CESifo WORKING PAPERS

10299
2023

February 2023

Sponsored Search Auction and the Revenue- Maximizing Number of Ads per Page

Pallavi Pal

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10299*

Sponsored Search Auction and the Revenue-Maximizing Number of Ads per Page

Abstract

In this paper, I derive a new method to identify the distribution of the advertiser's ad-value in the sponsored search auction, explicitly looking at weighted Generalized Second Price auction (GSP^w henceforth). Compared to previous literature, this method incorporates a weaker and more realistic assumption of 'incomplete information' on advertisers' private information. Additionally, I evaluate how much the advertisers shade their bid below their value, defined as bid shading amount. The results show that the bid shading is very small; the 50th percentile of the bid shading upper bound is below by 0.2% of their value. The low amount of bid shading is due to high competition intensity in the online ad market as the number of competing bids in the online ad market is very large. The bid shading calculation can also shed light on how the change of ad auction will impact the ad revenue.

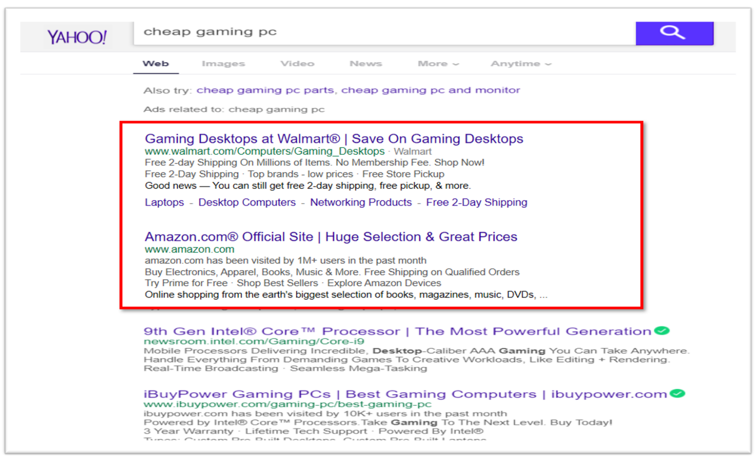
Pallavi Pal
Business School
Stevens Institute of Technology / NJ / USA
ppal2@stevens.edu

I am grateful to Jeitschko D. Thomas, Kyoo il Kim, Hanzhe Zhang, Roger J. Calantone, Michael Conlin, Jay Pil Choi, Chenhui Guo, Ashish Agarwal, David Croson and Vivek Bhattacharya for all the useful discussions and suggestions. I would also like to thank the Yahoo! research lab for providing the dataset used in this paper. I am grateful for comments by participants at Michigan State University, the Midwest Econometric Group, the American Marketing Association conference, Boston College and Grand Valley State University.

1 Introduction

Online advertising is a new but rapidly growing market. The rapid growth has led to a high volume of ads processed daily. For instance, Google delivers an astounding 29.8 billion ad impressions every single day.[1] Due to the high volume, online platforms use the auction as a method to price and sell ads. In this paper, I specifically look at the auction used for search ads. Search ads are paid search links that appear above generic[2] search results on major search engines such as Google, Yahoo! and Bing. Figure 1 gives an example of the Yahoo! result page.

Figure 1



Search ads use a Generalized Second Price (GSP henceforth) auction to sell ads. The ads are sold as contingent objects where the advertisers only pay if the consumer clicks on the ad. In 2005, the auction was modified to accommodate the difference in the probability of getting a click across ads. The click probability of each ad was calculated and termed as ad's '*quality score* .'The modified auction was referred to as weighted GSP as it uses bids weighted by their quality score. In this paper, I develop and analyze the theoretical and empirical model of advertisers' behavior in the weighted GSP auction and apply this model to Yahoo's ad dataset. Earlier papers in the literature such as Varian (2007) look at a perfect information case where the advertiser knows the bid and quality score of other advertisers. Recent papers have relaxed the perfect information assumption by introducing uncertainty in the entry or by looking at the simpler non-weighted GSP

[1] Google stats - https://venturebeat.com/2012/10/25/30-billion-times-a-day-google-runs-an-ad-13-million-times-it-works/

[2] Generic search results are non-sponsored links.

auction.[3] I relax the perfect information assumption by looking at an incomplete information case, where the advertiser only knows the distribution of other advertisers' value and quality score. The incomplete information equilibrium was first derived in the paper Gomes and Sweeney (2014) for the non-weighted GSP auction. This paper extends the literature by empirically looking at the incomplete information case in weighted generalized second price auctions. In general, auction models assume that bids are customized for a single auction.

The method derived can be used to infer advertisers' valuation from their observed bid and to compute counterfactual equilibrium outcomes for various economic questions. This paper applies the method to look at how much do advertisers shade their bid below their value, defined as bid shading.

There are multiple reasons why we would be interested in the bid shading amount in the online ads market. The online ad market is relatively new and has introduced a few unique features, such as a high volume of ads processed every day. The auctions used, such as GSP auction, are also relatively new auctions (GSP was first applied at Google, and subsequently theoretical literature on it began in 2007(few early papers are Varian (2007), and Edelman et al. (2007)). Thus, the literature on GSP auctions is less developed than other auction designs. For example, the BNE solution for incomplete information was solved in 2014 by Gomes and Sweeney (2014). Thus, apart from solving the BNE in the weighted version of GSP, i.e. GSP^w. It would also be interesting to see how much advertisers shade their bids when bidding in online ad auctions. This analysis is also essential as the ad market has recently seen a change in the type of auction used for online ads. Google has moved to first-price auctions for several ad offerings baring the search ads. The trend indicates that we might expect a change in search ads soon too. This paper contributes to this discussion by looking at how different the bids would be if we move from GSP^w to GFP^w. The main component that would vary in this regard is the bid shading in the two auctions. We provide the first part of this analysis by showing that the bid shading is minimal in the GSP^w auction due to the high number of advertisers – a unique feature to the online ad auction.

For estimation, I set up a structural model that accommodates the effect of consumer's click

[3] for example, Athey and Nekipelov (2010) looks at entry uncertainty and Gomes and Sweeney (2014) theoretically proves the equilibrium in the non-weighted GSP auction

behavior on the advertiser's equilibrium bid. The estimation is done in two steps. In the first step, I estimate the parameters that affect a consumer's click decision using a weighted logit model.[4] In the second step, I show how the partial identification method used in Haile and Tamer (2003) can be used here to derive bounds on the distribution of the advertiser's unobserved ad valuation.

The data set is provided by the Yahoo! research lab, with approximately 51 million observations. It covers all ads[5] displayed on Yahoo! search result page, over four months, for five major categories: *laptop, TV cable, cruise travel, collectible coins, and car insurance.* I have information about the number of displays, the number of clicks, ad description, and the ad position.[6] The data also provides the bids for each ad. Additionally, I can measure the number of advertisers per day and the number of ads per page.[7] Notice that the data does not provide information on the ad's quality score; I solve this limitation by estimating the quality score of an ad in terms of the advertiser's effect on consumer's click probability.

The findings show that the method provides close bounds on the distribution of advertisers' per click value. Apart from calculating the bid shading amount in monetary terms, we also calculate the bid shading amount as a percentage of the advertiser's value. For all product categories, 90% of the advertisers, the difference between their bid and value is less than 1.5 cents. Car Insurance and Coins seem to have the highest level of bid shading while laptop has the lowest. Although the bid shading mount is interesting to analyze, we need to be careful in interpreting the dollar amounts as the bid was re-scaled in this data in order to mask the actual bid amount. Thus, the bid shading in terms of value percentage is a more precise estimate. Similar to the monetary value, the bid in percentage term is also very low. For all categories, the 50^{th} percentile was below 0.2% of ad value. Car insurance and Coins have a higher percentage of bid shading compared to others.

This paper contributes to the work on estimating the unobserved advertisers' ad value using the equilibrium bid. In the theory literature, Edelman et al. (2007) (referred to as EOS) and Varian (2007) were among the first to derive the equilibrium bid. Although online ad auctions have received great attention in the theoretical literature, empirical research remains sparse; partially due

[4]For categories where the weighted logit does not converge, I use a linear probability model.

[5]The data excludes ads that appear for the search of brand names.

[6]ad description is measured as the words provided by the advertisers about the ad measured as keyword

[7]the number of ads on the first page is assumed to be seven ads unless observed less than seven ad positions. This is a common assumption made for papers using this data set from yahoo such Agarwal and Mukhopadhyay (2016)

to limited public data available on this topic. Börgers et al. (2013) analyzes Yahoo! data to estimate ad-position dependent value, and Yang et al. (2013) structurally estimate EOS's model. Athey and Nekipelov (2010) propose and estimate a structural model tailored to features of sponsored search auctions run by US search engines (such as Google or Microsoft).[8]Earlier papers in the literature such as Varian (2007) look at a perfect information case where the advertiser knows the bid and quality score of other advertisers. Recent papers have relaxed the perfect information assumption by introducing uncertainty in entry or by looking at the simpler non-weighted GSP auction.[9] A key motivation of this paper was to empirically estimate the advertiser's ad value under weaker information assumptions, specifically looking at the case of incomplete information. Gomes and Sweeney (2014) theoretically solved the equilibrium bid in the incomplete information case for a non-weighted GSP auction. This paper extends their work by looking at the incomplete information case in weighted GSP auction, which has a multi-dimensional type. The extension is nontrivial as the weight introduces a multi-dimensional type of advertiser. As pointed out in Gomes and Sweeney (2014), the extension to multi-dimensional private information is an important and challenging next step. Additionally, this paper provides a way to estimate the ad value. Thus, the paper further contributes to this literature by providing closed-form bounds on the equilibrium bids that give us partial estimates for the advertiser's ad value.

Additionally, this paper is related to econometric theory papers on partial identification methods. The methodology in this paper closely follows a method first proposed in Haile and Tamer (2003). Their paper shows how to estimate bounds on the distribution of object value in an English auction. I extend it and show how to apply the method in an online auction, i.e., a Generalized Second-Price auction.

The remainder of the paper is structured as follows: section 2 gives an overview of the market and presents the theoretical model, section 3 gives details about the data, section 4 specifies the econometric method, section 5 gives the results. Finally, section 6 summarizes the findings and discusses the broader consequences of this paper.

[8]Specifically, they accommodate uncertainty in advertisers' perceptions (due to randomness in an advertiser's quality score over time, as well as in the set of competitor bidding in the auction at any time).

