0, but the bid shading percentage increased significantly across higher percentiles. *Car insurance* and *cable TV* exhibited significantly less bid shading compared to other categories. The level of competition and position-specific click-through rates are closely linked to bid shading behaviors. *Car insurance* had a higher number of advertisers competing in the category, which limited the extent to which advertisers could shade their bids. *Cable TV* had a similar number of advertisers as *laptop*, but the position-specific click-through rate in this category decreased more sharply across positions, reducing the incentives for advertisers to shade their bids.

Furthermore, we investigate the incentive for a search engine to utilize score squashing through counterfactual analysis. Score squashing is a method for altering the relative importance of quality weights by raising the quality score to the power of $\theta \in [0, 1]$. Despite its potential impact on auction outcomes, prior literature has limited empirical evidence on the effects of score squashing. This paper conducts counterfactual experiments to determine the optimal squashing level and provides empirical evidence that the search engine can improve revenue through score squashing at the cost of advertiser profit and consumer welfare. We find that the impact of score squashing on the auctioneer's revenue is contingent upon market characteristics, with a tendency towards a higher value of the squashing parameter in markets with a more competitive environment.

Finally, the current methodology is further discussed concerning special features of the online auction market, including limited bid data, which is addressed in the Appendix A. We show that the identification of the advertiser's value can be achieved through observation of a single bid order statistic, thereby mitigating the issue of limited bid data, which has been known to cause bias as identified in Athey and Nekipelov (2010) (AN10 hereafter). Additionally, the extension of our framework to incorporate a reserve price is proposed.

The remainder of the paper is structured as follows. In the remainder of this section, we discuss the related literature. Section 2 gives an overview of the auction model, equilibrium analysis, and identification of valuations. Section 3 provides simulation experiments on generated auction data. Section 4 describes the data. Section 5 explains the estimation procedure. Section 6 presents the empirical results. Section 7 conducts counterfactual experiments. Finally, Section 8 summarizes the findings and discusses the broader consequences of the paper. Further discussions, additional empirical results, all theoretical proofs, and more details on data are provided in appendices in the Online Supplementary Material.

1.1. *Related Literature*

This paper contributes to the literature on the identification and estimation of advertisers' latent valuations in online search auctions. The early works of Edelman et al. (2007) (referred to as EOS) and Varian (2007) were pioneers in deriving ex-post Nash equilibria in GSP auctions under the assumption of "perfect information," where advertisers possess knowledge of their opponents' bids and quality scores.[2] Börgers et al. (2013) extended this analysis by estimating position-specific valuations. The subsequent literature has further investigated the theoretical aspects of GSP auctions under various information structures and model setups, including the Bayes-Nash equilibrium derived by GS14 under the incomplete information assumption, the impact of consumer search on auction design explored by Athey and Ellison (2011), and the bounding inefficiencies in GSP auctions analyzed by Caragiannis et al. (2015). Sun et al. (2014) derived the optimal reserve price in the GSP$^{\omega}$ auction but did not formally derive the equilibrium.

In the empirical literature, our study is broadly related to works that address issues related to GSP auctions, such as Decarolis and Rovigatti (2021), Svitak et al. (2021), Yao and Mela (2011), and Ghose and Yang (2009).[3] Specifically, our paper is closely related to papers that look at the estimation of advertiser's value in the GSP auctions. Empirical studies on estimating value in the weighted GSP auctions face several challenges, including multi-dimensional equilibrium bids, analytically difficult equilibrium bid equations, unobserved non-winning bids, and unknown finite sample properties of previously proposed estimators. As a result, empirical works in this area have been limited. Börgers et al. (2013) and Hsieh et al. (2015) look at estimating value distribution in the non-weighted GSP auction using perfect information setup derived in Varian (2007).[4]

---

[2]EOS and Edelman and Schwarz (2010) compare the static incomplete information GSP auction to a dynamic version of the auction.

[3]Decarolis and Rovigatti (2021) studied the impact of intermediary firms in the online search ads market on search engines' profit using average bid per keyword data. Svitak et al. (2021) investigated the product price effect of restrictions on online search ads through "non-brand bidding agreements." Ghose and Yang (2009) analyzed how keyword search and position characteristics impact auction results using bidding data from a large retail chain. Yao and Mela (2011) considered the GSP auction in a dynamic setting, but did not explicitly model the bidding behavior of advertisers.

[4]Hsieh et al. (2015) estimated a structural model to understand the impact of score weighting in the GSP auction when the click rate is not decreasing with ad-position using Chinese online auction data.

To the best of our knowledge, the only other work that estimates value in the GSP$^{\omega}$ auctions using equilibrium bidding behavior is AN10, which derived a unique ex-post Nash equilibrium under market uncertainty assuming that the advertiser knows the competitors' bids. Our approach differs from the earlier work in that it is based on a different set of assumptions, incorporating uncertainty in values and quality scores, and yielding a unique Bayes-Nash equilibrium. Additionally, our method can be extended to cases of limited data and binding reserve prices, which are not addressed by AN10, and our estimator can be more easily computed with a tractable analytic solution without a tuning parameter. In contrast, AN10's valuation estimator relies on a numerical derivative formula with a tuning parameter τ_N affecting the convergence rate, which follows the standard nonparametric rate $\sqrt{N\tau_N}$, while ours follows the $\sqrt{N}$ rate.

There have been experimental studies as well looking at the advertiser's bidding behavior. An early field experiment was analyzed in Ostrovsky and Schwarz (2011) to understand the gains from optimal reserve prices in GSP auctions. Two more recent experimental studies, Bae and Kagel (2019) and Che et al. (2017), were conducted to analyze bidding behaviors in sponsored search auctions, comparing bidding behaviors in static complete information GSP auctions and dynamic incomplete information auctions. Both studies found that the equilibrium in the static complete information case deviates from the Vickrey–Clark–Groves (VCG) equilibrium. These findings further highlight the importance of studying the equilibrium bid in the incomplete information setting considered in our paper.

Our study relates to the computer science literature and builds upon previous works. In particular, Mohri and Medina (2015) proposed an algorithm to estimate the reserve price, which relies on the equilibrium conditions in GS14 and is based on the bid distribution and other observables. However, we show that GS14 cannot be directly applied in weighted GSP auctions.[5] We present a formal proof of the equilibrium in the weighted GSP auction,

[5] According to Mohri and Medina (2015), the bid (b) in the weighted GSP auction is equal to the equilibrium bid in the non-weighted GSP auction divided by the quality score ($\beta(v)/s$). However, this statement contradicts the results from Sun et al. (2014), who showed that the bid in the weighted GSP auction is equal to the weighted

and derive a new estimator that is more tractable than the one proposed by Mohri and Medina (2015) as it does not require nonparametric density estimation.

2. AUCTION MODEL

In this section, we introduce a model of sponsored search auctions that is amenable to empirical analysis and assumes a standard setting of incomplete information. The search ad market is described in Subsection 2.1, which outlines both the consumer and advertiser perspectives. The consumer model develops the click-through rates and quality score, which are employed in the advertiser's bidding strategy. In Subsection 2.2, the equilibrium for the GSP^w auction is derived and nonparametric identification of the valuation and its distribution is established.

2.1. *Market Environment*

Consumer side

Each consumer, denoted by i, has a unit demand for a product or service and initiates their search by submitting a query through an online search engine. Once the results page appears with links related to the search query, the consumer decides whether to click on any of the relevant links and make a purchase from one of the selected links. The consumer considers the anticipated benefit from clicking on the ad, which is determined by the ad's visible attributes and its position on the results page.[6] It is worth noting that each click incurs a substantial search cost in terms of time spent by the consumer. Further details regarding the variables impacting the consumer's utility are discussed in the empirical section.

Let $U_{i,j}$ denote the expected utility derived by consumer i from clicking on ad j. If the consumer chooses not to click on the ad, she allocates her time to an outside good, resulting

value divided by the quality score ($\beta(v \times s)/s$). Furthermore, Sun et al. (2014) did not formally prove that the derived bid is the equilibrium bid in the weighted GSP auction.

[6]Other factors may influence the click decision such as the characteristics of generic links on the page, but their effect on consumer click behavior cannot be determined due to limitations in available data.

in a normalized utility of $\bar{U}_{i,j} \equiv 0$. In the equilibrium, consumers essentially click on all ad links as long as the benefit of a click outweighs the search cost.[7] Let $y^*_{i,j}$ denote the binary indicator capturing consumer i's click decision for ad j as follows:

$$y^*_{i,j} = \begin{cases} 1 & \text{consumer } i \text{ clicks on ad } j \text{ if } U_{i,j} > 0 \\ 0 & \text{if } U_{i,j} \leq 0 \end{cases}$$

The above equation is used to estimate the predicted click probability for ad j in ad-position k. We assume that the ad-specific and position-specific effects are multiplicatively separable. This assumption was adopted by other papers in the literature for identification of the quality of advertiser (for instance, see Varian (2007) and AN10). Thus, for all ad j and position k, the click probability can be written as:

$$\text{Click Probability} = s_j \times c_k \tag{1}$$

where s_j is the effect of advertisement j (so called quality score) and c_k is the effect of ad-position k on the click probability, ceteris paribus. The set of click-through rates for all ad positions on the result page is denoted as $\mathcal{C} = \{c_1, ...c_K\}$. The click probability in Equation (1) is employed in the advertiser's profit maximization problem.

Advertiser side

Each advertiser, denoted by $j \in \mathcal{J} := \{1, \cdots, J\}$, places a single ad on the search engine, and the subscript j is used interchangeably for the advertiser and the ad. We assume each ad is separately optimized following AN10. The advertisements being sold in these auctions are contingent, meaning that advertisers only pay for their ad positions if a consumer clicks on the advertisement. To simplify notation, all ad-related terms such as bids, valuations, and prices are defined on a per-click basis and will be referred to as "value", "bid", and "price"

[7] Although most papers in the literature assume a single click per page, multiple clicks per consumer is more realistic in this market. This is evident in a new feature on Bing, which gives an option of opening a new tab every time you click on a link (See https://searchengineland.com/bing-is-testing-an-open-in-new-window-icon-in-the-search-results-301922).

in the remainder of this article. Advertiser j has an ad value, v_j, drawn independently from a common distribution F_v supported on $[\underline{v}, \bar{v}]$. In addition, each advertiser has a quality score, s_j, drawn independently from a common distribution F_s with bounded support on $[\underline{s}, \bar{s}]$. We further assume that v_j and s_j are independent. Advertiser j's type is given by (v_j, s_j). We define the *weighted value* as the product of the ad value and the quality score of the advertiser denoted as $\omega_j \equiv (s_j \times v_j)$, $\omega \sim F_w$. The potential number of advertisers is denoted by N, and the ads are sold through an auction mechanism.

Auction Setup

A GSP^w is held for each search query to sell ad positions on search results pages. Under the standard assumption of symmetric and independent private values, there are K ad positions to be auctioned and N potential advertisers participating in each auction. The bidding strategy for advertiser j is defined by a bid function $b_j \equiv b(v_j, s_j)$, where bj denotes the bid submitted by j based on his value, v_j, and quality score, s_j. To account for the effect of click probability on revenue, the auctioneer employs weighted bids instead of the original bids. The weighting of bids reflects the impact of the advertiser on the click rate, which is captured by the quality score, s_j. The weighted bid for advertiser j is expressed as $b_w(v_j, s_j) \equiv b_{j,w} = b_j \times s_j$.

Let G_b and G_w represent the distributions of the original bids and weighted bids, respectively. The k^{th} highest order statistic of the weighted bid is denoted as $b_w^{[k]}$. The ad positions are assigned in descending order of weighted bids, meaning that the k^{th} ad position is awarded to the advertiser with the k^{th} highest weighted bid. The price paid by the advertiser who wins the k^{th} position is equal to the $[k+1]^{th}$ highest weighted bid divided by his quality score.[8] The price for the same ad position may vary across advertisers. The rules of the auction can be summarized as follows:

$$k^{th} \text{ position allotted to } j \text{ if } \quad b_w^{[k+1]} \leq b_{j,w} \leq b_w^{[k-1]}$$

[8]For simplicity, it is assumed that there is no reserve price. Yahoo! had a fixed reserve price during the data period but the reserve price was not reported in the data.

$$\text{price for } k^{th} \text{ position paid by } j: \quad p_{k,j} = \frac{b_w^{[k+1]}}{s_j}.$$

This auction design is a generalization of the second-price auction. It is important to note, however, that in the GSP^w auction, advertisers do not submit their values as bids, unlike in the second-price auction. This is because, in the GSP^w auction, bidding less than one's value can result in a strategic advantage, as the bid influences both the winning probability and the price paid. This incentive does not exist in the second-price auction, as the bid only affects the allocation and not the price.

In the following, we present the information setup and assumptions necessary to derive the equilibrium.

ASSUMPTION 1: *(Incomplete information) Each advertiser knows their type, (v, s), but does not know the opponents' bids, quality scores, and values. They only know the primitive value and quality score distributions, F_v and F_s. The number of advertisers (N), click rates across ad positions ($\mathcal{C}$) and the number of ads per page (K) are common knowledge.*

ASSUMPTION 2: *The advertiser's weighted bid in the GSP^w auction is strictly increasing in his weighted value.*

Assumption 1 describes the standard incomplete information setting. Assumption 2 is necessary to guarantee the existence of an efficient equilibrium. It is noteworthy that the validity of Assumption 2 can be verified after estimating the values.

Profit maximization problem

Denote advertiser j's profit from winning k^{th} position as $\pi_{k,j}$. This quantity can be expressed as:

$$\text{profit from } k^{th} \text{ position:} \quad \pi_{k,j} = \underbrace{(c_k \times s_j)}_{\substack{\text{Prob. of click} \\ \text{at position } k}} \times \underbrace{(v_j - p_{j,k})}_{\substack{\text{Per click profit} \\ \text{at position } k}}. \tag{2}$$

Since each advertiser submits a single bid to acquire one of the available ad positions on the search result page, the equilibrium bid maximizes the expected total profit, which is defined as the sum of the product of the profit for each ad position and the probability of winning that position.

$$\text{profit from the auction:}\quad \Pi(b_j; v_j, s_j) = \sum_{k=1}^{K} \underbrace{P(b_{j,w} = b_w^{[k]})}_{\text{Prob. of winning position } k} \underbrace{\mathbb{E}(\pi_{k,j} | b_{j,w} = b_w^{[k]})}_{\text{Profit from position } k}$$

Using Equation (2), we obtain

$$\Pi(b_j; v_j, s_j) = \sum_{k=1}^{K} P(b_{j,w} = b_w^{[k]})(c_k \times s_j) \left(v_j - \mathbb{E}[p_{j,k} | b_{j,w} = b_w^{[k]}] \right). \tag{3}$$

We are interested in "bid shading," which is defined as the difference between the advertiser's bid and their ad value.

$$\text{Bid Shading} \equiv v_j - b_j$$

2.2. *Identification Analysis*

The primary focus of this paper is to estimate the distribution of advertisers' valuations, denoted as F_v. This distribution conveys advertisers' willingness to pay for an advertisement and serves as a basis for counterfactual analysis. Given assumptions 1 to 2, we first demonstrate the existence of a unique efficient equilibrium in a GSP^w auction. The equilibrium bid is determined by maximizing the profit function in Equation (3).

$$b(v_j, s_j) = \arg\max_{\hat{b}} \sum_{k=1}^{K} P(\hat{b} \times s_j = b_w^{[k]})(c_k \times s_j) \left[v_j - \mathbb{E}\left(\frac{b_w^{[k+1]}}{s_j} \middle| b_w^{[k]} = \hat{b} \times s_j \right) \right] \tag{4}$$

In a standard auction, the valuation is identified by inverting the equilibrium bid function. However, in a GSP$^\omega$ auction, the situation is more complex as the weighted bid, $b_w(s_j, v_j)$, is multi-dimensional and depends on both v_j and s_j. As a result, it is not possible to simply invert the bidding function to identify the valuation. Therefore, deriving the Bayes Nash Equilibrium (BNE) in a GSP$^\omega$ auction is a significant extension of the standard BNE derivation in a GSP auction and is considered an important and challenging problem. This extension is discussed in GS14 as a step forward in the field.