[9]for example, Athey and Nekipelov (2010) looks at entry uncertainty and Gomes and Sweeney (2014) theoretically proves the equilibrium in the non-weighted GSP auction

2 Market Environment and Theoretical Model

In this section, I present a model of online advertising. Subsection 2.1 and 2.2 set up the ad market by explaining both the consumer and the advertiser's side, respectively. The consumer's model derives the click rate and quality score that is used in the equilibrium bid equation. Subsection 2.3 derives the equilibrium for the GSP^w auction, the theoretical bounds on the value distribution, and bounds on the advertiser's bid shading amount.

2.1 Consumers

Each consumer i has a unit demand for a product/service and consequently starts the search by putting a query on an online search engine. Once the result page displays all links related to the search query, the consumer clicks on all relevant links and purchases a good or service from one of the clicked links. In this section, I model the consumer's click decision.

The online environment motivates several considerations. Firstly, the consumer anticipates the derived click benefit by visible characteristics of the ad. Along with the visible ad characteristics, I assume that the consumer also believes that ads at a higher ad position are of higher quality and relevance. This belief probably stems from the consumer's observing that the search engine's algorithm assigns a higher ad-position to ads with higher quality scores, ceteris paribus. I also find evidence in the data to support this assumption. Another consideration is that each click requires the consumer to spend considerable time on it, which can be thought of as a search cost or time cost. Thus, we will add consumer characteristics and ad-position as variables that affect consumers' utility from clicking on an ad.[10] Further details about the variables impacting the consumer's utility is provided in the empirical section. Let the expected utility of consumer i receives from clicking in ad j is given as $U_{i,j}$. The following proposition shows consumers' click behavior.

Proposition 1. *Consumers in equilibrium may click on multiple ads per page.*

Essentially in the equilibrium, consumers click on all ad links where the benefit of a click is more than the search cost.[11] Let $y^*_{i,j}$ denote the binary variable capturing consumer i's equilibrium

[10]Note that there are other variables as well that might impact the click decision, such as generic links on the page. However, due to data limitation, we can not capture their impact on the click utility

[11]although most of the literature assumes a single click per page, multiple clicks per consumer is a more realistic

click decision for ad j , with $y^*_{i,j} = 1$ if consumer decides to click on the ad. Then the click decision can be written 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 in the empirical section to estimate consumers' probability of a click.

The estimation gives us the predicted probability of a click for an ad j in ad-position k, represented by $s_j \times c_k$. Assuming that the ad and ad-position effects are separable[12] the click probability can be rewritten as:

$$\text{Click Probability} \rightarrow s_j \times c_k \qquad \forall\ j \ \epsilon \mathscr{J} \ \&\ \text{k} \ \epsilon \ \{1, 2, ...K\} \tag{1}$$

where

s_j: The effect of advertisement j on probability of a click.

c_k: The effect of ad-position k on probability of a click. The set of click rate for all ad positions on a result page is denoted as $\mathcal{C} = \{c_1, ...c_K\}$. The click probability in equation(1) is used in advertiser's maximizing problem. In the empirical section, we will discuss how to estimate the components of the click probability, namely c_k and s_j.

2.2 Advertisers

Each advertiser denoted by $j \ \epsilon \ \mathscr{J} := \{1, ...J\}$ has an ad value, v_j, independently drawn from a distribution, F_v, with support $[\underline{v}, \bar{v}]$. In addition to the ad valuation, each advertiser has a quality score, s_j, drawn independently from a distribution F_s with support $[\underline{s}, \bar{s}]$.[13] The 'quality score' signifies the advertiser's click probability, ceteris paribus. Thus, advertiser j's type is given by both the ad value and the quality score, i.e. (v_j, s_j). Additionally, lets define the *weighted value* as the product of the value and the quality score of the advertiser. The weighted value is denoted as

situation in this market. This is evident by a new feature on Bing, which gives an option of opening a new tab every time you click on a link. Here is the link to the article: https://searchengineland.com/bing-is-testing-an-open-in-new-window-icon-in-the-search-results-301922

[12]This is a similar assumption adopted by various papers in the literature for identification of the quality of advertiser

[13]In this model, I assume that the the per click value, v_j and advertiser's quality score s_j are independent of each other.

$\omega_j \equiv (s_j \times v_j)$, $\omega \sim F_w(.)$.[14] Let the potential number of advertisers be denoted by $N \epsilon \{1, 2, ...\mathcal{N}\}$. Ads are sold through an auction mechanism.

Auction Setup

A single auction is held for each search query to sell all the ad positions on the search results page. Consider a standard symmetric independent private value paradigm, in which for each auction, there are K ad-positions to be auctioned and N potential advertisers. These auctions sell contingent ads, which means that advertisers only pay for the ad position if a consumer clicks on the ad. Thus, all ad-related terms such as bids, valuations, and prices are defined on a per-click basis (for ease of notation, the per-click ad value, per click ad bid, and per-click ad price are hereafter referred to as value, bid, and price).

The K ad-positions on a result page are allotted through a single GSP^w auction. The bidding strategy for an advertiser j consists of a bid function defined as $b_j = b(v_j, s_j)$, where $b(v_j, s_j)$ denotes the bid for j^{th} advertiser given that the advertiser's value is v_j and quality score is s_j.

Recall that the advertiser only pays for the ad if a consumer clicks on it as the price is per click. Thus, the auctioneer's revenue increases with the ad's click probability apart from the price. To accommodate the effect of click probability on revenue, the auctioneer uses weighted bids instead of bids. Bids are weighted by the advertiser's impact on click rate, captured by the quality score, s_j. The weighted bid is denoted as follows

$$b_w(v_j, s_j) \equiv b_{j,w} = b_j \times s_j$$

Let us denote the distribution of the bid as G_b and the distribution of the weighted bid as G_w. The weighted bids are ranked in descending order. The ordered weighted bids are then used for allocating and pricing the ad positions.[15] The order statistic of the weighted bid is denoted as $b_w^{[l]}$, which represents the l^{th} highest order statistic of the weighted bid. Additionally, $b_{-j,w}^{[k]}$ denotes the k^{th} highest weighted bid among all advertisers except j.

[14] The defined weighted value would be later used to reduce the dimensions of advertiser type from two dimension (v_j, s_j) to single dimension (w_j). Essentially lemma(1) shows that weighted value, w_j is sufficient to capture the advertisers type and its affect of the equilibrium bid.

[15] In the data, the quality score is calculated to capture the impact of an ad on click rate. However, I assume one ad per advertiser in each auction market in this paper. Thus, ad and advertiser scores are interchangeable.

The auction assigns ad positions in descending order of advertisers weighted bids. Essentially, allotting k^{th} ad-position to the advertiser with k^{th} highest weighted bid. For example, the top ad-position goes to the advertiser with the highest weighted bid; the second ad-position goes to the second-highest, and so on. The price paid varies depending on the winning ad position and the quality score of the advertiser. Specifically, if an advertiser wins the k^{th} ad-position, his price would be equal to the $[k+1]^{th}$ highest weighted bid divided by his quality score.[16] The auction rules can be summarized as follows

$$\begin{aligned} &\text{ad-position } k \text{ allotted to advertiser } j \text{ if} \quad b^{[k]}_{-j,w} \leq b_{j,w} \leq b^{[k-1]}_{-j,w} \\ &\text{Price for ad-position } k \text{ (given to adv } j\text{):} \qquad p_k = \frac{b^{[k+1]}_{w}}{sj} \end{aligned}$$

Here $b^{[k]}_{-j,w}$ is the k^{th} highest weighted bid among advertiser $j's$ competitors and $b^{[k+1]}_{w}$ gives the $[k+1]^{th}$ weighted bid among all advertisers. As the name suggests, this auction design is a general version of the second-price auction. However, it is important to note that, unlike the second-price auction, advertisers do not bid their value in the GSP^w auction. The reason is that, in the GSP^w auction, an advertiser can gain from bidding less than his value as the bid affects the winning probability and impacts the price paid. This incentive is absent in the second-price auction as the bid only impacts the allocation, not the price paid.

Observation 1. *An important feature of this auction is that advertisers do not bid for a single ad position; instead, they submit a single bid for all the ad positions on the result page. Furthermore, the auction allocates the ad positions in the descending order of advertisers' weighted bids.*

Main Assumptions:

I will now introduce the two main assumptions used for deriving the bounds on the value distribution.

Assumption 1. *The information set assumes the standard incomplete information case.*

For deriving the bounds and the equilibrium bid, we need to state the information available to advertisers when they place their bids. Assumption 1 explains that, in this model, we assume

[16] I assume no reserve price for simplicity. The reserve price was not reported in the data, and in the time period used for the study, yahoo had a fixed reserve price that did not change throughout the data time period.

that the advertisers have *'incomplete information'* about the market variables. An incomplete information case is a more realistic and weaker assumption in the auction literature than in other cases, such as perfect information or complete information.

Advertiser's information setup This model looks at the incomplete information case. Each advertiser knows their type, i.e., (v_j, s_j) but does not know other advertisers' bids, quality scores, or values. They only know the primitive distributions, namely the value distribution F_v, the quality score distribution F_s, and the weighted value distribution F_ω. Apart from this, the number of advertisers (N) and ads per page (K) are common knowledge.

Assumption 2. *Advertisers' weighted bids in the GSP^w auction are strictly increasing in the advertiser's weighted value.*

Assumption 2 is intuitive, and as we will see later, it would be necessary to guarantee the existence of an equilibrium.[17] It is important to mention here that the equilibrium bid exists in the incomplete information case. However, the bid does not have a closed-form, and it is an N-P hard problem to estimate the value from the equilibrium bid equation empirically. Thus, in this paper, I use inequalities to bound the value in terms of observed data to partially estimate the bounds on the value distribution. We can essentially study the market under a weaker – and realistic assumption (i.e., incomplete information) by looking at partial identification instead of point identification.

Profit maximization problem

Let us define advertiser j's profit from winning ad position k as $\pi_{k,j}$, which can be expanded as follows:

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

where v_j denotes the ad value of advertiser j and $p_{j,k}$ is the price paid by advertiser j for winning ad-position k.[18] The term $c_k * s_j$ denotes the probability of a click on ad j in ad-position k; recall that s_j is the quality score, i.e. advertiser's affect on click probability and c_k is the effect of the

[17] See appendix(8) for details on equilibrium bid

[18] Recall that the price paid depends not only on other advertisers' weighted bids and the ad position but also on the winning advertiser's quality score. Thus the price for the same ad position may vary across advertisers

ad position on the click rate. Here the click probability term $c_k * s_j$ assumes, similar to previous literature [Varian (2007),Athey and Nekipelov (2010)], that the click probability is multiplicatively separable in the effect of the ad-position and the advertiser.[19]

Note that till now, we have defined the profit per ad position. However, from observation 1 we know that advertisers submit a single bid to win one of the available ad positions on the search result page. Thus, the equilibrium bid maximizes the expected profit, given as the sum of the profit from each ad-position times the probability of winning that ad-position. As shown below:

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

using eqn 2, we get

$$\rightarrow \Pi(b_j; v_j, s_j) = \sum_{k=1}^{K} Prob(b^{[k]}_{-j,w} \leq b_{j,w} \leq b^{[k-1]}_{-j,w})\Big((s_j \times c_k)\{v_j - \mathbb{E}[p_{j,k}|b^{[k]}_{-j,w} \leq b_{j,w} \leq b^{[k-1]}_{-j,w}, s_j]\}\Big) \tag{3}$$

The expected profit from ad-position k is the product of the probability of winning ad-position k and the expected profit from winning ad-position k. The probability of winning is equal to the probability that the weighted bid b^w_j is less than $(k-1)^{th}$ highest weighted bid and more than k^{th} highest weighted bid, i.e. $b^{[k]}_{-j,w} \leq b_{j,w} \leq b^{[k-1]}_{-j,w}$.

2.3 Equilibrium Analysis

The main interest of this paper is to estimate the distribution of advertisers' value, $F_v(.)$. The value distribution characterizes advertisers' willingness to pay for an ad and is used for counterfactual analyses. For instance, in the application section, we will use the estimated value distribution to compare bids in the GSP^w auction to the bids in GFP^w auction. In this section, I theoretically derive the bounds on the ad value.

Lower bound on value distribution

To obtain the lower bound on the value distribution $F_v(.)$, we need to look at the equilibrium bid. In this section, I first discuss the profit-maximizing objective function and then derive theorem 1,

[19]Note, I assume each ad is separately optimize, thus each ad is treated from a separate advertiser. This is similar to the assumption made in paper Athey and Nekipelov (2010)

which defines the inequality used for the lower bound.

Using equation(3), we can show that the equilibrium bid maximizes the following objective function:

$$b(v_j, s_j) = Arg\max_{\hat{b}} \sum_{k=1}^{K} Prob(b^{[k]}_{-j,w} \leq \hat{b} * s_j \leq b^{[k-1]}_{-j,w}) s_j c_k \left[v_j - \mathbb{E}\left(\frac{b^{[k+1]}_w}{s_j} \middle| b^{[k]}_w = s_j * \hat{b} \right) \right]$$

In the standard auction, this is solved by inverting the bid and using the value distribution. However, in this case the weighted bid $b_{j,w}(s_j, v_j))$ is multi-dimensional as it depends on the value v_j as well as the score s_j.

This paper overcomes the problem of the non-invertible weighted bid by proving equivalence between a GSP^w auction and a *non-weighted* GSP auction where the advertisers' value is equal to the weighted value. Specifically, lemma 1 shows that for any advertiser j, the equilibrium weighted bid in the GSP^w auction is equivalent to his equilibrium bid in GSP auction in which his value is replaced by the weighted value, ω_j. In other words, the bid in GSP serves a similar purpose as the weighted bid in GSP^w auction. The difference is that in GSP auction, the bid can be written as a single dimension function given as follows:

$$b^{GSP}_w(\omega_j) \rightarrow \mathscr{R}_+, \text{ where } \omega_j = s_j \times v_j$$

$b^{GSP}(\omega_j)$. In the next lemma, I formally prove 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, i.e. $b^{GSP^w}_w(v_j, s_j)$, is equivalent to the equilibrium bid function in a GSP auction where the value is replaced by the weighted value, i.e. $b^{GSP}_w(\omega_j)$.*

$$b^{GSP}(\omega_j) = b^{GSP^w}_w(v_j, s_j), \qquad \forall j \text{ where } \omega_j = s_j \times v_j$$

The lemma shows that the weighted bidding function is equivalent to a function that is dependent only on single-dimension i.e. advertiser's weighted value ω_j. Additionally, it shows that at equilibrium we can rewrite the weighted bid $b_w(v_j, s_j)$ as a function of only weighted value , i.e. $b_w(\omega_j)$. This simplification comes in handy for the proof of bounds and equilibrium. We can see below how it simplifies the maximization problem:

$$b_w(\omega_j) = Arg\max_{\hat{b}_w} \sum_{k=1}^{K} Prob(b^{[k]}_{-j,w} \leq \hat{b}_w \leq b^{[k-1]}_{-j,w})c_k\left[v_j - \mathbb{E}\left(\frac{b^{[k+1]}_w}{s_j}\middle| b^{[k]}_w = \hat{b}_w\right)\right]$$

inverting the probabilities in terms of weighted value using lemma(1) and assumption(2), we get

$$b_w(\omega_j) = Arg\max_{\hat{b}} \sum_{k=1}^{K} Prob(\omega^{[k]}_{-j} \leq b^{-1}_w(\hat{b}_w) \leq \omega^{[k-1]}_{-j})c_k\left[\omega_j - \mathbb{E}\left(b^{[k+1]}_w\middle| b^{[k]}_w = \hat{b}_w\right)\right] \tag{4}$$

Where $\omega^{[k]}_{-j}$ signifies the k^{th} highest weighted value among advertiser j's competitors. We can derive the equilibrium bid using lemma 1 and assumptions 1 & 2. However, as stated earlier, the equilibrium bid cannot be used directly to estimate the ad value as it is an N-P hard problem. Therefore, I have provided the equilibrium bid and the proofs in the appendix(8)). Next, we derive lower and upper bounds on the advertiser's value which we will use in the empirical section to derive the distribution bounds. We first derive the inequality used for the lower bound of the value distribution.

Observation 2. *It is important to note here that the upper (lower) bound on the CDF of the value distribution uses the lower(upper) threshold on the value.*[20]

In this theoretical section, we derive the inequality on the advertiser's value. In the empirical section, we derive the bounds on the value distribution from the theoretical inequality equations of the ad value.

Theorem 1. *Under assumption 1 and 2, advertiser's value can be bounded in terms of observed variables:*[21]

$$b_j\left(1 + \Phi(F_w, \omega_j, \mathcal{C}|K, N)\right) \geq v_j \quad \forall j \epsilon \mathcal{J} \tag{5}$$

where the term $\Phi(F_w, \omega_j, \mathcal{C}|K, N)$ depends on the click rate across ad-position, namely $\mathcal{C} = \{c_1, ...c_K\}$, the advertiser j's weighted value, the weighted value distribution F_w, the number of ads on the page K and the total number of advertisers N. The interpretation of $\Phi(F_w(.), c, b_{j,w}, K)$

[20] This is implied from basic properties of probability.

[21] proof in appendix(8)

is more clear once we rearrange equation 5 as

$$\frac{v_j - b_j}{b_j} \leq \Phi(F_w, \omega_j, \mathcal{C}|K, N) \quad \forall j \epsilon \mathcal{J} \tag{6}$$

Thus, $\Phi(F_w(.), c, b_{j,w}, K)$ can be interpreted as the upper bound on the bid shading amount represented in terms of the bid percentage. The term can be expanded as follows:

$$\Phi(F_w, \omega_j, \mathcal{C}|K, N) = \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1 - F_w(\omega_j))^{k-2} F_w(\omega_j)^{N-k}}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} \left[(N-k)(1 - F_w(\omega_j)^{k-1} F_w(\omega_j)^{N-k-1} - (k-1)(1 - F_w(\omega_j)^{k-2}) F_w(\omega_j)^{N-k} \right]}$$

Notice that the function $\Phi(F_w, \omega_j, \mathcal{C}|K, N)$ has variables that are not known to the econometrician such as the weighted value and the weighted value distribution F_w. However, we can substitute that with the bid distribution as the weighted bid is strictly monotonic. I will use the following equality conditions:[22]

$$G_w(b_w(\omega_j)|N) = F_w(\omega_j|N)$$

The above can be used to rewrite the function in terms of observed distribution of the weighted bid as follows:

$$\begin{aligned} &\Phi(G_w, b_{j,w}, \mathcal{C}|K, N) \\ &= \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} (k-1)(1 - G_w(b_{j,w}))^{k-2} G_w^{N-k}(b_{j,w})}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} \left[(N-k)(1 - G_w(b_{j,w}))^{k-1} G_w^{N-k-1}(b_{j,w}) - (k-1)(1 - G_w(b_{j,w}))^{k-2} G_w^{N-k}(b_{j,w}) \right]} \end{aligned} \tag{7}$$

Upper Bound on value distribution

To obtain the upper bound on the distribution F_v, we show that advertisers always bid less than their value, as shown in the theorem below

[22]Similar conditions were used in paper Guerre et al. (2000) to empirically estimate value distribution in first price auction

Lemma 2. *Advertisers do not bid more than their value for the ad-click.*

$$v_j \geq b_j \quad \forall j \tag{8}$$

The above lemma can be derived from the equilibrium bid, which shows that the advertisers shade their bid below their value. Equation (8) will be used to derive the upper bound on the value distribution. More details in the estimation section.

The inequalities in equation (8) and (5), reproduced below, will be used in the estimation section to put bounds on the value distribution.

$$b_j\left(1 + \Phi(G_w, b_{j,w}, \mathcal{C}|K, N)\right) \geq v_j \geq b_j \quad \forall j \epsilon \beth \tag{9}$$

where $\Phi(G_w, b_{j,w}, \mathcal{C}|K, N)$ is defined by equation(7). In the empirical section, we look at how to use the bounds on the ad value to partially estimate the value distribution.

Bid Shading

The degree by which the bid is shaded below the advertiser's value is known as the '*bid shading*' amount. It is the amount by which the bid is less than value.

$$\text{Bid Shading} \equiv v_j - b_j \qquad\qquad \forall j \epsilon \mathbf{J}$$

We will first try and bound the bid shading bounds using theorem 2. As can be seen, by equation 6 we can bound the bid shading amount as follows,

$$0 \leq v_j - b_j \leq \Phi(G_w, b_{j,w}, \mathcal{C}|K, N) \times b_j \quad \forall j \epsilon \mathcal{J} \tag{10}$$

The above equation gives us bounds on the bid shading amount in the GSP^w auction. The lower bound is trivial and the bound that we particularly interested in is the upper bound. Specifically, closer teh upper bound is to the zero the more information these bounds prpvide about the bidding strategy of advertisers.