To address the issue of the non-invertible weighted bid, we prove the equivalence between a GSP^w auction and a *non-weighted* GSP auction in which an advertiser's value is equal to the weighted value. Specifically, we establish that for any advertiser j, the equilibrium weighted bid in the GSP^w auction is equivalent to their equilibrium bid in the GSP auction, where their value has been substituted by the weighted value, ω_j. Thus, the bid in the GSP auction performs a similar function as the weighted bid in the GSP^w auction, with the difference being that the bid in the GSP auction can be represented as a univariate function as shown below.

$$b^{GSP}(\omega_j) \to \mathbb{R}_+, \text{ where } \omega_j \equiv s_j \times v_j.$$

In the following lemma, we formally demonstrate the equivalence between the weighted bid in GSP^w to the bid in GSP.

LEMMA 1: *The equilibrium weighted bid function in GSP^w auction, $b_w^{GSP^w}(v_j, s_j)$, is equivalent to the equilibrium bid function in a GSP auction, $b_w^{GSP}(\omega_j)$, where the value is replaced by the weighted value.*

$$b^{GSP}(\omega_j) = b_w^{GSP^w}(v_j, s_j), \quad \forall j \in \mathcal{J}$$

This lemma establishes that the weighted bid function can be expressed as a function of a single dimension, i.e. the advertiser's weighted value, ω_j. Therefore, at equilibrium, the weighted bid $b_w(v_j, s_j)$ can be represented as $b_w(\omega_j)$. This simplification greatly facilitates the proof of the existence of an efficient equilibrium and identification of the value distribution. As demonstrated below, under Assumption 2, the maximization problem can be modified by inverting the weighted bid in the probability of winning the position k and incorporating the quality score s_j inside the last bracket in Equation (4).

$$b_w(\omega_j) = \arg\max_{\hat{b}_w} \sum_{k=1}^{K} P(b_w^{-1}(\hat{b}_w) = \omega^{[k]}) c_k \left[\omega_j - \mathbb{E}\left(b_w^{[k+1]} \middle| b_w^{[k]} = \hat{b}_w \right) \right], \tag{5}$$

where $\omega^{[k]}$ denotes the k^{th} highest weighted value. We can now derive the equilibrium bidding strategy using Lemma 1 and Assumptions 1–2.

LEMMA 2: *Let Assumptions 1–2 hold. Then, the unique efficient symmetric Bayesian Nash equilibrium of the GSP$^\omega$ auction is given by the following bidding strategy for all $N \geq 2$:*

$$b_w(\omega) = \omega - \Gamma(\omega) - \sum_{n=1}^{\infty} \int_0^{\omega} M_n(\omega,t)\Gamma(t)dt, \quad \forall \omega \sim F_w(.), \tag{6}$$

where

$$\Gamma(\omega) = \frac{\sum_{k=1}^{K} c_k \binom{N-2}{k-1}(k-1)(1-F_w(\omega))^{k-2}\int_0^{\omega} F_w^{N-k}(x)dx}{\sum_{k=1}^{K} c_k \binom{N-2}{k-1}(1-F_w(\omega))^{k-1}F_w^{N-k-1}(\omega)},$$

$$M_1(\omega,t) = \frac{\sum_{k=1}^{K} c_k \binom{N-2}{k-1}(k-1)(1-F_w(\omega))^{k-2}F_w^{N-k-1}(t)f(t)}{\sum_{k=1}^{K} c_k \binom{N-2}{k-1}(1-F_w(\omega))^{k-1}F_w^{N-k-1}(\omega)},$$

$$M_n(\omega,t) = \int_0^{\omega} M_1(\omega,\varepsilon)M_{n-1}(\varepsilon,t)d\varepsilon, \text{ for } n \geq 2.$$

Lemma 2 provides a formula for calculating the equilibrium weighted bid based on the weighted value, as well as the distribution and density functions of the weighted values. However, the complexity of the functional form presents challenges in terms of determining the advertiser's valuation through inversion of the equilibrium bidding function, due to the presence of an infinite sum and multiple layers of integral. Furthermore, the distribution and density functions of the weighted bid are never known to the econometrician. To address these obstacles, an alternative method for deriving the advertiser's valuation is proposed that does not require inversion of the equilibrium bid, the solution of differential equations, or knowledge of the distribution and density functions. The following theorem demonstrates the identification of the advertiser's valuation.

THEOREM 1: *Under Assumptions 1–2, the advertiser's value, v, is identified by:*

$$v = b + \Phi(G_w, b, s|\mathcal{C}, K, N) \tag{7}$$

where

$$\Phi(G_w, b, s|\mathcal{C}, K, N) = \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-G_w(b_w))^{k-2} \int_0^{b_w} G_w(u)^{N-k} du}{s \sum_{k=1}^{K} c_k \binom{N-1}{k-1} G_w(b_w)^{N-k-1} (1-G_w(b_w))^{k-2} \Big[(N-k)(1-G_w(b_w)) - (k-1)G_w(b_w) \Big]}$$

given the quality score (s)*, the equilibrium bid* (b)*, the distribution function of equilibrium weighted bids* (G_w)*, the number of available ad positions* (K)*, click rates across ad positions* $(\mathcal{C})$*, and the number of advertisers* (N)*.*

The detailed proof is presented in Appendix B. A brief overview of the key insights underlying the theorem is given here. The proof employs an indirect approach, utilizing the Revelation Principle, to derive the equilibrium bid. As a result, the following equation for the equilibrium bid is established:

$$\sum_{k=1}^{K} c_k \omega \frac{d\xi(\omega)}{d\omega} = \sum_{k=1}^{K} c_k \binom{N-1}{k-1} b_w(\omega)(N-k) F_w(\omega)^{N-k-1} f_w(\omega)(1-F_w(\omega))^{k-1}$$
$$- \sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} f_w(\omega) \int_0^{\omega} b_w(x)(N-k) F_w(x)^{N-k-1} f_w(x) dx, \quad (8)$$

where $\xi(\omega) \equiv P(\omega^{[k+1]} \leq \omega \leq \omega^{[k-1]})$. We apply integration by parts and integration by substitution to the integral part of Equation (8). Then we replace $F_w(\omega)$ with the observed weighted bid distribution $G_w(b_w)$ using Assumption 2. By canceling out the density function on both sides and rearranging, the desired outcome is obtained.

Identification under an alternative assumption

In the context of online auctions, it may be more realistic to assume that advertisers, through repeated interactions, develop a belief about the distribution of weighted bids rather than the primitive weighted value distribution. In such a scenario, the assumption stated below can serve as a replacement for Assumption 1.

ASSUMPTION 3: *Each advertiser knows their type, (v, s), but does not know other advertisers' bids, quality scores, or values. They only know the primitive weighted bid distribution, G_w. The number of advertisers (N), click rates across ad positions (c_k), and the number of ads per page (K) are common knowledge.*

The following theorem demonstrates that utilizing Assumption 3 in place of Assumption 1 yields identical results in Theorem 1.

THEOREM 2: *Under Assumptions 2–3, the advertiser's equilibrium bid (b) satisfies the following equation:*

$$b = v - \Phi(G_w, b, s|\mathcal{C}, K, N) \tag{9}$$

given the quality score (s), the valuation (v), the distribution function of equilibrium weighted bids (G_w), the number of available ad positions (K), click rates across ad positions $(\mathcal{C})$, and the number of advertisers (N).

The term $\Phi(G_w, b, s|\mathcal{C}, K, N)$ represents the bid shading amount of the advertiser by definition. It can also be expressed in terms of the value percentage as follows:

$$\frac{v_j - b_j}{v_j} = \frac{\Phi(G_w, b_j, s_j|\mathcal{C}, K, N)}{v_j}.$$

While the bid shading amount is of interest, the focus is placed on the bid shading in relation to its value. It is important to note that the actual bids are re-scaled in the data, and therefore, the bid shading percentage provides more accurate information. In practice, the equilibrium weighted bid distribution $G_w(\cdot)$ is unknown and must be estimated using a consistent estimator, such as the empirical distribution based on the data. The integral component $\int_0^{b_{j,w}} G_w(u)^{N-k} du$ can be numerically integrated using trapezoidal sums.[9]

REMARK 1: *(Consistency)* The only statistical uncertainty in the estimated value comes from $\hat{G}_w$. The empirical distribution uniformly converges to the true distribution at the

[9] The 'trapz' command in MATLAB is used for numerical integration in our calculations.

$\sqrt{n}$-rate by the Glivenko-Cantelli Theorem. Given the uniform convergence, the integral part also converges to the true value. Therefore, the valuation is consistently estimated by the proposed estimator. Unlike the nonparametric estimator in Guerre et al. (2000), our estimator does not involve a nonparametric density estimator of observed bids. Therefore, we do not have to introduce trimming to correct the boundary problem. The distribution and density functions of the valuation can be estimated by the empirical distribution and a kernel density estimator. The estimator of the valuation converges to the true value as fast as the empirical distribution and much faster than the kernel density estimator so that the nonparametric distribution and density estimators enjoy the standard convergence rates.

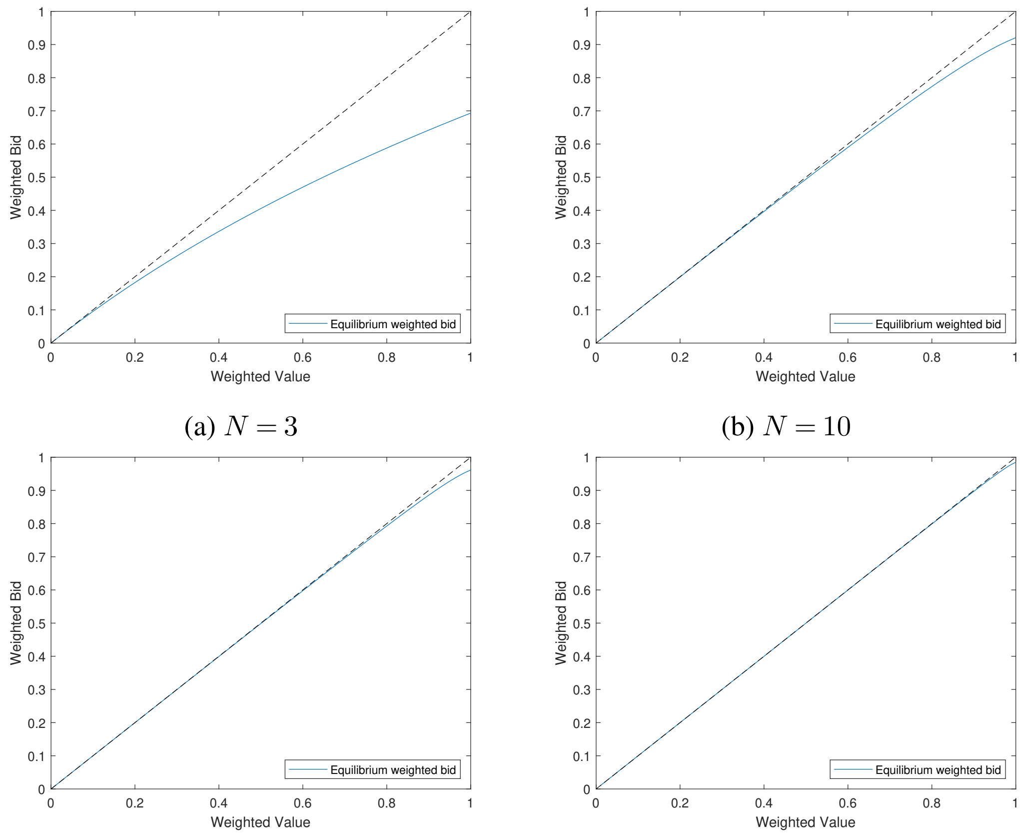
3. SIMULATIONS

We conduct simulation experiments on known data generating processes (DGPs) to examine how our estimator performs across varying numbers of bidders. To derive the analytical solution for the equilibrium bid, we focus on auctions with two ad positions. For $K = 2$ and arbitrary N, we obtain the following first-order differential equation:

We conduct simulations on a set of known data generating processes (DGPs) to assess the performance of our estimator under different numbers of advertisers. To derive a theoretical solution for the equilibrium bid, we focus specifically on the case of auctions with two ad positions. For $K = 2$ and an arbitrary N, we derive the following first-order differential equation:

$$b'_w(\omega) = \frac{F_w(\omega)(c_1 - c_2) + (N-2)(1 - F_w(\omega))c_2}{F_w(\omega)c_1 + (N-2)(1-F_w(\omega))c_2} + \frac{(N-2)f_w(\omega)\left[c_1 + \left((N-3)\left(\frac{1}{F_w(\omega)} - 1\right) - 2\right)c_2\right]}{F_w(\omega)c_1 + (N-2)(1-F_w(\omega))c_2}(\omega - b_w(\omega)). \quad (10)$$

We use this equation to calculate the equilibrium weighted bid given the advertiser's weighted value, click rates, and the distribution of weighted values. The solution to the equation depends on the parameters c_1, c_2, N, and F_w. For sample sizes $n = 300$ and $n = 1200$, we simulate the auctions n/N times with $N = 3, 10, 20$ and 50, resulting in a consistent number of observations for each value of N. We first draw the weighted value



(a) $N = 3$ (b) $N = 10$

(c) $N = 20$ (d) $N = 50$

FIGURE 1.—Equilibrium weighted bids given weighted values. The dashed line in each plot is a 45-degree line, which represents truthful bidding.

ω_j from the uniform distribution on the unit interval so that the weighted value distribution function is simplified to $F_w(\omega) = \omega$. Quality scores, s_j, are drawn from $Unif[0.5, 1]$. We set $c_1 = 1$ and $c_2 = 0.5$. We compute the weighted bids ($b_{j,w}$), and then estimate their values using the proposed method. This process is repeated 500 times for each combination of n and N.