Although the bid shading amount is interesting by itself, we can also look at the bid shading as a percentage of value. Recall that the actual bids are re-scaled in our data, thus the bid shading amount provides a more accurate information. We can bound the bid shading as a percentage of the advertiser's value as follows

Lemma 3. *The percentage increase in advertiser's bid compared to it value can be bounded as follows*

$$0 \leq \frac{v_j - b_j}{v_j} \leq \Phi(G_w, b_{j,w}, \mathcal{C}|K, N) \quad \forall j \epsilon \mathcal{J}$$

This follows directly from equation 10 and lemma 1.[23]

In the empirical section we will estimate the bounds. The results show that the bid shading is on average below 1% of the advertiser's value. The next step to this analysis woudl be to compare the bid shading between GSP^w and GFP^w.

3 Data

The data set is provided by Yahoo!.[24] The data covers all search queries for four months, from January 2008 to April 2008. The sample covers all search ads[25] in 5 categories, namely *Laptop*, *TV Cable, Cruise, Collectible Coins* and *Car Insurance*. Each category is used as a separate data set, and the results are obtained separately for each of them. The advantage of data from multiple product categories is that we can compare the results across product categories after the estimation to see whether the results are sensitive to product characteristics. The total observations in the raw data set are 77,850,272. After cleaning the data, we divide the data into two parts. The first part has consumer side information such as clicks and number of ad displays, and the second part has advertiser's side information such as bid and ad description. For this analysis, I limit my sample to ads on the first page of the search result.[26] For more information on data cleaning refer to appendix 8.3.

[23] please refer to appendix for the proof

[24] the data was provided as part of the *Yahoo! Research Alliance Webscope program*. The data was part of the Advertising & Markets Data and, more specifically ?A3. Yahoo! Search Marketing Advertiser Bid-Impression-Click data on competing Keywords

[25] Search of specific brands names are removed from the data.

[26] A similar restriction was followed in Athey and Nekipelov (2010)Athey and Nekipelov (2010), who use Bing data

Consumers side data: There are 51.7 million observations on the consumer side with information about ad display and the consumer's click response for each ad-position–advertiser–keyword combination. For every ad, the advertiser specifies a list of words related to the ad, which are referred to as keywords. The auctioneer uses these keywords to search queries to the most relevant ads. For example, an ad by Amazon in the laptop category will report that Amazon's matched keyword was 'cheap gaming laptop,' and it got 100 displays in the 1^{st} ad-position, which translated into five clicks. In the data, we have information for each day-id-keyword combination. Thus, to do analysis at the ad level, I define an ad as the set of keywords for which the advertiser had the same bid on a given day. There are on average 0.34 million consumer search queries per day in the mentioned five categories. Table(1) gives the list of variables used and Table(3) gives the summary statistics.

Using the clicks and display information, we can deduce the click rate for each ad, which is measured as the ratio of the number of clicks over the number of displays. The summary statistic shows an average of 0.016 click rate, implying that, on average, 1.6% of the ads got a click. Additionally, the keyword (matched words between ad and search) approximates the type of search. The data has $648,515$ unique keywords. The number of words in the keyword, referred to as keylength, can be used as an approximation for the length of the search query.[27] Previous papers in the literature have noted that longer search queries are typically associated with more focused search intent and can thus be more valuable for advertisers.[28] The summary statistics show the maximum value of keylength is 19, with an average of 3.3 words per keyword. Another variable used is the popularity of the keyword. The popularity of a keyword is measured by the daily count of ad displays that were matched using that keyword. This controls for the possible effect of the popularity of the search. On average, a single keyword matched 5055.59 ads per day.

Advertisers side: The data is aggregated on the id-day level, with 5.5 million observations. For each ad, I have information on the bid for the ad, the number of times the ad won an auction, the average winning ad position, and the total number of advertisers shown in a day.[29] Table 2

[27]Keylength can be used as a proxy for longer search as the longer keywords are more likely to be matched by longer search queries

[28]for instance, Ramaboa, Kutlwano KKM, and Fish, Peter (2018) look at differences in consumers with different search lengths

[29]I assume that the total number of potential ads is equal to the total number of ads that won at least once in a

lists the variables, and Table 4 provides the summary statistics. The bid is measured in terms of cents. To mask the actual amount, Yahoo! scaled all bids by an unknown amount. I subtract the bid with an amount close to the lowest value. Thus, the bid can be taken as the lower bound on the actual bid. The average bid is 4.9 cents. Through the data, I can measure how many times an advertiser had a winning ad in the auction and what was his corresponding bid.

4 Econometric Method:

In this section, I describe the estimation method. An auction in a given period is considered as a market, denoted by m. Thus, the variables will now have an additional dimension as the observations are recorded for multiple markets (i.e. multiple auctions over time). The estimation steps are as follows:

- Step M.1: Estimate consumers click probability.
- Step M.2: Estimate lower and upper bounds on the advertiser's value distribution.

4.1 Step M.1: Estimate consumer click probability

The main aim of this step is to derive ad-position and advertisers' effect on the click probabilities. The ad-position effect on click probability are used as a measure for click rate for each ad-position.[30] The advertiser's effect on click probability is used to create the measure for ad quality.[31] We use the utility detived from clicking on an ad from teh theory section to run a linear probability model. The conditional probability of getting a click can be denoted as

$$\mathbb{E}(y|x) = \rho_0 + \rho_1 dP_{i,j,m} + \rho_2 dA_j + \rho_2 z_i + \rho_3 z_m + \epsilon_{i,j,m} \tag{11}$$

where dP and dA represent ad-position dummy and advertiser dummy respectively. Apart from this the control variables include consumer specific variables (z_i) such as search popularity measure

day

[30]click rate is the expected percentage of clicks received in each ad-position

[31]Note: The quality score is calculated by the auctioneer (i.e. the search engine) and then reported to the advertiser. During the time period of the data used in this paper, the quality score captured the advertiser's impact on click probability. Thus, the assumption that the quality score is equal to the impact of advertiser on click probability is true in the data set. Note that the current quality score (used by Yahoo, Google, and Bing) uses ad display characteristics and other consumer characteristics to compute the quality score. Note that the results of this paper would still hold, even in the current definition of quality score.

as well as the keylength which captures how detailed is the search.[32] Lastly the term $\epsilon_{i,j,m}$ is the idiosyncratic shock. Using linear probability model, we can predict the probability that consumer i chooses to click ad j in market m. Specifically, the parameters estimated in step 1 are:

- Click rate of ad-position k ($\hat{c}_k$): This is measured as the predicted probability of a click in ad-position k. The click rate is then scaled to be between $[0, 1]$ by subtracting the minimum and dividing it by the click rate of position 1 in each market.
- Quality measure ($\hat{s}_j$): The quality measure used the predicted probability of a click for advertiser j, denoted by s_j. This measure is then scaled to be between $[0, 1]$ by subtracting the minimum and dividing it by the highest value.[33]

4.2 Step M.2: Estimate lower and upper bounds on advertiser's value distribution

This step involves estimating the distribution of advertisers value, $F_v(.)$.[34] The distribution is partially identified, implying that only the upper and lower bound on the value distribution is identified.

This step uses inequality equation for the ad value from equation 9, as reproduced below:

$$b_{j,m}\Bigg(1 + \Phi(G_w, b_{j,w,m}, \mathcal{C}_m|K, N)\Bigg) \geq v_{j,m} \geq b_{j,m} \quad \forall j \epsilon \daleth \tag{12}$$

where function $\Phi(.)$ depends on the click rate vector, namely $\mathcal{C}_m = \{c_{1,m}, ...c_{k,m}, ...c_{K,m}\}$, the advertiser j's bid, the bid distribution G_w, the number of advertisers N and the number of ads on the page K.[35] From the bounds on value in equation(12). Define $H_\phi(.)$ as the distribution of $b_j\Bigg(1 + \Phi(G_w, b_{j,w}, \mathcal{C}_m|K, N)\Bigg)$ and recall that $G_b()$ is the distribution of $b_{j,m}$. Similar to Haile and Tamer (2003), we can use equation(12) to imply the following first-order stochastic dominance

[32] search popularity is measured as the proportion of times the keyword appeared in the search result page relative to total search queries in the category

[33] Note we re-scaled the minimum to be 0.001 as zero quality score would mean advertiser did not have any incentive to bid and thus, making the observed bid inconsistent

[34] Note, that this step uses a non-parametric estimation method since in this step the goal is to estimate a distribution and not a parameter.

[35] We assume that the potential number of advertisers is the same across markets, which is observed as the full set potential advertisers

relation:

$$H_\phi(.) \leq F_v(.) \leq G_b(.) \tag{13}$$

I use kernel estimation to estimate the cdf, as shown below:

$$\hat{G}_b(h) = \frac{1}{\delta_h} \sum_j \sum_m 1\{\hat{b}_{j,m} \leq h\} K\left(\frac{\hat{b}_{j,m} - h}{\delta_h}\right) \tag{14}$$

Similar equation is used for the upper bound, giving the final estimate as:

$$\hat{H}_\phi \leq F_v \leq \hat{G}_b \tag{15}$$

5 Results

Advertiser's willingness to pay for an ad and bid shading

In this section, I analyze the advertiser's benefit from an ad. This step derives the distribution for advertisers exclusive ad value.

Advertiser's ad value: Using equation(12), I get the bounds on the advertiser's maximum willingness to pay for an ad or ad value. The distribution bounds are estimated for each category, as shown in the graph 3. The bounds are tight for all of the categories, implying that inequality is sufficient for inference. The tight bounds are also a result of the bid shading being very small as a result of high number of competitors in online advertising market. Using equation(6), table 6 summarizes the upper bound on the bid shading across bid percentile.

The graph 2 plots for all product categories, the upper bound estimates for the cumulative distribution function(cdf) of the ad value. It seems that the ad value follows a log-normal distribution, with the variance varying across categories.

Bid Shading: Recall that bid shading is the amount by which the advertiser shades his bid relative to his ad value. We calculate the bid shading in value percentage term and also in terms of actual amount (measured in cents). Table 7 shows the upper bound on the bid shading amount across different product categories for different percentile. We see some common patterns across the product categories. For all product categories, 90% of the advertisers, the difference between their bid and value is less than 1.5 cents. Car Insurance and Coins seem to have the highest level of bid shading while the laptop has the lowest.

Although the bid shading mount is interesting to analyze, we need to be careful in interpreting the dollar amounts as the bid was re-scaled in this data in order to mask the actual bid amount. Thus, I also calculate the upper bound on the bid shading as a percentage of ad value. Table 6 plots the bid shading in terms of percentage of ad value. Similar to the previous table, the bid in percentage term is also very low. For all categories, the 50^{th} percentile was below 0.2% of ad value. Car insurance and Coins have a higher percentage of bid shading compared to others.

Lastly, we also graph the bid shading amount as a function of the weighted bid. Graph 4 shows that bid shading is increasing in the weighted bid. This also provides evidence that our assumption 2 holds in the data. Additionally, the graph shows there is no strictly monotonic pattern between the bid shading percentage and the quality score. Further analysis is needed to explore the relationship between bid shading and quality score.

The results show that the bid shading amount is low when we look at GSP^w auction in a market with large volume and ads and advertisers. This provides initial evidence towards what we can expect if we change the auction design for an online ad. The small deviation of the bid from the advertiser's value provides initial evidence for minimal change in the bid and thereby the ad revenue when the auction design is changed. Further research is needed to explore this question in depth.

6 Conclusion

This research looks at deriving the advertiser's distribution in a generalized second price auction. The empirical contribution of this paper is that it estimates the advertiser's ad value using weaker assumptions on equilibrium. Using the equilibrium conditions, I estimate bounds on the distribution of the ad value. This result is essential as it helps to simulate how the advertiser's payment behavior changes with change in the market environment. Unlike the consumer side, for which the search engine can make a randomized controlled trial to test the changes, a similar approach is tough on the advertisers' side. This is because the advertisers' reaction to changing market factors such as pricing mechanisms is usually slower. Additionally, frequent changes in the environment can make advertisers leave the ad platform due to an increase in difficulties. Thus, companies such as Microsoft and Google often estimate the advertiser's unobserved parameters such as ad value and then simulate their best response to changes in the environment. Thus, the estimation of bounds on the distribution holds importance in this market and can be used to simulate revenue implications of changes in the market. I find that the estimated distribution is close to the log-normal distribution.

These results are further used to evaluate bid shading, which shows how much advertisers shade their bid below their ad value. The results show that due to intense competition in the online ad market, the bid shading amount is very small. For 90% of the advertisers, the bid shading is less than 1.5% of their ad value. This result has implications for the impact of change in auction design on ad revenue. Specifically, the results indicate that the change in the auction might have minimal impact on ad revenue. However, further research is needed to explore this result in more depth.

References

Agarwal, A. and T. Mukhopadhyay (2016). The impact of competing ads on click performance in sponsored search. Information Systems Research 27(3), 538–557.

Aggarwal, G., J. Feldman, and S. Muthukrishnan (2006). Bidding to the top: Vcg and equilibria of position-based auctions. In International Workshop on Approximation and Online Algorithms, pp. 15–28. Springer.

Aggarwal, G., A. Goel, and R. Motwani (2006). Truthful auctions for pricing search keywords. In Proceedings of the 7th ACM conference on Electronic commerce, pp. 1–7. ACM.

An industry survey conducted by PwC (2017). Iab internet advertising revenue report. https://www.iab.com/wp-content/uploads/2017/12/IAB-Internet-Ad-Revenue-Report-Half-Year-2017-REPORT.pdf,.

Aseff, J. and H. Chade (2008). An optimal auction with identity-dependent externalities. The RAND Journal of Economics 39(3), 731–746.

Athey, S. and G. Ellison (2011). Position auctions with consumer search. The Quarterly Journal of Economics 126(3), 1213–1270.

Athey, S. and D. Nekipelov (2010). A structural model of sponsored search advertising auctions. In Sixth ad auctions workshop, Volume 15.

Bajari, P., H. Hong, and S. P. Ryan (2010). Identification and estimation of a discrete game of complete information. Econometrica 78(5), 1529–1568.

Bajari, P., S. Houghton, and S. Tadelis (2004). Bidding for incomplete contracts.

Baye, M. R., D. Kovenock, and C. G. De Vries (1996). The all-pay auction with complete information. Economic Theory 8(2), 291–305.

Börgers, T., I. Cox, M. Pesendorfer, and V. Petricek (2013). Equilibrium bids in sponsored search auctions: Theory and evidence. American economic Journal: microeconomics 5(4), 163–87.

Constantin, F., M. Rao, C.-C. Huang, and D. C. Parkes (2011). On expressing value externalities in position auctions. In AAAI.

Deng, C. and S. Pekec (2013). Optimal allocation of exclusivity contracts. Work.

Desai, P. S., W. Shin, and R. Staelin (2014). The company that you keep: when to buy a competitor's keyword. Marketing Science 33(4), 485–508.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. American economic review 97(1), 242–259.

Edelman, B. and M. Schwarz (2010). Optimal auction design and equilibrium selection in sponsored search auctions. American Economic Review 100(2), 597–602.

Fox, J. T. (2018). Estimating matching games with transfers. Quantitative Economics 9(1), 1–38.

Fox, J. T. and P. Bajari (2013). Measuring the efficiency of an fcc spectrum auction. American Economic Journal: Microeconomics 5(1), 100–146.

Gatti, N. and M. Rocco (2013). Which mechanism for sponsored search auctions with externalities? In Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems, pp. 635–642. International Foundation for Autonomous Agents and Multiagent Systems.

Ghosh, A. and M. Mahdian (2008). Externalities in online advertising. In Proceedings of the 17th international conference on World Wide Web, pp. 161–168. ACM.

Ghosh, A. and A. Sayedi (2010). Expressive auctions for externalities in online advertising. In Proceedings of the 19th international conference on World wide web, pp. 371–380. ACM.

Giotis, I. and A. R. Karlin (2008). On the equilibria and efficiency of the gsp mechanism in keyword auctions with externalities. In International Workshop on Internet and Network Economics, pp. 629–638. Springer.

Goldman, M. and J. Rao (2014). Position auctions in practice.

Goldman, M. and J. Rao (2016). Experiments as instruments: Heterogeneous position effects in sponsored search auctions. EAI Endorsed Trans. Serious Games 3(11), e2.

Gomes, R., N. Immorlica, and E. Markakis (2009). Externalities in keyword auctions: An empirical and theoretical assessment. In International Workshop on Internet and Network Economics, pp. 172–183. Springer.

Gomes, R. and K. Sweeney (2014). Bayes–nash equilibria of the generalized second-price auction. Games and Economic Behavior 86, 421–437.

Guerre, E., I. Perrigne, and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. Econometrica 68(3), 525–574.

Haile, P. A., H. Hong, and M. Shum (2003). Nonparametric tests for common values at first-price sealed-bid auctions. Technical report, National Bureau of Economic Research.

Haile, P. A. and E. Tamer (2003). Inference with an incomplete model of english auctions. Journal of Political Economy 111(1), 1–51.

Heish, Y., M. Shum, and S. Yang (2014). To score or not to score? structural estimation of sponsored search auctions. Technical report, mimeo.

Hortaçsu, A. and D. McAdams (2018). Empirical work on auctions of multiple objects. Journal of Economic Literature 56(1), 157–84.

Hsieh, Y.-W., M. Shum, and S. Yang (2015). To score or not to score? estimates of a sponsored search auction model.

Hummel, P. and R. P. McAfee (2014). Position auctions with externalities and brand effects. arXiv preprint arXiv:1409.4687.

Iyengar, G. and A. Kumar (2006). Characterizing optimal keyword auctions. In Second Workshop on Sponsored Search Auctions. Citeseer Ann Arbor, Michigan.

Izmalkov, S., D. Khakimova, and G. Romanyuk (2016). Position auctions with endogenous supply.

Jehiel, P. and B. Moldovanu (2005). Allocative and informational externalities in auctions and related mechanisms.

Jerath, K. and A. Sayedi (2015). Exclusive display in sponsored search advertising.

Jeziorski, P. and S. Moorthy (2017). Advertiser prominence effects in search advertising. Management Science 64(3), 1365–1383.

Jeziorski, P. and I. Segal (2015). What makes them click: Empirical analysis of consumer demand for search advertising. American Economic Journal: Microeconomics 7(3), 24–53.

Krasnokutskaya, E. (2004). Identification and estimation in highway procurement auctions under unobserved auction heterogeneity.

Lahaie, S. (2006). An analysis of alternative slot auction designs for sponsored search. In Proceedings of the 7th ACM Conference on Electronic Commerce, pp. 218–227. ACM.

Lu, S., Y. Zhu, and A. Dukes. Position auctions with budget-constrained advertisers.

Lu, Z. and C. Riis (2016). Bayes-nash equilibria in generalized second price auctions with allocative externalities.

Milgrom, P. (2007). Simplified mechanisms with applications to sponsored search and package bidding. Technical report, Working paper, Stanford University.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. Econometrica 70(2), 583–601.

Muthukrishnan, S. (2009). Bidding on configurations in internet ad auctions. In International Computing and Combinatorics Conference, pp. 1–6. Springer.

Narayanan, S. and K. Kalyanam (2015). Position effects in search advertising and their moderators: A regression discontinuity approach. Marketing Science 34(3), 388–407.

Ostrovsky, M. and M. Schwarz (2011). Reserve prices in internet advertising auctions: A field experiment. In Proceedings of the 12th ACM conference on Electronic commerce, pp. 59–60. ACM.

Qin, T., W. Chen, and T.-Y. Liu (2015). Sponsored search auctions: Recent advances and future directions. ACM Transactions on Intelligent Systems and Technology (TIST) 5(4), 60.

Ramaboa, K. K. and P. Fish (2018). Keyword length and matching options as indicators of search intent in sponsored search. Information Processing & Management 54(2), 175–183.

Sayedi, A., K. Jerath, and M. Baghaie (2018). Exclusive placement in online advertising. Marketing Science 37(6), 970–986.

Schurter, K. (2017). Identification and inference in first-price auctions with collusion. working Paper, University of Chicago.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. The RAND Journal of Economics 37(3), 619–640.

Syrgkanis, V., E. Tamer, and J. Ziani (2017). Inference on auctions with weak assumptions on information. arXiv preprint arXiv:1710.03830.