Figure 1 displays the equilibrium weighted bid as a function of the weighted value. As can be seen, the number of advertisers significantly influences the amount of bid shading. As the number of advertisers participating in the auction increases, the bid shading decreases rapidly. A summary of the selected quantiles of bid shading percentage is presented in Table I. The median bid shading percentage is 18.9% when $N = 3$, and this reduces to

TABLE I

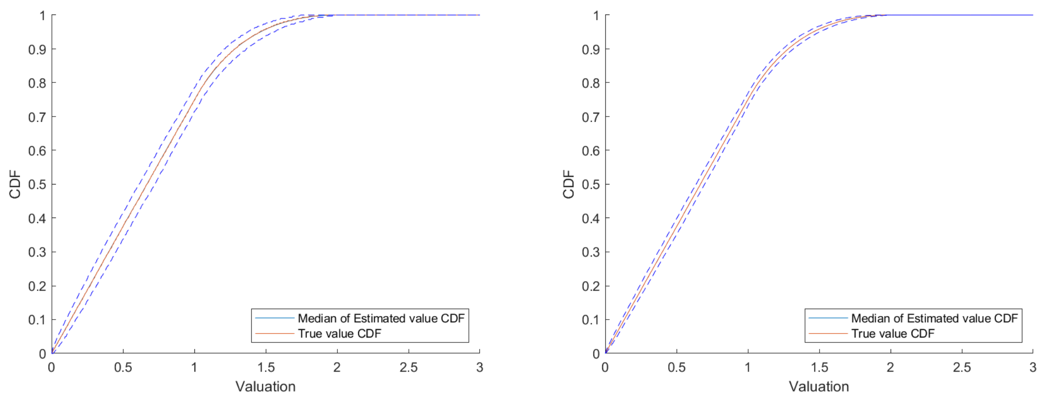
QUANTILES OF BID SHADING PERCENTAGE

Quantile	$N=3$	$N=10$	$N=20$	$N=50$
25%	10.734%	0.440%	0.095%	0.015%
50%	18.901%	1.202%	0.275%	0.042%
75%	25.389%	2.808%	0.732%	0.120%
90%	28.680%	4.970%	1.616%	0.316%
99%	30.495%	7.553%	3.445%	1.143%

1.2% when $N = 10$. For $N = 50$, even at the 90% percentile, the bid shading percentage is only 0.32%. The decrease in bid shading percentage with an increase in N is slower for higher quantiles. For example, when moving from $N = 3$ to $N = 50$, the 90% percentile bid shading decreased from 28.6% to 0.32%, while the 99% percentile bid shading decreased from 30.5% to 1.14%.

These findings suggest that the extent of bid shading is influenced by the number of participating advertisers. As N increases, bid shading approaches zero. This aligns with what is observed in other auction designs, such as the first price auction, where bid shading is known to be inversely related to the number of bidders, as described in detail in Krishna (2009), Ch 2.3. Additionally, we observe that the decrease in bid shading occurs more quickly for advertisers with lower weighted values compared to those with higher values. Despite having a high number of advertisers, non-zero bid shading persists for advertisers with high weighted values. This discrepancy between bid shading for high and low-valued advertisers may not be immediately evident but is rooted in the low probability of high-valued advertisers losing an auction. Since it is highly unlikely for a competitor to have a higher weighted value, the risk of not winning any ad positions from bidding low relative to their weighted value is minimal. Furthermore, this could potentially result in lower profits for the search engine for higher ad positions compared to lower ones.

The results of the simulation are presented in Figure 2, which displays the median, 5th and 95th percentiles of the estimated value distributions in comparison to the true distribu-



(a) $n = 300$ (b) $n = 1200$

FIGURE 2.—Finite sample performances. The dashed lines are the 5th and 95th percentiles of the estimated distributions across 500 replications.

tion for various sample sizes.[10] The median of the estimated distribution coincides with the true distribution. It also demonstrates the finite sample performances of the proposed estimator. The 5th and 95th percentiles are close to the true distribution even with a relatively small sample size $(n = 300)$.

These simulations indicate that the proposed estimator can effectively estimate the value and its distribution in a finite sample. The availability of extensive observations across repeated auctions in online auction data renders the proposed estimator capable of accurately estimating the unobserved values from bids. It is important to note that only two ad positions were considered in our simulation to make it computationally tractable. In actual search auctions, there are typically more than two ad positions, and if the click-through rates differ significantly across these positions, bid shading may be more pronounced than what is observed in the two-position case entertained here. We will further investigate the bid shading behavior using real-world data in the subsequent sections.

[10]The underlying value distribution remains constant regardless of the value of N, resulting in identical estimation results for different values of N.

4. DATA

For empirical analysis, we utilize a large data set obtained through the Yahoo Webscope Program.[11] The data covers all search queries over four months (January-April 2008) across five product categories: *Cruise, Car Insurance, Laptop, Cable TV,* and *Collectible Coins*. By using each category as a separate data set, we estimate the valuation distribution and bid shading for each category. This approach allows us to compare the underlying valuation distributions and bidding behaviors across different product characteristics and market environments.

The data include aggregated information on search keywords, bids, clicks, ad positions, and display frequencies. The keywords consist of a base category word along with one or more additional words.[12] The identity and keywords of the advertisers have been masked, though it is possible to track the same advertiser and keyword over time.[13] While the base categories are also anonymized, it is possible to make credible speculations about the categories based on their observed characteristics, as discussed in the Appendix F.

The raw data contains $77,850,272$ observations. To ensure accurate estimation, the data was restricted to ads displayed on the first two pages of search results, which aligns with the assumptions made in Agarwal and Mukhopadhyay (2016) and AN10. This restriction is based on the fact that the majority of clicks (93.8%) occur on the first page of search results, while click-through rates on subsequent pages are almost negligible. To eliminate

---

[11] Advertising & Markets Data, A3. Yahoo! Search Marketing Advertiser Bid-Impression-Click data on competing Keywords (version 1.0). The data set is not publicly available. However, researchers at accredited universities can request access to the data through Yahoo! Labs upon the Data Sharing Agreement.

[12] For instance, 'business laptop' and 'student laptop' fall under the base category of *laptop*. The utilization of additional words allows for more targeted advertisements, with an ad for 'business laptop' appealing to consumers specifically searching for business laptops and the keyword 'laptop' encompassing a broader search for any type of laptop requirement.

[13] Search queries that include specific brand names are removed from the data.

outliers, the data was further restricted to the month of February.[14] To identify the underlying value distribution and make it less likely for bids from multiple keywords to enter the same auction, the data was also restricted to the top 10% of the most popular keywords. This approach treats each category-day-keyword combination as an auction and assumes homogeneous value distribution across auctions. By focusing on the most popular keywords, we aim to rule out keywords with distinct value distributions and ensure that the most frequent searches are likely to share similar ad value.[15] Lastly, we drop keywords that did not have all ranks from 1 to 14 each day. The sample size of the restricted data remains substantial with $1,610,333$ observations.

The restricted data is separated into the consumer and advertiser sides. The data provider aggregated the information for each day-keyword-advertiser-position combination. In this context, advertisers specify a list of words related to the ad they are promoting. The auctioneer utilizes these keywords to match consumer search queries with the most appropriate ads. For instance, on January 1st, the data reports that the keyword "business laptop" specified by Amazon was displayed in the first position 100 times, resulting in 5 clicks from consumers. For the same keyword and advertiser combination, there is a separate observation for the ad displayed in the second position. This indicates that the data does not aggregate over different positions, but instead reports results for each position and keyword individually. This approach provides an advantage for the consumer side analysis as it allows for the examination of the advertiser-specific effect on the click probability and the position-specific click-through rates separately.

For the advertiser side, it is crucial to have information on how they optimize profits for each advertisement. In the decision-making process for setting a bid, advertisers consider

[14]This was done to remove high bid values placed by only 5 advertisers, which represented only 0.03% of the total number of advertisers in the sample. These high bid values were deemed outliers and their removal was necessary to obtain a more reasonable maximum bid compared to the average.

[15]For instance, in the *laptop* category, the search phrase 'high-end laptop' or 'expensive laptop' is associated with higher values and tends to be less sought-after. We try to rule out these keywords with likely distinct value distributions by focusing on the most frequent keywords. As the keywords are masked, it was not possible to remove heterogeneity across auctions or cluster keywords with similar values.

the expected click probabilities and prices for each keyword-based auction conducted over a day. As a result, the relevant variation in our data is the bid across keywords rather than across positions. Thus, the data on the advertiser side is aggregated by the combination of day, keyword, and advertiser. For instance, on January 1, the data reports that during an auction for the keyword “Best Laptop,” an ad by Amazon with a bid of 50 cents had an average position of 2.5, with 200 displays and 10 clicks.[16] The aggregation of data on the advertiser side aligns more closely with the profit-maximizing strategy of advertisers. The aggregated data enables us to measure the number of times an advertiser won an auction and their corresponding bid.[17]

The click rate is determined by the ratio of clicks to the number of displays for each ad. The average click rate across all ads is found to be 0.8%. The keyword associated with each ad provides insights into the type of search query that the ad is matched with longer keywords typically correspond to more specific search queries and are therefore considered to be more valuable for advertisers. This relationship has been previously documented in the literature (e.g. Ramaboa and Fish (2018)). We define the number of words in the keyword as keylength. There are 327 distinct keywords in the data, with a maximum of 5 words and an average of 2.4 words per keyword. To control for the influence of the popularity of a search query, we consider the popularity of keywords, measured as the daily average of the number of times an ad was matched using that keyword in the first position. On average, a single keyword is used to match 413 thousand ads per day. The descriptive statistics are reported in Table II.

[16]The bid is recorded in cents and is scaled by an unknown amount by Yahoo! for masking purposes. We re-scaled the bids by subtracting an amount close to the lowest value. The average bid in our data set is 2.3 cents.

[17]Some advertisers may have adjusted their bid throughout the day. To account for this, the bid for these advertisers was aggregated at the day level, representing 5.56% of the restricted data set. Additionally, there were 27 observations where the number of clicks exceeded the number of displays, which may be due to a recording error or the possibility of a consumer clicking on the same ad twice. The data does not provide sufficient information to thoroughly examine this behavior. This discrepancy would have impacted the calculation of the click rate, which is defined as the ratio of clicks to displays. To address this issue, the click rate was imputed as 1 if it exceeded 1.

TABLE II

DESCRIPTIVE STATISTICS

Variable	Cruise	Car Insurance	Laptop	Cable TV	Coin
Number of advertisers[1]	45	403	124	142	55
Keywords[2]	67	33	110	68	49
Keylength (mean)[3]	2.28	3.02	2.33	2.16	2.07
Keylength (max)	(4)	(5)	(4)	(4)	(4)
CTR (mean)[4]	0.87%	0.62%	0.99%	0.75%	0.94%
Bid (mean)[5]	1.15	6.94	0.83	0.90	0.46
Keyword popularity (mean)[6]	1355	9664	5514	1820	1857

Note: 1) The number of advertisers is computed as the maximum number of bidders across day-keyword combinations. 2) 'Keywords' means the number of the 10% most popular keywords in each category. 3) 'Keylength' is the average number of words across popular keywords. 4) 'CTR' is the average probability of being clicked. 5) 'keyword popularity' is the average popularity measure across popular keywords.

5. ESTIMATION PROCEDURE

This section describes the estimation procedure. We define the market, denoted by m, as the auction held for search terms related to a specific keyword within a day.[18] Thus, the variables are indexed by the additional dimension of the market, reflecting the fact that observations are collected from multiple markets (i.e. multiple auctions over time). The market is defined as the combination of category, day, and keyword. The estimation procedure involves two steps: 1) Estimating advertiser- and position-specific effects on click probability using information from the consumer side, and 2) Calculating the advertiser's value based on the observed bid using information from the advertiser side.

Step 1: Estimation of advertiser- and position-specific effects on click probability

The position-specific effect is employed as a metric to quantify the click-through rate (CTR) for each ad position. The advertiser's effect on the click probability is used to

[18] Our data is aggregated at keyword-day-advertiser-ad position level, thus we have to define the market at the day level. The aggregation is a data limitation not a limitation of the methodology. The suggested estimation method can also be applied to non-aggregated data if available.

compute the advertiser's quality score.[19] To analyze consumer click decisions, we use a linear probability model (LPM) and incorporate the aggregate CTR for each category-day-keyword-ad position-advertiser combination:

$$ctr_{k,j,m} = \beta_0 + \beta_1 z_m + \alpha_k + \gamma_j + \epsilon_{k,j,m} \tag{11}$$

where $ctr_{k,j,m}$ is the CTR, z_m are the market characteristics such as keyword popularity, the length of the search, and the weekend dummy, α_k is the ad-position fixed effect, γ_j is the advertiser fixed effect, and $\epsilon_{i,j,m}$ is the idiosyncratic shock. The parameters of the model are estimated using weighted least squares, where the weight is determined by the number of times the ad was displayed. This is because the click probabilities are more precisely estimated for more frequently displayed ads. Then, using the estimated parameters, two key parameters for the second step estimation are computed.

First, we compute the position-specific click rate ($\hat{c}_k$) using the estimated position-specific fixed effect, $\hat{\alpha}_k$. We normalize the click rates to be within the interval of $[0, 1]$ by subtracting the click rate of the 14^{th} position and dividing by the click rate of the first position in each category.

$$\hat{c}_k = \frac{\hat{\alpha}_k - \hat{\alpha}_{14}}{\hat{\alpha}_1 - \hat{\alpha}_{14}}$$

The estimated fixed effects for ad positions on the first page ($k = 1, \cdots, 7$) are highly significant (p-value < 0.000) but not significant for the second page ($k = 8, \cdots, 14$). As a result, we limit our analysis to only the first page of ad positions ($K = 7$), with the assumption that ad positions beyond the first page have negligible effects on advertiser payoffs. The estimated click rates across ad positions for each product category are presented in Table III. It was observed that there are substantial gaps between the click rates of the first and second positions in all categories. A decrease in click rate was also noted as the ad position became

[19]The quality score is computed and reported to the advertiser by the search engine. At the time of the data collection, the quality score captured the advertiser's impact on click probability and is equivalent to the estimated quality measure in this paper. It is worth noting that the current method of calculating the quality score, which is used by companies such as Yahoo!, Google, and Bing, includes the use of consumer characteristics and ad display characteristics.

lower, though the decline was not as steep as the drop between the first and second positions. In two categories, *car insurance* and *coin*, the decrease was not strictly monotonic. This pattern was consistent with the simple means of $ctr_{k,j,m}$ across ad positions.

Second, we determine the quality score ($\hat{s}_j$) using the estimated advertiser fixed effect, $\hat{\gamma}_j$. This quantity is then re-scaled to be between $[0, 1]$ by subtracting the minimum and dividing by the highest value.[20] The quantiles of the estimated quality scores are reported in Table IV.

TABLE III

POSITION SPECIFIC CLICK RATES ACROSS PRODUCT CATEGORIES

Ad-position	Cruise	Car Insurance	Laptop	Cable TV	Coins
2^{th}	0.602	0.633	0.671	0.457	0.739
3^{th}	0.474	0.494	0.554	0.403	0.627
4^{th}	0.231	0.330	0.505	0.242	0.733
5^{th}	0.167	0.271	0.413	0.187	0.591
6^{th}	0.099	0.328	0.314	0.153	0.526
7^{th}	0.083	0.267	0.274	0.128	0.316

TABLE IV

QUANTILES OF QUALITY SCORE ACROSS PRODUCT CATEGORIES

	Cruise	Car Insurance	Laptop	Cable TV	Coins
25%	0.038	0.022	0.071	0.189	0.040
50%	0.043	0.026	0.087	0.205	0.048
75%	0.047	0.032	0.103	0.222	0.062
90%	0.057	0.047	0.130	0.250	0.086
99%	0.119	0.084	0.234	0.389	0.249

[20] We transformed the minimum value to be equal to 0.001, as a quality score of zero implies that the advertiser had no motivation to bid.