Varian, H. R. (2007). Position auctions. international Journal of industrial Organization 25(6), 1163–1178.

Varma, G. D. (2002). Standard auctions with identity-dependent externalities. RAND Journal of Economics, 689–708.

Yang, S., S. Lu, and X. Lu (2013). Modeling competition and its impact on paid-search advertising. Marketing Science 33(1), 134–153.

Zhu, Y. and K. C. Wilbur (2011). Hybrid advertising auctions. Marketing Science 30(2), 249–273.

6.1 Figures and Tables

Table 1: **List of variables available for consumers side data**

Variable	**Description**
Day	The day of the month.
Advertiser ID	The id for each advertiser.
Clicks	The number of clicks received by an Ad.
Ad Displays	The number of times an ad was displayed.
Keyword	The keyword gives the specified words matched between search and the ad.
Keylength	The number of words specified in the keyword.
Keyword popularity	The sum of the number of times keyword is matched with a query for an ad.
Ad Position	The winning position of the advertisement on the search result page.
CTR	Click Through Rate
Search volume	Number of search per day

Total number of observations: 51,775,997. Observations are at day-advertiser-keyword - ad position level. The data is restricted to ads on first page.

Table 3: Summary Statistics for consumers side data

Variable	Mean/Range	Std. Dev/Max
Consumer's side variables:		
Keywords (ad description & search common words)	6,48,515(count)	-
Keylength (no. of words in keyword)	3.3	19(max)
CTR: Click Through Rate	.016	.1003495
Search Volume	335806	220222.8
Ad-position	3.7	7(max)
Keyword Popularity	5055.59	60326.36

Total number of observations: 51,775,997. Observations are at day-advertiser- keyword - ad position level. The data is restricted to ads on first page. Keyword popularity is calculated as sum of number of times keyword is matched with a query for an ad.

Table 2: **List of variables available on advertisers' side**

Variable	**Description**
Day	The day of the month.
Advertiser ID	The id for each advertiser.
Bid	The per click bid specified by an advertiser for an ad.
Ad	The combination of advertiser and keywords that have the same bid
Ad Displays	The number of times an advertisement was displayed in a specific ad postion.
Keyword	The keyword gives the specified words matched between search and the ad.
Ad Position	The winning position of the advertisement on the search result page.
Ad Specificity	Number of keywords specified for each ad.
Ad popularity	Number of times ad is displayed in a day
Max Keylength	Maximum length of keywords specified in an ad
N	Number of total potential ads in each category
Day popularity	No of approximate searches per day - approximated by no. of 1st ad positions in a day
New Years Day	Dummy for 1st January
Weekends	Dummy for days that fall on the weekend
MLK day	dummy for Martin Luther King Jr. Day

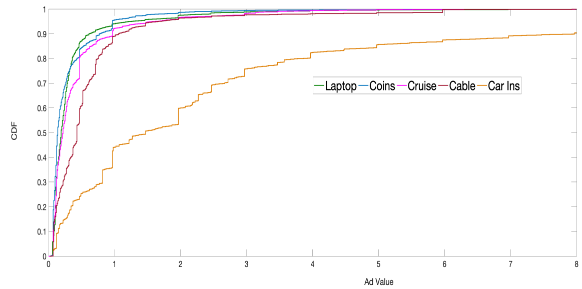
Total number of observations: 5,320,896. Observations are at Aggregated for each day-ad-ad position level. The data is restricted to ads on first page. Ad definition id-keyword combination for the same bid gives the advertisement

Table 4: Summary Statistics for advertisers' side data

Variable	Mean/Range	Std. Dev/Max
Advertiser's side variables:		
Bid[1]	1.28	4.38
Number of ads on the first page	7	0
Keywords (ad description & search common words)	6,48,515(count)	-
Number of ads	383005.4	176428.9
Ad Specificity	4.76	36.55
Ad popularity	176.73	3411.63
Max Keylength	3.47	1.17

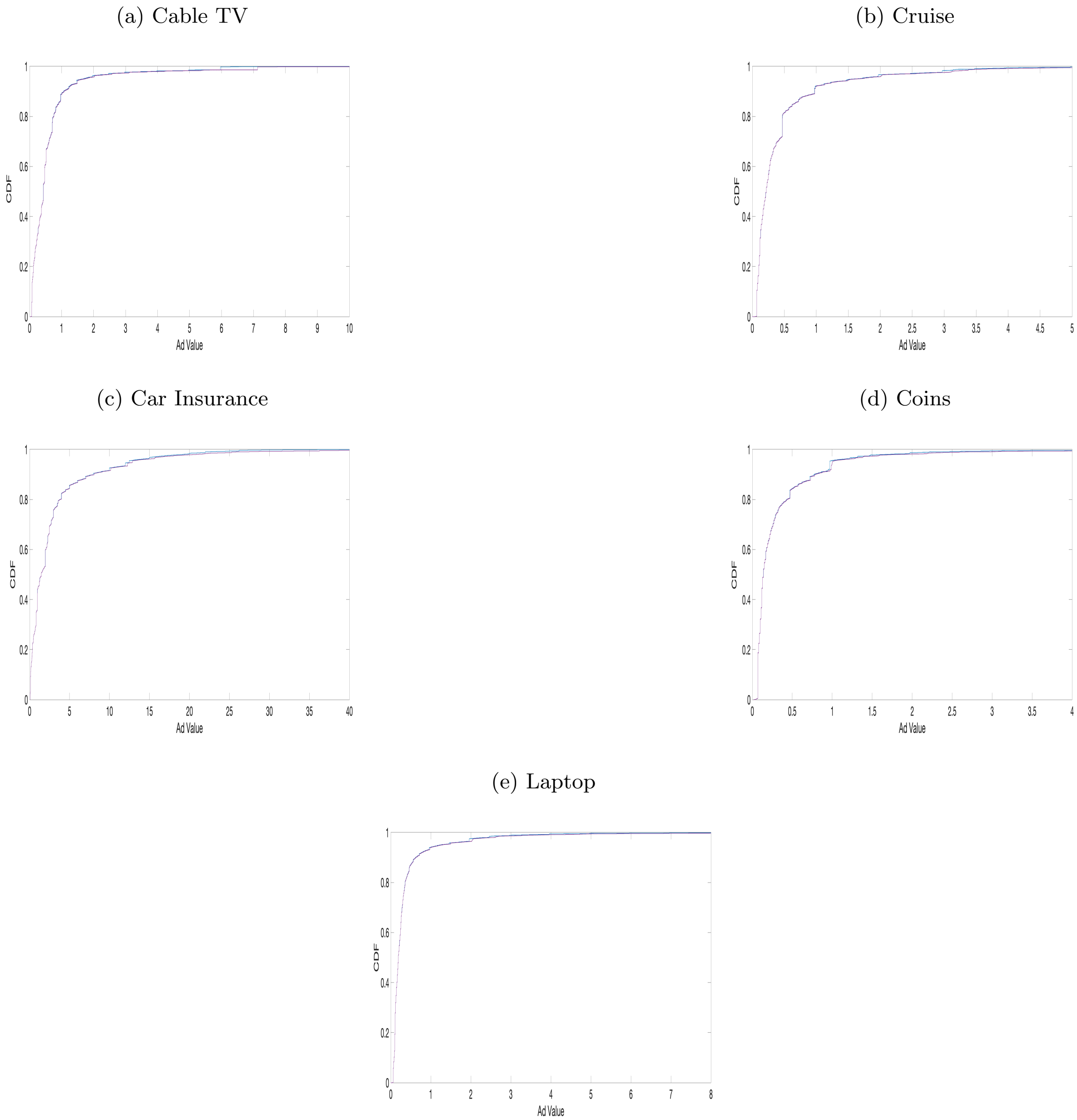
Total number of observations: 5,320,896. Observations are at Aggregated for each day-ad-ad position level. The data is restricted to ads on first page.

Figure 2: Upper bound for the cumulative distribution function of advertisers' ad value



The plot shows the upper bound of the estimated distribution of advertisers' ad value. The distribution of advertisers value for an ad differs across product category.

Figure 3: Cumulative distribution of advertiser's ad value



(a) Cable TV

(b) Cruise

(c) Car Insurance

(d) Coins

(e) Laptop

Notes: This graph plots the upper and lower bound estimated for the ad-value distribution. The x-axis plots the values and y-axis shows the corresponding cumulative distribution at each point.

Table 7: Upper bound on bid shading percentage across product categories

	Cruise	Car Ins	Laptop	Cable	Coins
Bid Shading (in cents):					
Quantiles					
25%	0.0005	0.0028	0.0006	0.0011	0.0056
50%	0.0027	0.0260	0.0035	0.0094	0.0247
75%	0.0141	0.1554	0.0179	0.0421	0.1723
90%	0.1091	1.2675	0.1070	0.2022	1.2770

The table summaries the statistics of the upper bound for bid shading, where bid shading is measured as a percentage of bid

Table 5: Quality measure across product categories

	Car-Insurance	Laptop	Cable	Cruise	Coins
Quality Score $(0-1)$**:**					
Mean	0.05	0.21	0.30	0.27	0.13
	(0.01)	(0.04)	(0.06)	(0.06)	(0.02)
Quantiles					
25%	0.00	0.05	0.11	0.06	0.05
50%	0.01	0.14	0.21	0.21	0.10
75%	0.06	0.29	0.46	0.42	0.16
90%	0.13	0.51	0.66	0.64	0.26

The table summaries the statistics of the predicted quality score. Quality score is the predicted click probability of advertisers. It is estimated on the consumer side in step 1 as the predicted effect of advertiser id on click probability.