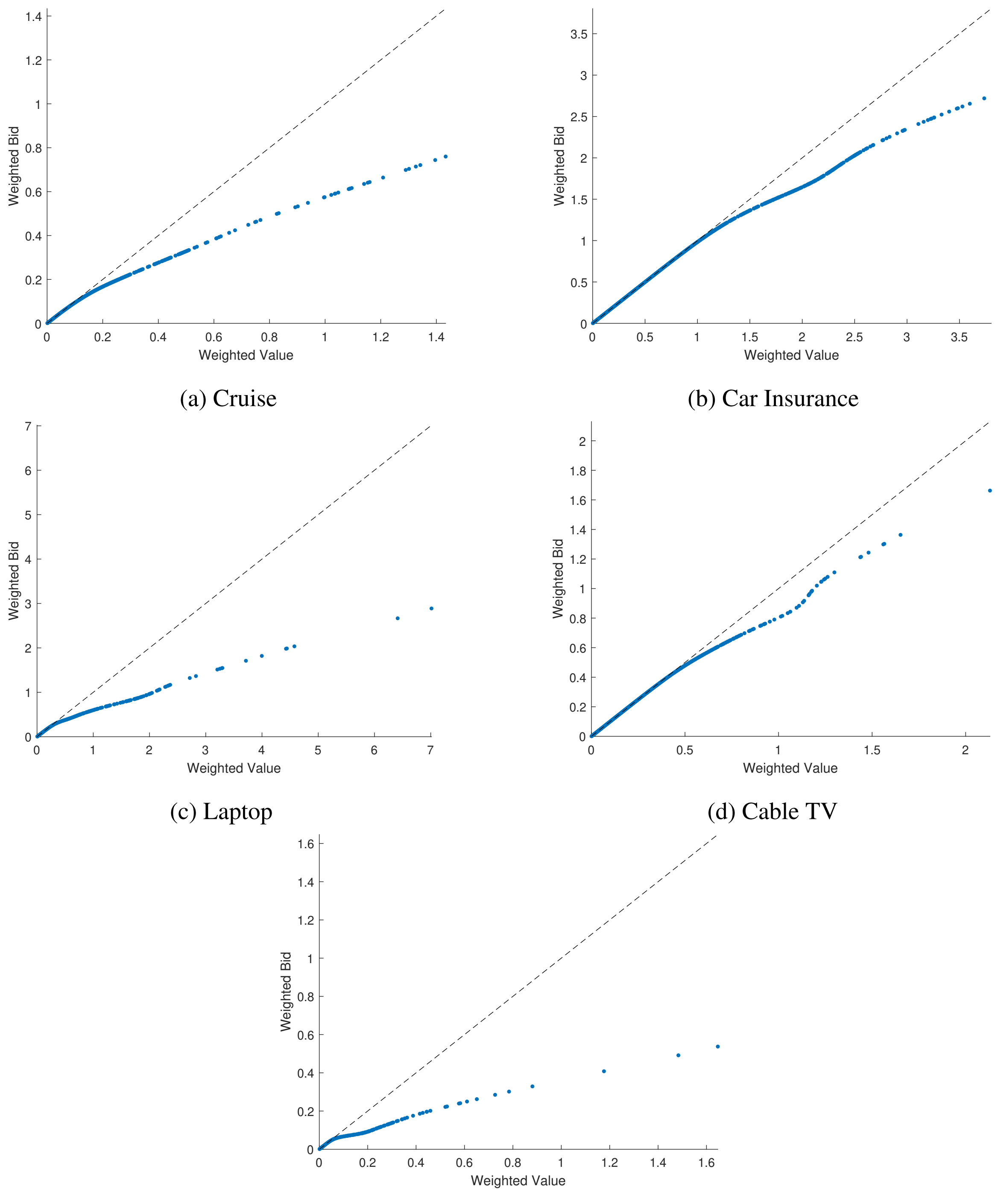
Step 2: Estimating the advertiser's value and its distribution

In this step, we obtain the weighted bids by multiplying the observed bids with the estimated quality scores ($\hat{b}_{j,w} = b_j \hat{s} j$). The distribution of weighted bids is then estimated by the empirical distribution ($\hat{G}_w$). The valuations (v_j) are computed using Equation (14) presented in Theorem 1. In this calculation, the unknown parameters G_w, $\mathcal{C}$, and s_j are replaced by their respective estimates $\hat{G}_w$, $\hat{\mathcal{C}}$, and $\hat{s}_j$. Finally, the distribution of values is estimated by the empirical distribution of the estimated values.

6. EMPIRICAL RESULTS

This section presents the estimation results of five product categories in the dataset. To evaluate the advertiser's benefit from an ad, we analyzed the estimated valuation and its distribution. The advertiser's maximum willingness to pay (value) for an ad is estimated first. The relationship between the weighted value and weighted bid is then verified by drawing the weighted bid as a function of the estimated weighted value in Figure 3. It is observed that for all product categories, the weighted bid is strictly monotone in the weighted value (Assumption 2 holds), thus ensuring the existence of a unique equilibrium in the analyzed auctions. The results show that the equilibrium weighted bidding function is close to the 45-degree line when the weighted values are low, but deviates significantly from the 45-degree line as the weighted value increases. The bid shading amount, as measured by the distance between the 45-degree line and the weighted bidding function, also increases monotonically with the weighted value. This implies that advertisers with higher weighted values can shade their bids more, thereby generating larger profits from sponsored search auctions.

Table V summarizes the quantiles of bid shading percentage across the product categories. The median bid shading percentage is low in all categories, with the highest median bid shading percentage being less than 0.6%. This indicates that a significant fraction of bid shading percentages are concentrated around 0. The extent to which advertisers can shade their bids varies across product categories, with advertisers in the *laptop*, *cruise*, and *coins* categories shading their bids more than their counterparts in the *car insurance* and *cable TV* categories.



(a) Cruise

(b) Car Insurance

(c) Laptop

(d) Cable TV

(e) Coins

FIGURE 3.—Relationship between weighted values and weighted bids. The dashed line is a 45-degree line which represents truthful bidding.

A larger dataset of product categories would have allowed for a more comprehensive analysis of the impact of different market characteristics on bid shading. With only five categories, however, we can still discuss some observed associations between market characteristics and bid shading. The number of advertisers and the position-specific click-through rates are two product market characteristics that are closely related to different bid shading behaviors. For example, the *car insurance* and *cruise* categories have similar quality score distributions as shown in Table IV, but the higher number of competing advertisers in *car insurance* results in less bid shading in that category compared to cruise. In contrast, while the *laptop* and *cable TV* categories have similar numbers of advertisers (124 and 142 respectively), the more dramatic drop in position-specific click rates in *cable TV* leads to less bid shading in that category compared to *laptop*. Similarly, the smaller decrease in click rates between positions in the *coins* category combined with the low number of advertisers (55) results in the highest bid shading percentages at the high percentiles among the product categories analyzed.

TABLE V

QUANTILES OF BID SHADING PERCENTAGE ACROSS PRODUCT CATEGORY

	Cruise	Car Insurance	Laptop	Cable TV	Coins
25%	0.161%	0.001%	0.012%	0.013%	0.093%
50%	0.568%	0.005%	0.054%	0.030%	0.401%
75%	4.065%	0.033%	0.388%	0.188%	1.827%
90%	10.051%	0.210%	3.247%	1.602%	12.444%
99%	36.606%	16.922%	39.099%	17.515%	54.158%

We also analyze the bid shading percentages across average ad positions in Table VI.[21] A similar exercise is also conducted in AN10. Compared to theirs, our bid shading percentages are much smaller across all the positions. Only the *coins* category shows similar

[21]We calculate the average ad position obtained by each bid and compute the mean bid shading percentages across different ranges of average positions.

bid shading percentages to their *phrase #1* (reported in AN10' Table 4) except the top and the bottom positions. Furthermore, their bid shading percentages are not monotonically decreasing across ad-positions, whereas ours drop monotonically. We demonstrate that the bid shading behavior can drastically vary with product characteristics. Our results are not directly comparable to theirs because AN10 used a different model and data.

TABLE VI

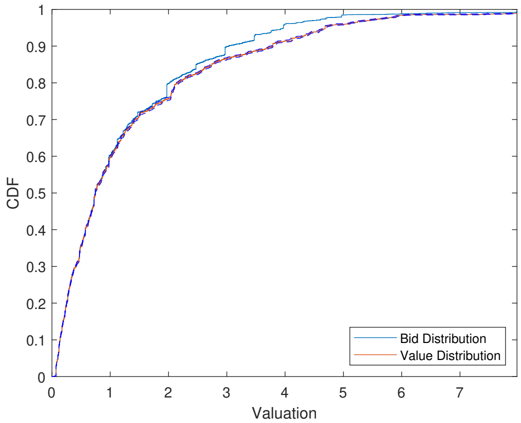
BID SHADING ACROSS AVERAGE RANKS

Average	Cruise	Car Insurance	Laptop	Cable	Coin
$[1,2)$	12.622%	2.213%	12.994%	7.307%	14.111%
$[2,3)$	7.485%	2.148%	6.935%	4.066%	11.538%
$[3,4)$	7.005%	1.409%	5.313%	2.576%	11.875%
$[4,5)$	5.776%	0.589%	3.460%	1.941%	8.768%
$[5,6)$	4.363%	0.263%	2.114%	1.619%	5.450%
$[6,7)$	2.926%	0.106%	1.093%	1.305%	3.311%

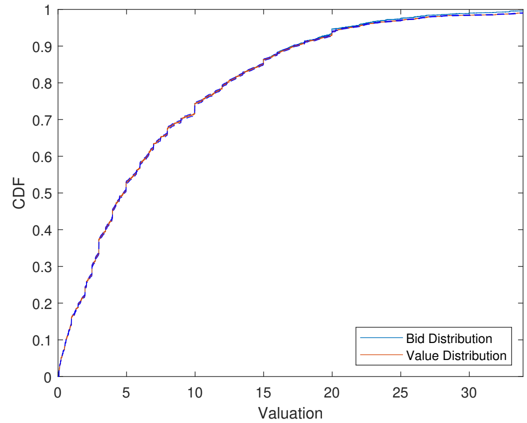
Note: The table summarizes the quantiles of bid shading in terms of the percentage of the corresponding estimated values.

We examine the distribution of valuations for each product category in Figure 4. Using empirical distributions of estimated values, we compare the estimated value distribution with the observed bid distribution. The figure demonstrates that the *car insurance* category generally has the highest per-click values, followed by *cruise*, while *coins* has the lowest. *cable TV* has a similar value distribution to *cruise* below the median, but with fewer advertisers with high per-click values. The value distribution in *laptop* falls between *coins* and *cable TV*, but this category has more advertisers with high values.

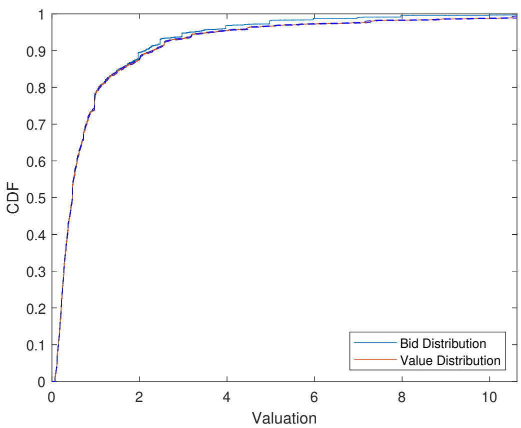
Our analysis further suggests that ad values in general follow a log-normal distribution, with varying mean and variance across categories. The Appendix C includes the approximated value distributions using log-normal distributions, which fit particularly well for the *cruise*, *cable TV*, and *coins* categories. This supports the extensive use of log-normal specifications in the theoretical auction literature for Monte Carlo simulations.



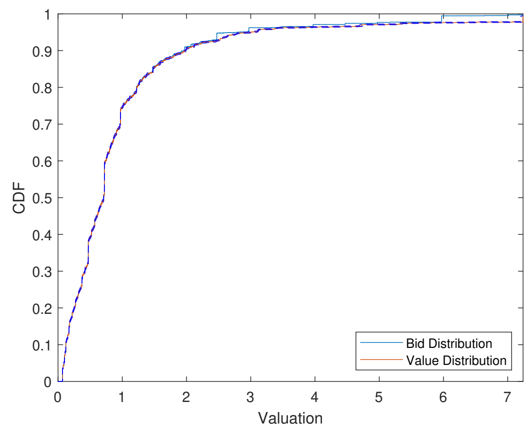
(a) Cruise



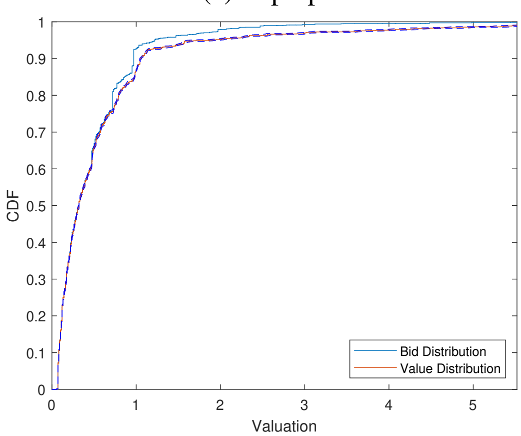
(b) Car Insurance



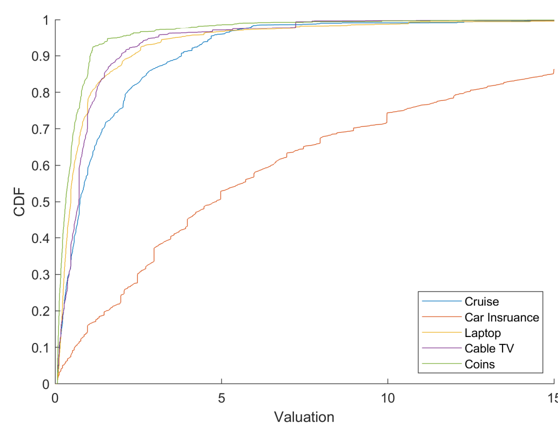
(c) Laptop



(d) Cable TV



(e) Coins



(f) Combined

FIGURE 4.—The cumulative distribution of advertiser's ad value. The blue dashed lines are the 90% confidence bands of the estimated valuation distribution obtained by bootstrap with 200 replications.

REMARK 2: **(Robustness checks)** We conduct several robustness checks by employing more flexible specifications for the first step estimation. The results are very similar to the main results. We also use the full keyword sample to see the keyword heterogeneity. The results show substantial keyword heterogeneity in the full sample. The robustness check results are provided in the Appendix D.

REMARK 3: **(Keyword heterogeneity)** we analyzed the data under the assumption of a common value distribution for different keywords, focusing on the top 10% most frequently searched keywords within each product category. However, the perceived value of a sponsored link click may vary depending on the specific keyword. For example, in the *laptop* category, an advertiser may expect higher revenue from a click on a link associated with a search query for “high-performance laptop” compared to “student laptop” or “budget laptop”. This suggests that the advertiser perhaps values some keywords more than others. Since the keywords in our data were masked, we were unable to cluster similar keywords based on their shared value distribution. Therefore, we assumed that the most commonly searched keywords had the same valuation distribution. However, in practice, one could apply the proposed framework using unsupervised machine learning techniques to cluster keywords with similar valuation distributions given more detailed keyword-specific information. Additionally, if the keywords were not masked, we could control for heterogeneity across auctions by conducting a hedonic bid regression analysis to remove the impact of observed auction features.

7. COUNTERFACTUAL ANALYSIS WITH SCORE SQUASHING

> *“When someone has a really high ad click probability, they’re very hard to beat, so it’s not a really competitive auction. So that they don’t just win [every auction], we do squashing. This makes the auction more competitive. It’s like handicapping. We handicap the people with the high click probability.”* – Preston McAfee[22]

---

[22]Chief economist of Yahoo! in 2010. See https://www.theregister.com/2010/09/16/yahoo_does_squashing/ for details on score squashing practice at Yahoo!.

In this section, we examine score squashing, a technique that reduces the weight placed on quality scores in search ad auctions. The practice of score squashing has gained recognition in the industry and has been employed by companies such as Yahoo! and by researchers at Google and Microsoft.[23] Our investigation focuses on the impact of score squashing on auction outcomes, including the search engine's revenue, the advertiser's profit, and consumer welfare.

We consider an alternative scoring mechanism, score squashing, which involves attaching monotonic, nonlinear transformations of weights (quality scores) to bids. The purpose of score squashing is to increase revenue at the cost of efficiency. The alternative score can be expressed as follows:

$$\text{New score} \rightarrow s_j^{SQ} = s_j^{\theta} \quad \forall j$$

where θ is the squashing factor in $[0,1]$. To examine the effects of score squashing, we first demonstrate that the equilibrium symmetric BNE bidding strategy derived in Section 2 still holds under the alternative scoring mechanism. This critical insight is based on the fact that Lemma 1 can still be used to derive the equilibrium, and the equivalence proven in Lemma 1 does not depend on how the quality score is calculated. As a result, the equilibrium bid in the case of score squashing is equivalent to the equilibrium bid in a GSP auction where the value is replaced by the new weighted value, $\omega_j^{SQ} = v_j \times s_j^{SQ}$. Utilizing this insight, we calculate the equilibrium bid and compare auction outcomes under different score squashing schemes.

Previous literature has explored the theoretical aspects of score squashing, but empirical studies on score squashing are scarce. Lahaie and Pennock (2007) and subsequent works such as Charles et al. (2016) have found that the impact of squashing depends on the ranking of advertisers based on quality scores, weighted bids, and squashed weighted bids. They have shown that revenue increases if all three rankings result in the same ranking of

[23] For instance, Google holds a patent, Provine et al. (2013), which describes methods for ranking search results using a boosting factor, such as a squashing function. Furthermore, researchers at search engines have published works such as Charles et al. (2016), Bergemann et al. (2022), and Lahaie and Pennock (2007).

advertisers in the top $K+1$ positions. This is because squashing does not alter the ranking and does not harm revenue by ranking a low-quality ad high, only impacting revenue through price increases for high-quality advertisers. Conversely, the impact of squashing on revenue is negative if both the weighted bid ranking and squashed weighted bid ranking differ from the quality score ranking. The effect on revenue is ambiguous if the weighted bid and squashed weighted bid ranking differ.

These insights provide an understanding of how squashing changes revenue through changes in winning advertisers and price increases for winners. However, drawing practical implications from these insights is challenging as a large number of ad auctions are held daily. In practice, it is more useful to examine how the outcome changes with auction characteristics at the search-topic level, such as the value and quality score distributions in each auction.

We aim to address the knowledge gap surrounding the effects of score squashing by presenting empirical evidence through a comprehensive counterfactual analysis. While AN10 has explored the impact of score squashing with a fixed squashing factor of $\theta = \frac{1}{2}$, our study extends this work by considering a wider range of squashing factors and examining the role of market-specific factors in determining the effects of squashing on auction outcomes. Our results indicate that different product categories exhibit varying revenue-maximizing squashing levels, and we provide further insight by investigating how changing market parameters such as the number of advertisers, click rates, and the distribution of value and quality scores can impact the optimal squashing level.

7.1. *Counterfactual simulations*

We conduct counterfactual experiments to examine the impact of varying levels of the score squashing factor θ on auction outcomes. The score squashing factor is varied in the range of $\theta \in \{0, 0.1, 0.25, 0.5, 0.75, 1\}$. We simulate 2000 auctions for each category and squashing factor combination and collect the resulting auction outcomes. To evaluate the performance of the auction, we calculate the average values of the search engine's revenue, advertisers' profit, and ad quality across all simulated auctions. Ad quality serves as a proxy measure of consumer welfare. We also consider the sum of revenue and profit to reflect the potential competition from other advertising platforms. Thus, in the long term, search engines might want to maximize a weighted average of their revenue and advertisers' profit.

The observed bids in the data cannot be used in our analysis, as the equilibrium bidding strategies vary with the auction mechanism (the squashing factor). We need to calculate the equilibrium bid function given the population distribution and density functions of the weighted value in each scenario. Due to the complexity of numerically implementing the equilibrium bid function in Lemma 2, an iterative approximation procedure was employed instead. The specifics of this procedure can be found in the Appendix E. Given that the value and quality score distributions in the data are well approximated by log-normal distributions, we assume that the value and quality scores follow log-normal distributions ($v_j \sim L(\mu_v, \sigma_v^2), s_j \sim L(\mu_s, \sigma_s^2)$). As a result, the weighted value with the squashing factor θ is also log-normally distributed ($w_j^{SQ} \sim L(\mu_v + \theta\mu_s, \sigma_v^2 + \theta^2\sigma_s^2)$). To estimate the means and variances of the log-normal distributions, we use maximum likelihood estimation (MLE). The estimated distributions approximate the empirical weighted value distributions well.

The algorithm determining the auction outcomes is outlined in the steps below. Given a product category and a value for the score squashing factor θ, the following steps are executed for each simulation iteration:

1. Use maximum likelihood estimation (MLE) to estimate the means and variances of the log-normal distributions of the valuation and quality score.
2. Solve for the equilibrium bid function using the weighted value distribution and position-specific click rates $\{\hat{c}_1, \cdots, \hat{c}_K\}$.
3. Draw N independent values and quality scores from estimated log-normal distributions:

$$v_j \sim L(\hat{\mu}_v, \hat{\sigma}_v^2), \quad s_j \sim L(\hat{\mu}_s, \hat{\sigma}_s^2)$$

4. Compute the weighted bids using the equilibrium bidding function and determine the auction outcome. Calculate the auction revenue, advertisers' profit, and ad quality for the top 7 ad positions.
5. Repeat Steps 3 to 4 for 2000 replications and compute the average revenue, profit, and ad quality.

7.2. *Results*

Figure 5 presents the counterfactual simulation results. The optimal level of squashing differs among categories, implying that a localized decision for squashing is more effective in maximizing revenue. In general, a well-chosen θ improves revenue at the cost of advertiser profit and consumer welfare. Only the *cable TV* category shows no improvement in revenue through squashing. The optimal squashing level tends to be lower in categories with a low number of advertisers (*cruise* and *coins*).

Advertiser profit is maximized without squashing, except for *cruise*, where the profit is second highest with no squashing, close to its maximum value. This result is expected as the increase in search engine revenue leads to higher prices paid by winning advertisers, leading to an inverse relationship between auction revenue and advertiser profit. Squashing reduces welfare monotonically as the value of θ decreases, leading to a decrease in average ad quality across all categories. This implies that the price increase induced by squashing has a greater impact than the decrease in quality (lower click probability) on revenue. The maximum sum of revenue and profit is achieved at $\theta = 0.5$ for *cruise*, $\theta = 0.75$ for *car insurance* and *laptop*, and with no squashing for the other categories. This suggests that the search engine has to consider a higher value of θ when caring about advertiser profit for long-term relationships.

Further, we investigate the heterogeneity in the impact of squashing across categories by examining the change in optimal squashing when switching one of the auction parameters between categories. We take *cruise* as the base category and examine the impact of replacing its value distribution, score distribution, position-specific click-through rates, or number of advertisers with another category. The results are presented in Figure 6.

We first focus on the impact of the number of advertisers. We switch the number of advertisers of *cruise* ($N = 45$) with that of *cable TV* ($N = 142$) and find that higher competition in the market results in an increase in both the revenue-maximizing squashing level (from 0.25 to 1) and the advertisers' profit-maximizing squashing level (from 0.5 to 1). However, the impact of competition is not monotonic. A similar switch with car insurance ($N = 403$) results in a lesser increase in the revenue-maximizing squashing level (from 0.25 to 0.75).

Next, we look at the impact of click-through rates. *Cruise* and *laptop* have similar changes in click rates from position 1 to 2, but the click rate decreases more sharply for *cruise*, with the last two positions getting 1% of clicks received by the top position, relative to *laptop* where the last two positions still get 27-31$\%$ of the click received by the first position. Surprisingly, switching the click rates between the two categories does not change the results. We found similar evidence when we switched *cruise*'s click rates with the other categories. Our findings suggest that position-specific click rates might have a marginal impact on the optimal squashing level.

Finally, we examine the impact of the value and quality score distributions. By switching the score distribution of *cruise* with *coins*, we find an increase in the revenue-maximizing squashing level (from 0.25 to 0.5), but no change in the advertisers' profit-maximizing squashing level. *Cruise* has a more right-skewed score distribution than *coins*, meaning that *coins* have more high-quality advertisers relative to *cruise* so that competition for top ad positions is fiercer in *coins*. On the other hand, switching the value distribution between *cruise* and *laptop* leads to an increase in both the revenue-maximizing squashing level (from 0.25 to 0.5) and the advertisers' profit-maximizing squashing level (from 0.5 to 1). Again, the value distribution of *cruise* is more right-skewed than *laptop* so this switch results in more competitive auctions. These results suggest that more competition results in an increase in the optimal squashing level for both the search engine and the advertisers.

In conclusion, our experiments shed light on the impact of various factors on the optimal value of θ in the advertising market. Our findings suggest that the number of advertisers, the quality score and valuation distributions have a significant impact on the optimal squashing level, while the impact of click-through rates is marginal. More competitive mar-

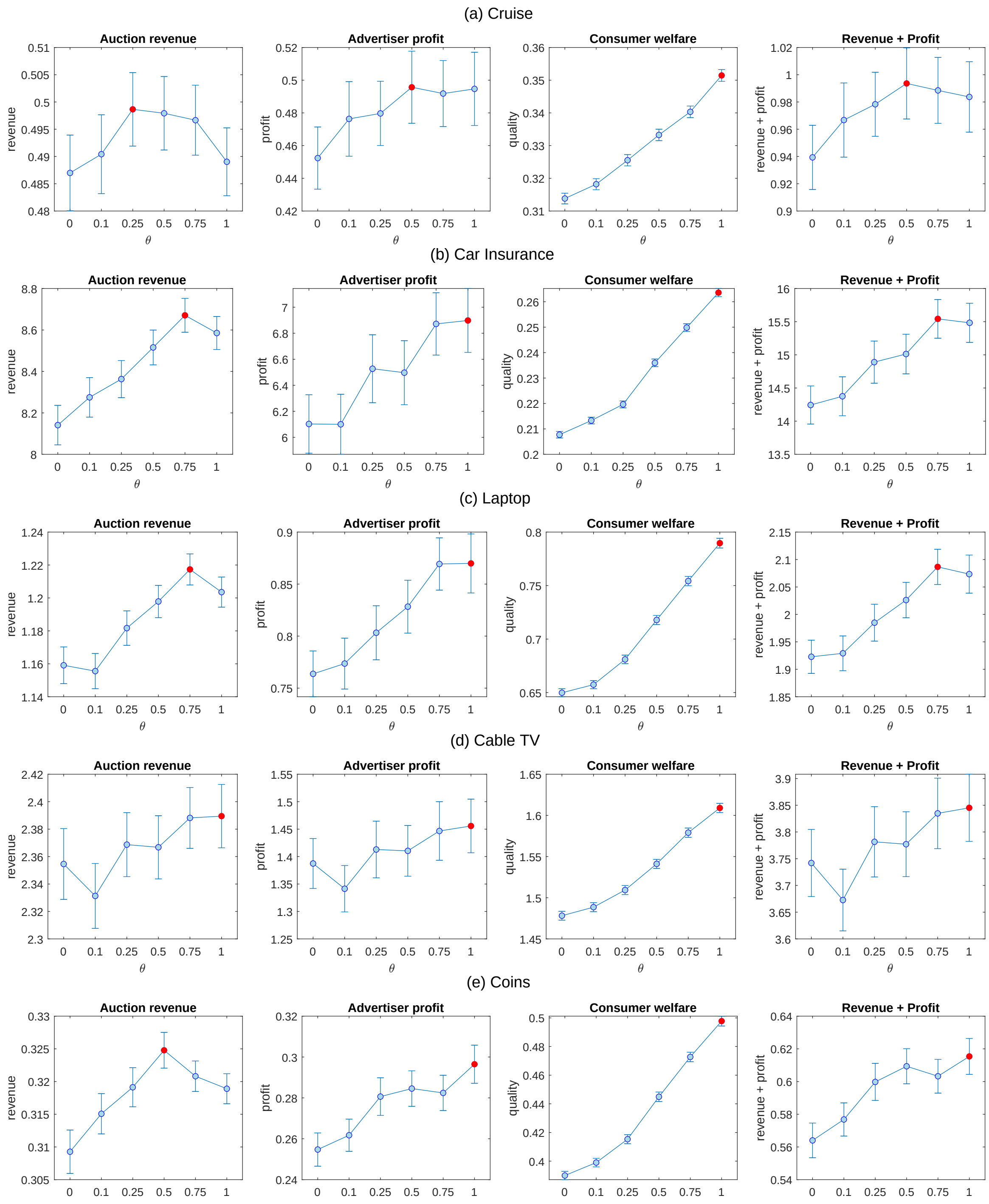


FIGURE 5.—Counterfactual simulations with score squashing. The dots are averages across 2000 simulated auctions. The red dots are maximums across values of θ. The ranges are 95% confidence intervals.

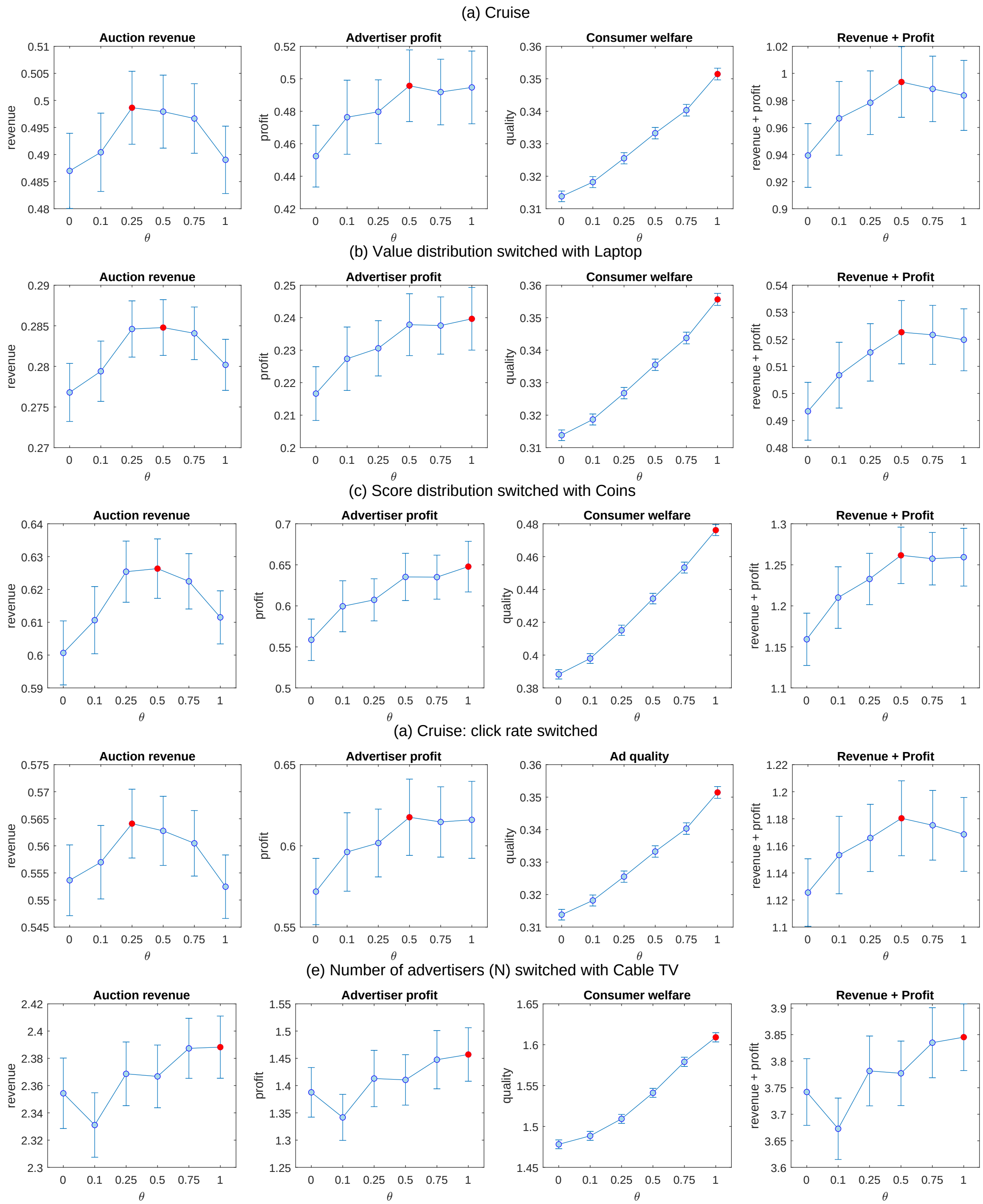


FIGURE 6.—*Cruise*: switching auction characteristics with other categories. The dots are averages across 2000 simulated auctions. The red dots are maximums across values of θ. The ranges are 95% confidence intervals.

ket environments in general lead to higher optimal squashing levels, although the number of advertisers has non-monotonic impacts on the optimal θ. Our counterfactual analysis provides evidence that the implementation of score squashing can increase auction revenue, however, it comes at the expense of advertisers' profit and consumer welfare. The optimal level of squashing varies based on specific auction characteristics. Our framework provides a comprehensive approach to determining the most appropriate level of squashing for a given auction. Search engines have the flexibility to adopt an objective function that reflects their priorities by choosing a weighted average of the auction outcomes.