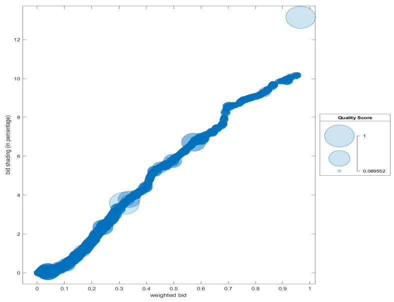
Table 6: Upper bound on bid shading percentage across product categories

	Cruise	Car Ins	Laptop	Cable	Coins
Bid Shading (in terms of value percentage):					
Quantiles					
25%	0.03%	0.05%	0.04%	0.04%	0.06%
50%	0.10%	0.17%	0.13%	0.14%	0.18%
75%	0.27%	0.50%	0.39%	0.40%	0.59%
90%	0.89%	1.50%	1.17%	1.18%	1.66%

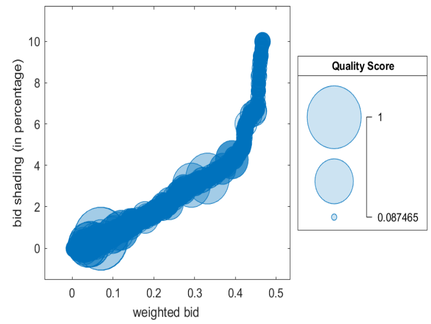
The table summaries the statistics of the upper bound for bid shading, where bid shading is measured as a percentage of bid

Figure 4: Upper bound estimated for the bid shading

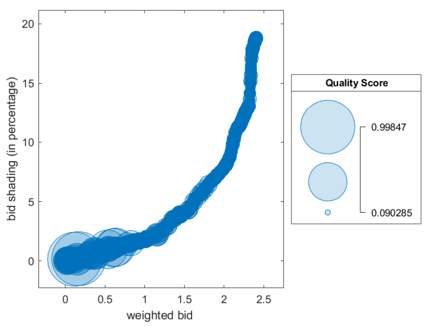
(a) Cable TV



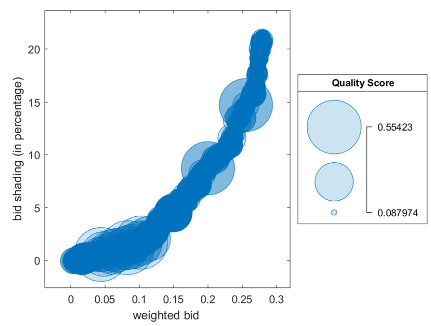
(b) Cruise



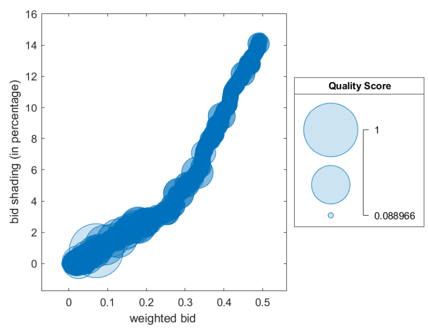
(c) Car Insurance



(d) Coins



(e) Laptop



Notes: This graph plots the upper bound estimated for the bid shading in terms of percentage of ad value on Y -axis and X-axis denotes the weighted bid. The size of the circle indicates the quality score.

7 Appendix

8 Theoretical Proofs

The outline of this section is as follows. First I provide proof of lemma (1) which shows that the equilibrium weighted bid b_w in GSP^w is equivalent to equilibrium bid in GSP auction where the value of the advertiser is replace by their weighted value ω. This equivalence will help us prove the equilibrium bid as well as prove the inequality that we use to bound the value distribution.

In order to make notations easy, I will be using the following notation for differentiation of functions

$f^{'}(x)$ refers to $\frac{d(f(x))}{dx}$ for any f, x

lemma(3) proof:

Proof. Form equation 10 we know the following

$$\frac{v_j - b_j}{b_j} \leq \Phi(G_w, b_{j,w}, \mathcal{C}|K, N) \quad \forall j \epsilon \mathcal{J} \tag{16}$$

The above equation gives the bid shadign amount as a percentage of the advertiser's bid. Next I will show that the same bounds can be used to bound the bid shading as a percentage of the advertiser's value. from lemma 1 we have

$$\begin{aligned} v_j &\geq b_j \\ \rightarrow \frac{1}{v_j} &\leq \frac{1}{b_j} \end{aligned} \tag{17}$$

Substituting equation 17 in equation 16 we get

$$\frac{v_j - b_j}{v_j} \leq \Phi(G_w, b_{j,w}, \mathcal{C}|K, N) \quad \forall j \epsilon \mathcal{J}$$

the above equation derives the upper bound. The lower bound is a direct consequence of lemma 1. □

lemma(1) proof:

Lemma(1) *The weighted GSP auction equilibrium weighted bid function* $b_w^{GSP^w}(v_j, s_j) \rightarrow \mathscr{R}_+$ *is equivalent to the equilibrium bid function in GSP auction with value replaced by the weighted value.*

$$b^{GSP}(\omega_j) = b_w^{GSP^w}(v_j, s_j), \qquad \forall j \ \textit{where}\ \omega_j = s_j \times v_j \tag{18}$$

In this lemma we need to show that the equilibrium weighted bid strategy is equal to the one used in GSP auction – $b_w^{GSP^w}(v, s) = b^{GSP}(\omega)$. Lets take an arbitrary advertiser j and suppose that all advertisers $l \neq j$ have the following weighted bid strategy equivalence

$$b_w^{GSP^w}(v_l, s_l) = b^{GSP}(\omega_l) \qquad \forall l \neq j \tag{19}$$

We will argue 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 :

$$b^{GSP^w} = Arg\max_{\hat{b}} \Pi(\hat{b}|v_j, s_j) = \max_{\hat{b}} \sum_{k=1}^{K} s_j c_k \left[v_j - \frac{\mathbb{E}\left(b_{-j,w}^{GSP^w,[k]} \middle| b_{-j,w}^{GSP^w,[k]} \leq \hat{b} * s_j \leq b_{-j,w}^{GSP^w,[k-1]} \right)}{s_j} \right] \times Prob(b_{-j,w}^{GSP^w,[k]} \leq \hat{b} * s_j \leq b_{-j,w}^{GSP^w,[k-1]}) \tag{20}$$

Now I will use the information that s_j is know to advertiser j and the auctioneer. Thus, the above problem can be rewritten to maximize $\hat{b}_w = \hat{b} \times s_j$ so that the optimal weighted bid is given by following equation:

$$\begin{aligned}
b^{GSP^w} &= \underset{\hat{b}_w}{\text{argmax}} \Pi(\hat{b}_w | v_j, s_j) \\
\rightarrow b^{GSP^w} &= \underset{\hat{b}_w}{\text{argmax}} \sum_{k=1}^{K} s_j c_k \left[v_j - \frac{\mathbb{E}\left(b_{-j,w}^{GSP^w,[k]} \middle| b_{-j,w}^{GSP^w,[k]} \leq \hat{b}_w \leq b_{-j,w}^{GSP^w,[k-1]} \right)}{s_j} \right] \\
&\times Prob(b_{-j,w}^{GSP^w,[k]} \leq \hat{b}_w \leq b_{-j,w}^{GSP^w,[k-1]}) \\
&\text{Using } \omega_j = v_j \times s_j \text{, we get} \\
\rightarrow b^{GSP^w} &= \underset{\hat{b}_w}{\text{argmax}} \sum_{k=1}^{K} s_j c_k \left[\omega_j - \mathbb{E}\left(b_{-j,w}^{GSP^w,[k]} \middle| b_{-j,w}^{GSP^w,[k]} \leq \hat{b}_w \leq b_{-j,w}^{GSP^w,[k-1]} \right) \right] \qquad (21) \\
&\times Prob(b_{-j,w}^{GSP^w,[k]} \leq \hat{b}_w \leq b_{-j,w}^{GSP^w,[k-1]}) \qquad (22)
\end{aligned}$$

Recall that, as shown in equation(19), we assumed that everyone except advertiser j has their weighted bid equal to the equilibrium bid in GSP auction. Thus in the above equation can be rewritten as

$$b^{GSP^w} = \underset{\hat{b}_w}{\text{argmax}} \sum_{k=1}^{K} s_j c_k \left[\omega_j - \mathbb{E}\left(b_{-j}^{GSP,[k]} \middle| b_{-j}^{GSP,[k]} \leq \hat{b}_w \leq b_{-j}^{GSP,[k-1]} \right) \right] \times Prob(b_{-j}^{GSP,[k]} \leq \hat{b}_w \leq b_{-j}^{GSP,[k-1]}) \tag{23}$$

Now, consider a non-weighted GSP auction which has the advertisers' value equal to the weighted value – *advertisers' value is* ω_j. Recall, that in the non-weighted GSP auction the allocation is done according to 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 :

$$\begin{aligned}
b^{GSP} =& \underset{\check{b}}{\text{argmax}} \Pi(b_w | \omega_j, s_j) \\
\rightarrow b^{GSP} =& \underset{\check{b}}{\text{argmax}} \sum_{k=1}^{K} s_j c_k \left[\omega_j - \mathbb{E}\left(b_{-j}^{GSP,[k]} \middle| b_{-j}^{GSP,[k]} \leq \check{b} \leq b_{-j}^{GSP,[k-1]} \right) \right] \times Prob(b_{-j}^{GSP,[k]} \leq \check{b} \leq b_{-j}^{GSP,[k-1]})
\end{aligned} \tag{24}$$

The optimization problem in equation(23) and (24) 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 will also have

$b^{GSP}(\omega_j) = b_w^{GSP_w}(v_j, s_j)$. 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. te here instead of using uniquness , I can also show the equation for all i , then the system of equations coincides

8.1 Equilibrium Bid

Theorem 2. *If assumptions (1) and (2) holds, then the unique symmetric Bayesian Nash equilibrium of the weighted GSP auction is given by the following bidding strategy:*

$$b_w(\omega) = \omega - \Gamma(\omega) - \sum_{n=1}^{\infty} \int_0^{\omega} M_n(\omega, t)\Gamma(t)dt \qquad \forall \omega \sim F_w(.) \tag{25}$$

where

$$\Gamma(\omega) = \frac{\sum_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-2} \int_0^{\omega} F^{N-k}(x)dx}{\sum_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-1} F^{N-k-1}(\omega)}$$
$$M_1(\omega, t) = \frac{\sum_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-2} F^{N-k-1}(t)}{\sum_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-1} F^{N-k-1}(\omega)}$$
$$M_n(\omega, t) = \int_0^{\omega} M_1(\omega, \epsilon) M_{n-1}(\omega, \epsilon) d\epsilon \qquad \forall n \geq 2$$

The above proposition shows the equilibrium weighted bid[36].

Proof From lemma(1) we know that the equilibrium weighted bid function for the GSP^w auction with values equal to $v = \{v_1, ..v_j..v_N\}$ and weight equal to $s = \{s_1, ..s_j, ..s_N\}$ is equal to the equilibrium bid of a GSP auction with the value equal to $v * s \equiv \omega = \{\omega_1, ..\omega_j, ..\omega_N\}$. Thus, solving for the equilibrium weighted bid in GSP^w auction is equivalent to solving for the equilibrium bid in the GSP auction with type ω instead of v. The proof of the equilibrium weighted bid then follows same steps as shown in paper- Gomes and Sweeney (2014) Gomes and Sweeney (2014).