8. CONCLUSION

Our proposed framework provides a valuable tool to estimate the value distribution in the GSP^{ω} auctions under more realistic assumptions than the prior literature. The estimated value distribution is used for several counterfactual analyses by the auctioneer (i.e., the search engine), such as calculating the reserve price, analyzing the impact of a change in an auction design on the advertisers, and calculating the optimal score squashing level. The use of counterfactual simulations allows for a deeper understanding of how the advertiser's bidding behavior may respond to changes in market mechanisms. Unlike the consumer side, for which search engines can conduct randomized controlled trials to examine the effect of a potential change, experiments are hard to implement on the advertiser side because the advertisers' responses to changes in market factors, such as pricing mechanisms, are usually slower. In addition, frequent changes in the market environment can motivate advertisers to leave the advertising platform due to increased difficulties.

Our model addresses the uncertainty in competitors' values and quality scores by incorporating the incomplete information assumption, adding further robustness to the framework. Further research could consider extending the model to include additional sources of uncertainty, such as entry decisions and position-specific valuations. Such extensions would be promising avenues for future work.

REFERENCES

AGARWAL, ASHISH AND TRIDAS MUKHOPADHYAY (2016): "The impact of competing ads on click performance in sponsored search," *Information Systems Research*, 27 (3), 538–557. [21]

ATHEY, SUSAN AND GLENN ELLISON (2011): "Position auctions with consumer search," *The Quarterly Journal of Economics*, 126 (3), 1213–1270. [6]

ATHEY, SUSAN AND DENIS NEKIPELOV (2010): "A structural model of sponsored search advertising auctions," in *Sixth ad auctions workshop*, vol. 15. [5, 1]

BAE, JINSOO AND JOHN H KAGEL (2019): "An experimental study of the generalized second price auction," *International Journal of Industrial Organization*, 63, 44–68. [7]

BERGEMANN, DIRK, PAUL DÜTTING, RENATO PAES LEME, AND SONG ZUO (2022): "Calibrated Click-Through Auctions," in *Proceedings of the ACM Web Conference 2022*, 47–57. [33]

BÖRGERS, TILMAN, INGEMAR COX, MARTIN PESENDORFER, AND VACLAV PETRICEK (2013): "Equilibrium bids in sponsored search auctions: Theory and evidence," *American economic Journal: microeconomics*, 5 (4), 163–87. [6]

CARAGIANNIS, IOANNIS, CHRISTOS KAKLAMANIS, PANAGIOTIS KANELLOPOULOS, MARIA KYROPOULOU, BRENDAN LUCIER, RENATO PAES LEME, AND ÉVA TARDOS (2015): "Bounding the inefficiency of outcomes in generalized second price auctions," *Journal of Economic Theory*, 156, 343–388. [6]

CHARLES, DENIS, NIKHIL R DEVANUR, AND BALASUBRAMANIAN SIVAN (2016): "Multi-score position auctions," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, 417–425. [33]

CHE, YEON-KOO, SYNGJOO CHOI, AND JINWOO KIM (2017): "An experimental study of sponsored-search auctions," *Games and Economic Behavior*, 102, 20–43. [7]

DECAROLIS, FRANCESCO AND GABRIELE ROVIGATTI (2021): "From mad men to maths men: Concentration and buyer power in online advertising," *American Economic Review*, 111 (10), 3299–3327. [6]

EDELMAN, BENJAMIN, MICHAEL OSTROVSKY, AND MICHAEL SCHWARZ (2007): "Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords," *American economic review*, 97 (1), 242–259. [5]

EDELMAN, BENJAMIN AND MICHAEL SCHWARZ (2010): "Optimal auction design and equilibrium selection in sponsored search auctions," *American Economic Review*, 100 (2), 597–602. [6]

GHOSE, ANINDYA AND SHA YANG (2009): "An empirical analysis of search engine advertising: Sponsored search in electronic markets," *Management science*, 55 (10), 1605–1622. [6]

GOMES, RENATO AND KANE SWEENEY (2014): "Bayes–nash equilibria of the generalized second-price auction," *Games and Economic Behavior*, 86, 421–437. [3]

GUERRE, EMMANUEL, ISABELLE PERRIGNE, AND QUANG VUONG (2000): "Optimal nonparametric estimation of first-price auctions," *Econometrica*, 68 (3), 525–574. [3, 17]

HAILE, PHILIP A AND ELIE TAMER (2003): "Inference with an incomplete model of English auctions," *Journal of Political Economy*, 111 (1), 1–51. [2]

HSIEH, YU-WEI, MATTHEW SHUM, AND SHA YANG (2015): "To score or not to score? Estimates of a sponsored search auction model," *Estimates of a Sponsored Search Auction Model (February 1, 2015). USC-INET Research Paper*, (15-09). [6]

KRISHNA, VIJAY (2009): *Auction theory*, Academic press. [19]

LAHAIE, SÉBASTIEN AND DAVID M PENNOCK (2007): "Revenue analysis of a family of ranking rules for keyword auctions," in *Proceedings of the 8th ACM Conference on Electronic Commerce*, 50–56. [33]

MILGROM, PAUL AND ILYA SEGAL (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70 (2), 583–601. [5]

MOHRI, MEHRYAR AND ANDRÉS MUNOZ MEDINA (2015): "Non-parametric revenue optimization for generalized second price auctions," *arXiv preprint arXiv:1506.02719*. [7, 8]

OSTROVSKY, MICHAEL AND MICHAEL SCHWARZ (2011): "Reserve prices in internet advertising auctions: A field experiment," in *Proceedings of the 12th ACM conference on Electronic commerce*, 59–60. [7]

PROVINE, JOHN R, CLAUDIO BASILE, AND ALEXANDER L VASSERMAN (2013): "Modifying search result ranking based on corpus search statistics," US Patent 8,359,309. [33]

RAMABOA, KUTLWANO KKM AND PETER FISH (2018): "Keyword length and matching options as indicators of search intent in sponsored search," *Information Processing & Management*, 54 (2), 175–183. [23]

SUN, YANG, YUNHONG ZHOU, AND XIAOTIE DENG (2014): "Optimal reserve prices in weighted GSP auctions," *Electronic Commerce Research and Applications*, 13 (3), 178–187. [6, 7, 8]

SVITAK, JAN, JAN TICHEM, AND STEFAN HAASBEEK (2021): "Price effects of search advertising restrictions," *International Journal of Industrial Organization*, 77, 102736. [6]

VARIAN, HAL R (2007): "Position auctions," *international Journal of industrial Organization*, 25 (6), 1163–1178. [5, 6, 9]

YAO, SONG AND CARL F MELA (2011): "A dynamic model of sponsored search advertising," *Marketing Science*, 30 (3), 447–468. [6]

Online Supplementary Material for Kim and Pal (2023)

This online supplementary material presents extensions of our methodology in Appendix A, all the theoretical proofs in Appendix B, log-normal approximations of estimated valuation distributions in our empirical application in Appendix C, robustness checks in Appendix D, the iterative approximation procedure for the equilibrium bid function used in the counterfactual analysis in Appendix E, and details on identification of product categories in our data in Appendix F.

APPENDIX A: EXTENSIONS

The proposed method possesses a key advantage in its versatility in incorporating additional features present in online search auctions. This study specifically examines two noteworthy cases. In the first case, we only observe winning bids, while the latter case involves the presence of a fixed reserve price that is observable.

A.1. *Case 1: Limited bid data*

In sponsored search auctions, a full set of bids is sometimes not available to the researcher. For instance, Athey and Nekipelov (2010) (AN10 hereafter) discusses a potential missing data problem, as they only observe a subset of weighted bids in each auction.[24] This can potentially lead to bias in the estimation of the empirical weighted bid distribution. In this section, we provide one possible way to resolve this bias. The advantage of search bid data is that we observe the weighted bids as well as their order statistic. We can infer the order of an observed weighted bid by the winning position. Under this setting, we need very limited bid information for our proposed method, as long as we know the total number of advertisers. Consider the case where we only observe the weighted bid of the advertiser who won the first position. Recall, in terms of the order statistic notation,

[24] AN10 observes only the winning bids and the highest (in terms of weighted bid) non-winning bid. In our data, We observe a large set of non-winning bids (beyond the first search result page) so this problem is negligible.

we observe $b_w^k \sim G_{w,k:N}(\cdot)$, where $G_{w,k:N}(\cdot)$ is the distribution of the kth highest order statistic.

To further exploit the order statistic, we use the following equivalence for the order statistic

$$G_{w,k:N}(b_w) = \int_0^{G_w(b)} \frac{N!}{(N-k)!(k-1)!} t^{N-k}(1-t)^{k-1} dt$$
$$= \Gamma(G_w(b_w); N-k+1, k). \tag{12}$$

We can use the above equation to see the relationship between the order statistic distribution and the primitive distribution given an incomplete beta function, represented here as $\Gamma(z; a, b)$ with $a = N-k+1$ and $b = k$.

Next, we can use Equation (12) and a well-known property of order statistic i.i.d variables, which states that given a known number (N) of i.i.d. draws of weighted bids from $G_w(\cdot)$, the order statistic distribution can be used to identify the parent distribution $G_w(\cdot)$ using the following equation:

$$G_w(b_w) = \Gamma^{-1}\Big(G_{w,k:n}(b_w); N-k+1, k\Big). \tag{13}$$

where Γ^{-1} is the inverse of the incomplete beta function defined in Equation (12).[25] Once we have estimated the weighted bid distribution, all the other steps follow the same as our main results.

A.2. *Case 2: Reserve Price*

We explore the case in which the data have a known fixed reserve price, denoted by r. Assume that the search engine sets a fixed reserve price observed by everyone. In such a case, we can still derive the value given observable quantities. However, the solution for the value is now obtained by modifying the equation in Theorem 1. The following theorem proves the identification of the valuation in the presence of a reserve price.

[25] For further details of using this property in an auction context, refer to Lemma 2 of Haile and Tamer (2003).

THEOREM 3: *Under Assumptions 1–2, the advertiser value, v, is identified by:*

$$v = b + \Phi(G_w, b, s|\mathcal{C}, K, N, r) \tag{14}$$

where

$$\Phi(G_w, b, s|\mathcal{C}, K, N, r) = \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1 - G_w(b_w))^{k-2} \left(\int_r^{b_w} G_w(u)^{N-k} du - r G_w(r)^{N-k} \right)}{s_j \sum_{k=1}^{K} c_k \binom{N-1}{k-1} G_w(b_w)^{N-k-1} (1 - G_w(b_w))^{k-2} \Big[(N-k)(1 - G_w(b_w)) - (k-1) G_w(b_w) \Big]}$$

given the quality score (s), the equilibrium bid (b), the distribution function of equilibrium weighted bids (G_w), the number of available ad positions (K), click rates across ad positions $(\mathcal{C})$, the number of advertisers (N), and the fixed reserve price (r).

We have now a lower bound on the price which is the reserve price. More details are provided in the proof in appendix B. Note that bids below the reserve price are not observed. In such a case, one can use the strategy proposed in the previous subsection (A.1) to derive the weighted bid distribution. If the number of advertisers is also not observed, one can assume that number of advertisers is equal to the maximum number of observed bids across auctions as a proxy for the number of advertisers. A similar proxy variable was used in Guerre et al. (2000).

APPENDIX B: THEORETICAL PROOFS

Proof of Lemma 1: Consider an arbitrary advertiser j and suppose that all advertisers $l \neq j$ have the following weighted bid strategy equivalence

$$b_w^{GSP^w}(v_l, s_l) = b^{GSP}(\omega_l) \qquad \forall l \neq j \tag{15}$$

We will show that, in this case, the equilibrium weighted bid for advertiser j is also equal to the equilibrium bid in the GSP auction.

Using Equation (3), we can show that the equilibrium bid b^{GSP^w} for advertiser j is given as:

$$b^{GSP^w} = arg\max_{\hat{b}} \Pi(\hat{b}|v_j, s_j)$$
$$= arg\max_{\hat{b}} \sum_{k=1}^{K} s_j c_k \left[v_j - \frac{\mathbb{E}\left(b_w^{GSP^w,[k+1]} \middle| \hat{b} \times s_j = b_w^{GSP^w,[k]} \right)}{s_j} \right] \times P(\hat{b} \times s_j = b_w^{GSP^w,[k]}). \tag{16}$$

As s_j is known to j and the auctioneer, the above problem can be rewritten to maximize $\hat{b}_w = \hat{b} \times s_j$:

$$b^{GSP^w} = arg\max_{\hat{b}_w} \sum_{k=1}^{K} c_k \left[w_j - \mathbb{E}\left(b_w^{GSP^w,[k+1]} \middle| \hat{b}_w = b_w^{GSP^w,[k]} \right) \right] \times P(\hat{b}_w = b_w^{GSP^w,[k]}). \tag{17}$$

Given the Equation(15), the above equation can be rewritten as:

$$b^{GSP^w} = arg\max_{\hat{b}_w} \sum_{k=1}^{K} c_k \left[w_j - \mathbb{E}\left(b_w^{GSP,[k+1]} \middle| \hat{b}_w = b_w^{GSP,[k]} \right) \right] \times P(\hat{b}_w = b_w^{GSP,[k]}). \tag{18}$$

Now consider a GSP auction in which the advertiser j's value is replaced by the weighted value (ω_j). Recall, that in the non-weighted GSP auction, the allocation is done according to the ranking of the bids and the price is equal to the highest bid below you. The equilibrium bid b^{GSP} for advertiser j is given as :

$$b^{GSP} = arg\max_{\hat{b}_w} \sum_{k=1}^{K} c_k \left[w_j - \mathbb{E}\left(b_w^{GSP,[k+1]} \middle| \hat{b}_w = b_w^{GSP,[k]} \right) \right] \times P(\hat{b}_w = b_w^{GSP,[k]}). \tag{19}$$

The optimization problems in Equations (18) and (19) are equivalent. Thus, for any arbitrary j, if all other advertisers $l \neq j$ have $b^{GSP}(\omega_l) = b_w^{GSP_w}(v_l, s_l)$, then advertiser j also has $b^{GSP}(\omega_j) = b_w^{GSP_w}(v_j, s_j)$. Hence, the equilibrium weighted bid in GSP^w is equal to the equilibrium bid of a GSP auction where the values are replaced by the weighted value. *Q.E.D.*

Proof of Theorem 1: By Lemma 2, we know an efficient equilibrium exists in this auction. In an efficient allocation, an advertiser with value v and quality score s, hence weighted value $\omega = v \times s$, wins the k-th ad position with probability:

$$\zeta_k(\omega) \equiv P(w^{[k+1]} \leq w \leq w^{[k-1]}) = \binom{N-1}{k-1} (1 - F_w(\omega))^{k-1} F_w^{N-k}(\omega). \tag{20}$$

Using Revelation Principle, an advertiser with weighted value ω has payoff that satisfies

$$\omega = \underset{\hat{\omega}}{\operatorname{argmax}} \sum_{k=1}^{K} \zeta_k(\hat{\omega}) c_k \left[\omega - \mathbb{E}\left(b_w(\omega^{[k+1]}) \middle| \omega^{[k+1]} \leq \hat{\omega} \leq \omega^{[k-1]} \right) \right] \tag{21}$$

Applying the envelop theorem (see Milgrom and Segal (2002)) in the payoff function in Equation (21), we have:

$$\frac{d}{d\omega} \Pi(\omega) = \sum_{k=1}^{K} c_k \zeta_k(\omega) \tag{22}$$

and also using the Fundamental Theorem of Calculus, we get

$$\Pi(\omega) = \Pi(\underline{\omega}) + \sum_{k=1}^{K} c_k \int_0^{\omega} \zeta_k(x) dx. \tag{23}$$

As $b_w(\omega)$ is increasing, a bidder with type $\underline{\omega}$ never has a non-zero payoff – $\Pi(\underline{\omega}) = 0$, we have

$$\Pi(\omega) = \sum_{k=1}^{K} c_k \int_0^{\omega} \zeta_k(x) dx. \tag{24}$$

Furthermore, using Equations (21) and (24), we obtain

$$\sum_{k=1}^{K} c_k \int_0^{\omega} \zeta_k(x) dx = \sum_{k=1}^{K} c_k \zeta_k(\omega) \left[\omega - \mathbb{E}\left(b_w(\omega^{[k+1]}) \middle| \omega^{[k+1]} \leq \omega \leq \omega^{[k-1]} \right) \right]$$

which can be rearranged as

$$\sum_{k=1}^{K} c_k \left[\zeta_k(\omega)\omega - \int_0^{\omega} \zeta_k(x) dx \right] = \sum_{k=1}^{K} c_k \zeta_k(\omega) \mathbb{E}\left(b_w(\omega^{[k+1]}) \middle| \omega^{[k+1]} \leq \omega \leq \omega^{[k-1]} \right).$$

Using integration by parts on the left-hand side we derive

$$\sum_{k=1}^{K} c_k \int_0^{\omega} x \frac{d\zeta_k(x)}{dx} dx = \sum_{k=1}^{K} c_k \zeta_k(\omega) \mathbb{E}\left(b_w(\omega^{[k+1]}) \middle| \omega^{[k+1]} \leq \omega \leq \omega^{[k-1]} \right).$$

Opening up the expectation on the right-hand side,

$$\sum_{k=1}^{K} c_k \int_0^{\omega} x \frac{d\zeta_k(x)}{dx} dx = \sum_{k=1}^{K} c_k \zeta_k(\omega) \frac{\int_0^{\omega} b_w(x)(N-k)F_w(x)^{N-k-1} f_w(x) dx}{F_w(\omega)^{N-k}}$$

and substituting $\zeta_k(\omega)$ using Equation (20) in the right hand side yields,

$$\sum_{k=1}^{K} c_k \int_0^{\omega} x \frac{d\zeta_k(x)}{dx} dx = \\ \sum_{k=1}^{K} c_k \binom{N-1}{k-1} \int_0^{\omega} b_w(x)(N-k)F_w(x)^{N-k-1} f_w(x)(1-F_w(\omega))^{k-1} dx.$$

Differentiating both sides we get:

$$\sum_{k=1}^{K} c_k \omega \frac{d\zeta_k(\omega)}{d\omega} = \sum_{k=1}^{K} c_k \binom{N-1}{k-1} b_w(\omega)(N-k)F_w(\omega)^{N-k-1} f_w(\omega)(1-F_w(\omega))^{k-1} \\ -\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} f_w(\omega) \int_0^{\omega} b_w(x)(N-k)F_w(x)^{N-k-1} f_w(x) dx. \tag{25}$$

The integral part in Equation (25) can be re-written using integration by parts:

$$\int_0^{\omega} b_w(x)(N-k)F_w(x)^{N-k-1} f_w(x) dx = b_w(\omega)F_w(\omega)^{N-k} - \int_0^{\omega} b_w'(x)F_w(x)^{N-k} dx \tag{26}$$

Under Assumption 2, the integral in the last term of Equation (26) can be replaced by

$$\int_0^{\omega} b_w'(x) G_w(b_w(x))^{N-k} dx$$

where G_w is the distribution function of weighted bids. Integration by substitution yields:

$$\int_0^{\omega} b_w'(x) G_w(b_w(x))^{N-k} dx = \int_0^{b_w(\omega)} G_w(u)^{N-k} du$$

and therefore plugging the above expression into Equation (25) using Equation (26),

$$\sum_{k=1}^{K} c_k \omega \frac{d(\zeta_k(\omega))}{d\omega} = \sum_{k=1}^{K} c_k \binom{N-1}{k-1} b_w(\omega)(N-k)F_w(\omega)^{N-k-1} f_w(\omega)(1-F_w(\omega))^{k-1}$$

$$- \sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} f_w(\omega) b_w(\omega) F_w(\omega)^{N-k}$$

$$+ \sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} f_w(\omega) \int_0^{b_w(\omega)} G_w(u)^{N-k} du.$$

Now the last step is to open up $\frac{d(\zeta_k(\omega))}{d\omega}$ and to rearrange.

$$\omega = b_w(\omega) + \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} \int_0^{b_w(\omega)} G_w(u)^{N-k} du}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} F_w(\omega)^{N-k-1}(1-F_w(\omega))^{k-2} \Big[(N-k)(1-F_w(\omega)) - (k-1)F_w(\omega)\Big]}$$

Divide both sides by quality score s:

$$v = b + \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} \int_0^{b_w(\omega)} G_w(u)^{N-k} du}{s \sum_{k=1}^{K} c_k \binom{N-1}{k-1} F_w(\omega)^{N-k-1}(1-F_w(\omega))^{k-2} \Big[(N-k)(1-F_w(\omega)) - (k-1)F_w(\omega)\Big]}$$

The only unobservable object in the solution is $F_w(x)$. By replacing it with $G_w(b_w(x))$, the desired expression in the theorem is obtained. *Q.E.D.*

Proof of Theorem 2: The advertiser's expected payoff with weighted value ω is given by

$$\Pi(\omega) = \sum_{k=1}^{K} \zeta_k(\omega) c_k \left[\omega - \mathbb{E}\left(b_w(\omega^{[k+1]}) \middle| \omega^{[k+1]} \le \omega \le \omega^{[k-1]}\right)\right].$$

Suppose that the advertiser maximizes his expected payoff given ω and the weighted bid distribution G_w by choosing optimal b_w. The payoff function is re-written as

$$\Pi(b_w; \omega) = \sum_{k=1}^{K} \binom{N-1}{k-1} (1-G_w(b_w))^{k-1} G_w^{N-k}(b_w) c_k \left[\omega - \mathbb{E}\left(b_w^{[k+1]} \middle| b_w^{[k+1]} \le b_w \le b_w^{[k]}\right)\right]$$

$$=\sum_{k=1}^{K}\binom{N-1}{k-1}\left(1-G_w(b_w)\right)^{k-1}G_w^{N-k}(b_w)c_k\left[\omega-\frac{\int_0^{b_w}x(N-k)G_w(x)^{N-k-1}g_w(x)dx}{G_w^{N-k}(b_w)}\right]$$

By integration by parts, the last integral term in the equation above becomes

$$\int_0^{b_w}x(N-k)G_w(x)^{N-k-1}g_w(x)dx=b_wG_w^{N-k}(b_w)-\int_0^{b_w}G_w^{N-k}(x)dx,$$

and then the payoff is:

$$\Pi(b_w;\omega)=\sum_{k=1}^{K}\binom{N-1}{k-1}c_k\left(1-G_w(b_w)\right)^{k-1}G_w^{N-k}(b_w)\left[\omega-b_w+\frac{\int_0^{b_w}G_w^{N-k}(x)dx}{G_w^{N-k}(b_w)}\right]$$

Derive the first order condition by differentiating $\Pi(b_w;\omega)$ w.r.t. b_w :

$$(\omega-b_w)\sum_{k=1}^{K}\binom{N-1}{k-1}c_k\left(1-G_w(b_w)\right)^{k-2}G_w^{N-k-1}(b_w)\Big[(N-k)\left(1-G_w(b_w)\right)-(k-1)G_w(b_w)\Big]$$
$$=\sum_{k=1}^{K}\binom{N-1}{k-1}c_k(k-1)\left(1-G_w(b_w)\right)^{k-2}\int_0^{b_w}G_w^{N-k}(x)dx$$

and therefore

$$b_w=\omega-\frac{\sum_{k=1}^{K}\binom{N-1}{k-1}c_k(k-1)\left(1-G_w(b_w)\right)^{k-2}\int_0^{b_w}G_w^{N-k}(x)dx}{\sum_{k=1}^{K}\binom{N-1}{k-1}c_k\left(1-G_w(b_w)\right)^{k-2}G_w^{N-k-1}(b_w)\Big[(N-k)\left(1-G_w(b_w)\right)-(k-1)G_w(b_w)\Big]}.$$

Now dividing both sides with quality score, s, yields the desired result. *Q.E.D.*

Proof of Theorem 3: Using the same argument as used in the proof of Theorem 1, given the fixed reserve price r, the advertiser with the weighted value ω has payoff that satisfies

$$\omega=\underset{\hat{\omega}}{\operatorname{argmax}}\sum_{k=1}^{K}\zeta_k(\hat{\omega})c_k\left[\omega_j-\mathbb{E}\left(Max\{b_w(\omega^{[k+1]}),r\}\middle|\omega^{[k+1]}\leq\hat{\omega}\leq\omega^{[k-1]}\right)\right]\quad(27)$$

Additionally, similar to Equation (24) derived in the proof of Theorem 1, here again we derive the advertiser's payoff function to be equal to

$$\Pi(\omega) = \sum_{k=1}^{K} c_k \int_r^{\omega} \zeta_k(x)dx. \tag{28}$$

Using Equations (27) and (28), we obtain

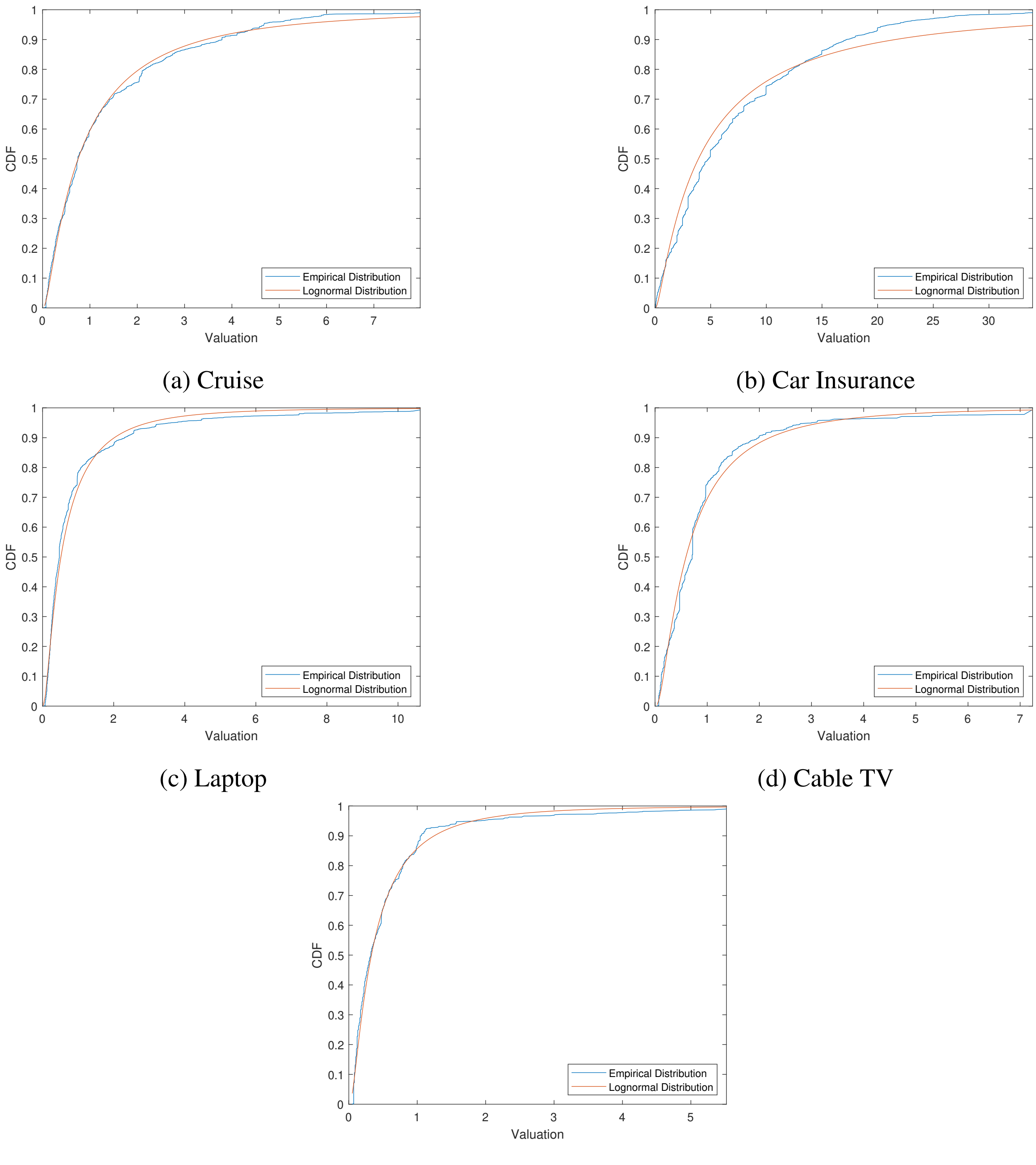
$$\sum_{k=1}^{K} c_k \int_r^{\omega} \zeta_k(x)dx = \sum_{k=1}^{K} c_k \zeta_k(\omega) \left[\omega - \mathbb{E}\left(Max\{b_w(\omega^{[k+1]}), r\} \middle| \omega^{[k+1]} \leq \hat{\omega} \leq \omega^{[k-1]} \right)\right]$$

After differentiating the above equation and algebraic manipulations we yield

$$\begin{aligned}
&\sum_{k=1}^{K} c_k \omega \frac{d\zeta_k(\omega)}{d\omega} = \sum_{k=1}^{K} c_k \binom{N-1}{k-1} b_w(\omega)(N-k)F_w(\omega)^{N-k-1} f_w(\omega)(1-F_w(\omega))^{k-1} \\
&- \sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} f_w(\omega) \int_r^{\omega} b_w(x)(N-k)F_w(x)^{N-k-1} f_w(x)dx \\
&- \sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} f_w(\omega) \int_0^{r} r(N-k)F_w(x)^{N-k-1} f_w(x)dx
\end{aligned} \tag{29}$$

Then the same steps used in the proof of Theorem 1 obtain the desired result. *Q.E.D.*

APPENDIX C: LOG-NORMAL APPROXIMATION OF THE VALUE DISTRIBUTION



(a) Cruise (b) Car Insurance

(c) Laptop (d) Cable TV

(e) Coins

FIGURE C.1.—The fitted log-normal distribution of advertiser's ad value. The blue lines are the empirical distributions and the red lines are fitted log-normal distributions.