8.2 Proof for lower bounds on advertiser's value i.e. theorem(1)

Proof. For this proof only, we will take the following notation for the probability of winning position k

$$Prob(\omega_{-j}^{[k]} \leq \omega \leq \omega_{-j}^{[k-1]}) \equiv \zeta_k(\omega)$$

[36] Once the lemma(1) is used to make the objective function depend on weighted value only , the subsequent proof of the equilibrium is similar to BNE derived in Gomes and Sweeney (2014)

By theorem(2), we know an efficient equilibrium exist. Using Revelation Principle, advertiser $j's$ payoff given the reported type as $\hat{\omega}$ will satisfy the following

$$\omega_j = \underset{\hat{\omega}}{\text{argmax}} \sum_{k=1}^{K} \zeta_k(\hat{\omega}) c_k \left[\omega_j - \mathbb{E}\left(\frac{b_{-j,w}(\omega_{-j}^{[k]})}{s_j} \middle| \omega_{-j}^{[k]} \le \hat{\omega} \le \omega_{-j}^{[k-1]} \right) \right] \tag{26}$$

Applying the envelop theorem (see Milgrom and Segal (2002)) in the payoff function in equation(26), we have :

$$\frac{d}{d\omega}\Pi(\omega)\bigg|_{\omega=\omega_j} = \sum_{k=1}^{K} c_k \zeta_k(\omega)$$

and also using the Fundamental Theorem of Calculus, we get

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

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

$$\rightarrow \Pi(\omega_j) = \sum_{k=1}^{K} c_k \int_0^{\omega_j} \zeta_k(\omega) dx \tag{27}$$

Furthermore, using equation(26)) and equation (27), we get

$$\sum_{k=1}^{K} c_k \int_0^{\omega_j} \zeta_k(\omega) dx = \sum_{k=1}^{K} c_k \zeta_k(\omega) \left[\omega_j - \mathbb{E}\left(\frac{b_{-j,w}(\omega_{-j}^{[k]})}{s_j} \middle| \omega_{-j}^{[k]} \le \omega \le \omega_{-j}^{[k-1]} \right) \right]$$

$$\rightarrow \sum_{k=1}^{K} c_k \left[\zeta_k(\omega)\omega_j - \int_0^{\omega_j} \zeta_k(\omega) dx \right] = \sum_{k=1}^{K} c_k \zeta_k(\omega) \mathbb{E}\left(\frac{b_{-j,w}(\omega_{-j}^{[k]})}{s_j} \middle| \omega_{-j}^{[k]} \le \omega \le \omega_{-j}^{[k-1]} \right)$$

Using intergration by parts on the left hand side we get

$$\sum_{k=1}^{K} c_k \int_0^{\omega_j} x \frac{d(\zeta_k(x))}{dx} dx = \sum_{k=1}^{K} c_k \zeta_k(\omega) \mathbb{E}\left(\frac{b_{-j,w}(\omega_{-j}^{[k]})}{s_j} \middle| \omega_{-j}^{[k]} \le \omega \le \omega_{-j}^{[k-1]} \right)$$

opening up right hand side, we get

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

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

Differentiating both sides we get:

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

As the last term is negative, we can get the following inequality

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

Opening up $\frac{d(\zeta_k(\omega))}{d\omega}$ and rearranging the terms we get

$$\Rightarrow \omega_j \leq b_{j,w}\left(1 + \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} F_w(x)^{N-k}(k-1)(1-F_w(\omega))^{k-2}}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} \left[(N-k)F_w(x)^{N-k-1}(1-F_w(\omega))^{k-1} - F_w(x)^{N-k}(k-1)(1-F_w(\omega))^{k-2}\right.}\right)$$

Canceling quality score from both sides for weighted value and weighted bid, we get

$$\Rightarrow v_j \leq b_j\left(1 + \frac{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} F_w(x)^{N-k}(k-1)(1-F_w(\omega))^{k-2}}{\sum_{k=1}^{K} c_k \binom{N-1}{k-1} \left[(N-k)F_w(x)^{N-k-1}(1-F_w(\omega))^{k-1} - F_w(x)^{N-k}(k-1)(1-F_w(\omega))^{k-2}\right.}\right)$$

Hence proved

□

8.3 Data cleaning

The data set is provided by Yahoo! as part of the Yahoo! Research Alliance Webscope program. It gives details about five different categories, namely laptop, cable, coins, cruise, and car insurance, over 123 days from January 2008 to April 2008. The data has information about keywords, bid, clicks, ad position, and display frequency. The keywords in the data set include one of the base category word 'coin,' 'laptop,' 'cable,' and 'car insurance.' Apart from the base category word, the keywords also include one or more additional words. For example, 'business laptop' and 'student laptop' are two keywords within the base category laptop. The additional words provide a more targeted ad. For instance, an ad 'business laptop' targets consumers that are specifically looking for business usage; however, keyword 'laptop' captures a broader search for any laptop need of a consumer. The maximum number of words in a keyword is 19.

Another key characteristic of this data is that the keywords, and the advertiser's id is masked. This means the actual identity of the advertisers and keywords is masked; however, I can track the same advertiser and keywords across time. More details on identifying the base categories are given in the appendix(8.4). To do the analysis, I restrict the dataset to ads on the first page and consider only the first seven ads. The total observations in the raw data set are 77,850,272. After restricting the data to the first page, the data has 51,775,997. Additionally, the data is divided into two parts according to the aggregation needed for consumers or advertisers analysis. On the consumer's side, the analysis focuses on the click decision on the consumer. The data provider, Yahoo! Research Lab, aggregated the data for each day-keyword-advertiser-position combination. Let us take an example of observation, for January 1st 2008, the data reports that the keyword 'business laptop' specified by Amazon that was displayed in the first position, got 100 displays and five clicks. For the same keyword and advertiser, i.e. Amazon ad with keyword 'business laptop,' I will have a different observation for the ad displayed in the second position. This means that the data does not aggregate over different positions and reports results for each position and keyword separately. This is an advantage for the consumer side, as the different keywords can help in capturing the difference in consumer search. Thus, the total observation on the consumer side analysis is 51,775,997.

On the other hand, for analysis on the advertiser's side, I need information on how the advertiser maximized profit for each ad. While deciding the bid, the advertiser accommodates the expected clicks and price for each ad across different winning positions over a day. This means here the relevant variation is the bid over ads and not the keywords. Therefore, on the advertisers side, I aggregate the data on day-ad-advertiser-position combination.[37] An example of an observation is; on January 1st, 2008, the data reported that an ad by Amazon had an average position of 2.5, got 200 displays, and 10 clicks. As the model assumes that the advertiser maximizes profit over an ad, the results from aggregation are more aligned with the advertiser's profit-maximizing strategy.[38] The aggregated data gives 5,306,217 observations.

The data is aggregated at day-ad-position level. Thus, I assume that the bid is aggregated at day level. However, there are some advertisers that seems to change their bid over day. I aggregate the bid for such advertisers at day level. This represents 5.5% of the full dataset. Additionally, for few observations the number of clicks are more than the number times ad was displayed. This is likely either a recording error or a case where the same ad was clicked twice by the consumers. The data does not provide enough information to explore this behavior further. This also has a direct impact on the click rate calculation, which recall is the ratio of click to number of displays. To handle this error, I impute the click rate as 1 if it is greater than 1. This represents about 0.01%

[37]Recall the definition of an ad is the set of keywords for which the advertiser has specified the same bid.

[38]Note that the advertiser can always specify a different bid for each keyword, and in this case, each keyword would be treated as a separate ad in this data set. So the aggregation is only on the keywords for which the advertiser has specified the same bid, meaning the bid was maximized over all the keyword that has the same bid per day.

of the full dataset.

8.4 Within product category variation :

This data has varied product categories, namely 'laptop', 'cable', 'cruise', 'coin' and 'car insurance'. Additionally, recall that the keywords are declassified and thus the product categories are also declassified.[39] To overcome this we analyze the differences in the deidentified category and match it to the closest possible category among ['laptop', 'cable', 'cruise', 'coin' and 'car insurance'] according to the observed features. Table(8) gives a summary of how variables differ among categories. Additionally the table(8) shows the corresponding mean value for all features for different categories.

Let us first look at features of category 1. This category is characterized by very high average bid and relatively small number of competitors compared to other categories; this is consistent with the car insurance product category. They are known to be the industry with one of the highest pay per click.[40] This is due to the high profit margins in auto insurance industry(which is result of it being a highly concentrated market). Other observations about this market which makes it consistent with the car insurance is that there are no keywords with one word, again this is consistence with car insurance since you have to atleast type two words 'car' + 'insurance'. It also has two words that occur in all keywords which makes it most likely be "car" and "insurance"

The next category that stands out is category 2, which is characterized by high number of advertisers and high number of search queries per day. Due to its high volume of consumer searches this is likely a consumer good, which makes it closest to 'Laptop' category in the data.[41]
Apart from this the other category that is easy to identify is category 0. This category has high number of advertisers and a high click rate. This is again a popular category, and thus it is best matched with the category ''Cruise'. Due to its low value for search volume , bid and clicks, it is likely to be the less popular category in the data i.e. 'Coins'. The next category is 'Coins', relative to other categories, Coins category has a lower search volume, less competition and lower bid. Thus, 'Coins' fits Category 4, that has the lowest value in all four variables reported in table 8. Lastly, 'Cable' is also a less popular and lower priced category but it is relatively more popular than 'Coins'. Therefore, we will match Cable with category 3. The table below summarizes the findings. Note although these claims are just approximation, we will use them for the rest of the analyzes. Even if there is some error in identifying the category , we can still use the features of the

[39] The categories are identified through specificity of the keyword. The data has four single word keyword which identify the four categories, and the fifth category is identified by the two word keyword that has the highest frequency

[40] refer to these articles for more information : - https://www.adgooroo.com/the-most-expensive-keywords-in-paid-search-by-cost-per-click-and-ad-spend/ and http://www.automotivedigitalmarketing.com/photo/1970539:Photo:28810

[41] as that is the only consumer good category in the data

category and interpret how and why the results might differ for categories with different features.

Table 8: **Feature of Different Categories**

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

Showing mean value for each category