APPENDIX D: ROBUSTNESS CHECK

D.1. *Nonparametric control for keylength*

In the first step of our estimation procedure, we only control for keyword popularity, the length of a keyword (keylength), and the weekend dummy to estimate the position-specific click rates and quality scores. Here we further exploit the discrete nature of the keylength variable to nonparametrically control for the length of keywords. We generate the keylength dummy for each discrete value and include them in the first step regression. The more flexible specification yields almost identical results to our baseline results.

D.2. *Including full keyword dummies*

We also include the full set of keyword dummy variables to control for keyword-fixed effects. This is to control for heterogeneous keyword-specific effects of keyword characteristics including the length of a keyword and popularity on click probability. Under this richer specification, we obtain the new estimates of position-specific click rates and quality scores, given which the valuations are estimated. The new results for the value distribution, bid shading, click rates, and quality scores are very similar to our baseline results except for '*cruise*' and '*car insurance*'. The new position-specific click rates for these categories are lower than the baseline results so bid shading percentages become also smaller as shown in Tables D.I-D.II. The quality scores are overall estimated to be higher than the baseline estimates in all the categories. The other 3 categories, however, still show almost identical bid shading percentages to the baseline results as the position-specific click rates remain almost the same. These results again confirm the importance of position-specific click rates for bidding behaviors.

D.2.1. *Larger samples with full sets of keywords*

We restrict the data set to the top 10% most popular keywords only to make sure that the underlying valuation distributions across markets are similar. Here we consider the full sets of keywords in our estimation. This empirical strategy results in much larger sample sizes so that the underlying valuation distributions are more precisely estimated. Note that

TABLE D.I

POSITION SPECIFIC CLICK RATES WITH A FULL SET OF KEYWORD DUMMIES

Ad-position	Cruise	Car Insurance	Laptop	Cable TV	Coins
2^{th}	0.509	0.614	0.690	0.460	0.733
3^{th}	0.395	0.480	0.586	0.390	0.633
4^{th}	0.093	0.313	0.442	0.222	0.653
5^{th}	0.054	0.243	0.353	0.164	0.517
6^{th}	0.001	0.298	0.249	0.124	0.429
7^{th}	0.013	0.237	0.221	0.098	0.243

Note: The click rate of the first position is 1 due to re-scaling and hence omitted in the table.

TABLE D.II

QUANTILES OF BID SHADING PERCENTAGE WITH A FULL SET OF KEYWORD DUMMIES

	Cruise	Car Insurance	Laptop	Cable TV	Coins
25%	0.128%	0.001%	0.010%	0.010%	0.078%
50%	0.578%	0.006%	0.048%	0.029%	0.414%
75%	3.082%	0.032%	0.412%	0.172%	1.900%
90%	6.859%	0.132%	3.217%	1.623%	8.367%
99%	18.848%	10.977%	34.898%	17.449%	54.508%

Note: The table summarizes the quantiles of bid shading in terms of the percentage of the corresponding estimated values.

the included keywords are now much more heterogeneous and hence the common value distribution assumption is more likely to be violated. Nonetheless, this exercise can shed some light on keyword heterogeneity. We expect that the popular keywords correspond to higher valuations than less popular keywords.

Tables D.III-D.V show the position-specific click rates, quantiles of quality scores, and quantiles of bid shading percentages estimated on the full sample. Figure D.1 compares the empirical distributions of the estimated valuations for the popular keywords and the full sample. The estimated bid shading percentages on the full sample are overall smaller than those for the popular keywords. This is partly due to the lower position-specific click rates. The full sample contains many keywords that are rarely searched by consumers. The results

TABLE D.III

POSITION SPECIFIC CLICK RATES WITH THE FULL SAMPLE

Ad-position	Cruise	Car Insurance	Laptop	Cable TV	Coins
2^{th}	0.577	0.580	0.584	0.421	0.653
3^{th}	0.430	0.419	0.430	0.346	0.497
4^{th}	0.186	0.233	0.371	0.152	0.532
5^{th}	0.115	0.171	0.215	0.093	0.377
6^{th}	0.041	0.179	0.088	0.041	0.255
7^{th}	0.022	0.111	0.035	0.013	0.055

Note: The click rate of the first position is 1 due to re-scaling and hence omitted in the table.

TABLE D.IV

QUANTILES OF QUALITY SCORE ACROSS PRODUCT CATEGORIES

	Cruise	Car Insurance	Laptop	Cable TV	Coins
25%	0.034	0.023	0.020	0.035	0.017
50%	0.037	0.028	0.024	0.039	0.021
75%	0.040	0.032	0.029	0.041	0.026
90%	0.048	0.044	0.035	0.045	0.032
99%	0.083	0.082	0.063	0.073	0.102

Note: The table summarizes the quantiles of the estimated quality scores. Quality score is the advertiser-specific effect on click probability.

show that ads related to unpopular keywords are less likely to be clicked when they are displayed. Another reason for less bid shading is the lower weighted value. Quality scores and valuations are overall lower on the full sample. As advertisers with higher weighted values can shade their bids more, the overall bid shading is smaller on the full sample.

The results on the full sample clearly show heterogeneity across keywords. The valuation distributions of the popular keywords are strictly below their counterparts from the full samples. The gaps are larger for '*cruise*' and '*car insurance*' categories. This implies that keyword heterogeneity is more outstanding in these categories. We speculate that the two categories have more diverse underlying products in terms of price. For instance, the price

TABLE D.V

QUANTILES OF BID SHADING PERCENTAGE WITH THE FULL SAMPLE

	Cruise	Car Insurance	Laptop	Cable TV	Coins
25%	0.105%	0.002%	0.010%	0.009%	0.070%
50%	0.684%	0.007%	0.047%	0.029%	0.445%
75%	3.125%	0.032%	0.349%	0.149%	2.272%
90%	11.852%	0.208%	3.035%	1.360%	9.872%
99%	33.871%	11.787%	25.818%	11.957%	41.889%

Note: The table summarizes the quantiles of bid shading in terms of the percentage of the corresponding estimated values.

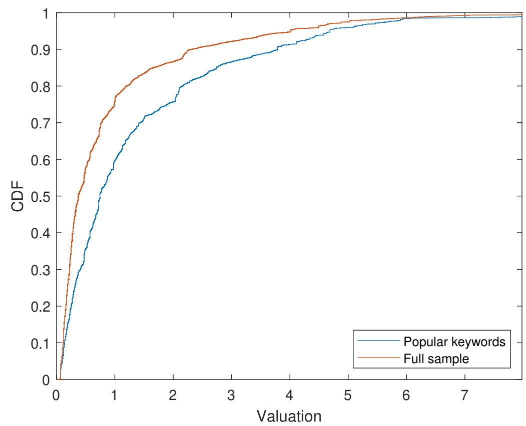
gap between 'luxury cruise' and 'budget cruise' would be much larger than the gap between 'high-end laptop' and 'budget laptop.' These results support our baseline empirical strategy that focuses on popular keywords.

APPENDIX E: ITERATIVE PROCEDURE FOR THE EQUILIBRIUM BID

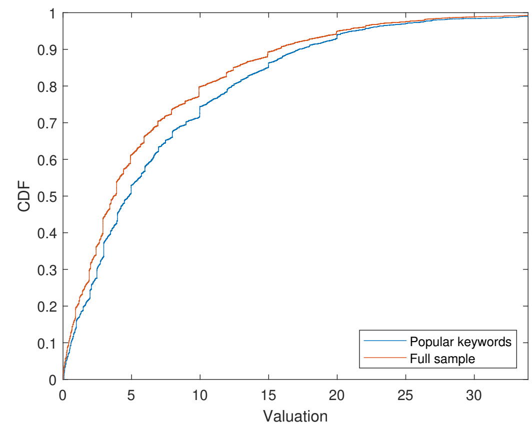
This section explains the iterative procedure used in our counterfactual analysis to approximate the equilibrium bid function. From Equation (25) in the proof of Theorem 1, one can derive the following equation:

$$b_w(\omega) = \omega - \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} \delta(\omega|b_w, F_w, f_w, N, k)}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} F_w(\omega)^{N-k-1}(1-F_w(\omega))^{k-2} \Big[(N-k)(1-F_w(\omega)) - (k-1)F_w(\omega)\Big]}$$

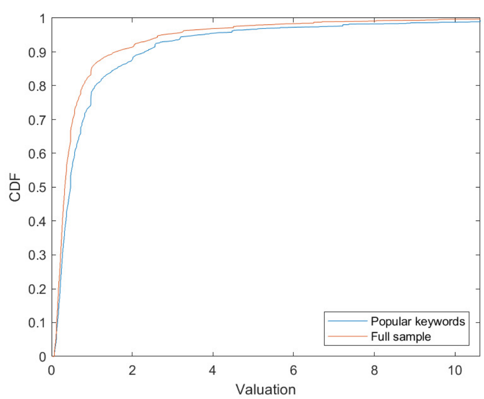
where $\delta(\omega|b_w, F_w, f_w, N, k) = b_w(\omega)F_w(\omega)^{N-k} - \int_0^\omega b_w(x)(N-k)F_w(x)^{N-k-1}f_w(x)dx$. We do not know about b_w so we start with our initial guess $b_w^{(1)}(\omega) = \omega$. We plug the initial guess in the RHS of the above equation and check whether the resulting function is sufficiently close to the guess. If not, we update our guess and repeat the step. More formally, for each n stage, using the guess $b_w^{(n)}$, we compute the RHS of the equation to



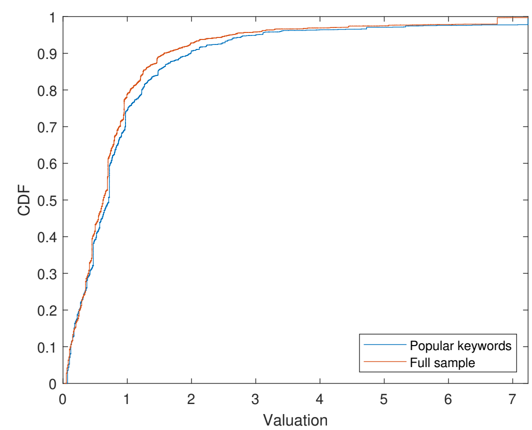
(a) Cruise



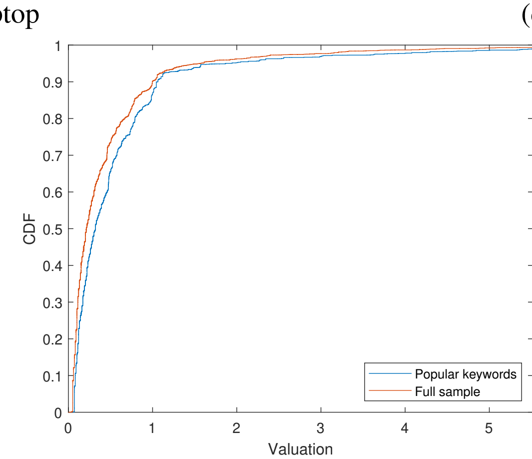
(b) Car Insurance



(c) Laptop



(d) Cable TV



(e) Coins

FIGURE D.1.—The valuation distributions for the top 10% popular keywords and the full sample.

obtain $b_w^{(n*)}$ as follows.

$$b_w^{(n*)}(\omega) = \omega - \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1-F_w(\omega))^{k-2} \delta(\omega | b_w^{(n)}, F_w, f_w, N, k)}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} F_w(\omega)^{N-k-1} (1-F_w(\omega))^{k-2} \Big[(N-k)(1-F_w(\omega)) - (k-1) F_w(\omega) \Big]}$$

At the end of each stage, we check whether $\sup_\omega |b_w^{(n*)}(\omega) - b_w^{(n)}(\omega)| < \varepsilon$ for some small $\varepsilon > 0$. If the condition is met, we stop and use $b_w^{(n*)}$ as the approximated equilibrium bid function. Otherwise, we update our guess $b_w^{(n+1)} = ab_w^{(n)} + (1-a)b_w^{(n*)}$ for some $a \in (0,1)$ and iterate the steps until convergence. As $n \to \infty$, $b_w^{(n)}$ converges to β_w from above. We set $\varepsilon = 10^{-4}$ and $a = 0.9$. In every case we consider, the convergence is quickly achieved with n less than 100.

APPENDIX F: IDENTIFICATION OF PRODUCT CATEGORIES

The raw dataset we use in the paper has multiple product categories, namely *cruise, car insurance, laptop, cable*, and *coins*. Additionally, the keywords are declassified and thus the product categories are also declassified (stated by numbers 0-4).[26] To overcome this problem, we analyze the differences in the categories and match each of them to the closest possible category among *cruise, car insurance, laptop, cable*, and *coins* according to the observed features. Table F.I gives a summary of how variables differ across categories. This table also shows the corresponding means of all features in different categories.

First, look at the features of category 1. This category is the easiest to identify because there are no keywords with one word. The base keyword consists of two words which must be '*car insurance*'. We can see that all the keywords in this category share the same two-word base keyword. This category is characterized by the very high average bid as well as the relatively small number of competitors compared to the other categories. This is consistent with the *car insurance* category. They are known to be the industry with one

[26]The categories are identified through the base keywords. The data has four single-word base keywords which identify the four categories, and one category is identified by a two-word base keyword.

of the highest prices per click. This is due to the high profit margins in the auto insurance industry which is a highly concentrated market.

The next category that stands out is category 2, which is characterized by a high number of advertisers and a high number of search queries per day. Due to its high volume of consumer searches, this is likely a consumer good. Therefore, it is closest to the '*laptop*' category as that is the only consumer good category in the data. Another category that is easy to identify is category 0. This category has a high number of advertisers and a high average click-through rate. A key feature of this category is more detailed searches that have longer keyword lengths. This is again a popular category with detailed search, and thus it is best matched with the '*cruise*' category.

Category 4 has the lowest average bid and search volume, as well as the fewest number of advertisers. It is most likely to be the least popular category in the data and thus likely '*coins*'. Lastly, '*cable TV*' is also a less popular and less expensive category but it is relatively more popular than '*coins*'. Therefore, we match '*cable TV*' with category 3. The table below summarizes the findings. Although these claims are just from our speculation, we use this classification for the main analysis in the paper. Even if there is some error in identifying the categories, we can still use the features of the category and interpret how and why the results might differ for categories with different features.

TABLE F.I

FEATURES OF DIFFERENT CATEGORIES

Category	**Description**	**CTR** (%)	**bid** (cent)	**adv**	**search** (million)	**length**
Cruise (Cat 0)	high competition & detailed search	1.28	0.51	6223	1320	3.18
Car insurance (Cat 1)	highest bids & high concentration	0.44	3.59	3815	2509	3.75
Laptop(Cat 2)	popular & high competition	1.33	0.45	4764	2913	3.03
Cable (Cat 3)	less popular & high bids	0.64	0.77	4703	1874	3.02
Coins (Cat 4)	low value across variables	1.36	0.36	3330	784	2.83

Note: This table shows the summary of key features of product categories in the raw data set. 'CTR' is the average click-through rates, 'bid' is the average bid, 'adv' is the number of advertisers, 'search' is the daily average search volume, and 'length' is the average word counts across keywords within each category.