Revenue-Maximizing Number of Ads per Page in the Presence of Market Externalities*

Pallavi Pal†

For Latest Version Click Here

Abstract

Firms use advertising as a medium to gain a competitive advantage, which is negatively affected if the ad appears alongside their rival's ad—a form of externality. The multiple ad display setting on search engines, such as Google and Yahoo!, introduces such externalities in the market. In this paper, I estimate a structural model based on a novel data set of Yahoo! ads to (i) quantify the effect of externality on an advertiser's willingness to pay and (ii) simulate the revenue-maximizing number of ads for a search engine. First, I find that externality depends on the quality and quantity of competing ads. For example, an advertiser's willingness to pay decreases by 18.5 percent due to the addition of a second high-quality ad, but only by 0.15 percent due to the addition of a seventh low-quality ad. Second, the counterfactual results suggest that the revenue-maximizing number of ads per page differs across the ad product category, with the average being five ads per page, and implementing the suggested number of ads would lead to a 4.5 percent increase in revenue, on average. These results provide evidence in support of recent changes in the online advertising market; for example, Microsoft introduced a service called RAIS that provides advertisers with an option of an exclusive ad display.

JEL: D44, C14, C57

Keywords: online advertising, partial identification methods, nonparametric estimation, auctions.

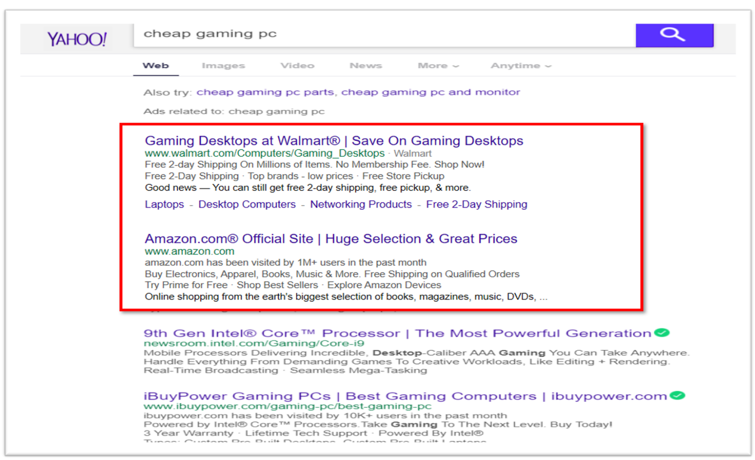
*I am grateful to Jeitschko D. Thomas, Kyoo il Kim, Hanzhe Zhang, Roger J. Calantone, Michael Conlin, Jay Pil Choi, (Julian) Chenhui Guo, Ashish Agarwal, and Vivek Bhattacharya for all the useful discussions and suggestions. I would also like to thank the Yahoo! research lab for providing the dataset used in this paper. I am grateful for comments by participants at Michigan State University, the Midwest Econometric Group, the American Marketing Association conference and Grand Valley State University. Note that all views This is a work in progress, do not circulate without permission

†Author. palpalla@msu.edu; Department of Economics, Michigan State University, 466 W. Circle Dr., East Lansing, MI 48824.

1 Introduction

Online advertising is a new but rapidly growing market. Back in 1998, when Google was a small startup company, there were on average 10,000 search queries made per day; fast-forward to 2019, Google recorded an average of 5.5 billion searches per day.[1] Furthermore, this rapid growth is accompanied by new and evolving ad features. One such example is the pay per click feature, wherein the advertiser pays for the click on the ad link and not for the ad display.[2] In such a dynamic environment, even a feature as simple as the number of ads per page can have significant implications. This paper focuses on the implications of the number of search ads per page on the search engine's ad revenue. Search ads are paid search links that appear above generic[3] search results on major search engines such as Google, Yahoo! and Bing. Figure 1 gives an example of the Yahoo! result page.

Figure 1



The motivation for looking at the number of ads per page is twofold. First, analyzing this feature can help us derive the revenue-maximizing number of ads per page, which can provide suggestions on how to improve the current practices. Second, the empirical literature on online advertising does not address how the number of ads impacts the ad price and, thereby, the search engine's revenue. This paper fills this gap by quantifying the revenue impact of the number of ads

[1] https://martechseries.com/mts-insights/guest-authors/googles-seo-strategy-is-constantly-changing-four-ways-small-businesses-can-keep-up/

[2] This is true for most of the online advertising platforms such as Google, Yahoo! Facebook, Amazon, and eBay.

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

per page.

The search engine faces a trade-off while deciding the number of ads per page. The addition of one more ad leads to an increase in the quantity of ads sold; however, it also leads to a decrease in the ad price. Thus, the revenue-maximizing number of ads is an empirical question that is derived after quantifying the change in price due to a change in the number of ads. Note that the price for each ad is decided through an auction mechanism. The change in equilibrium price is broadly due to the auction design, as well as due to the externality exerted from other ads on the page. The externality is defined as a side effect of the presence of other competing ads on an advertiser's benefit from the ad. I will first talk about the externality in the market and then give details of the auction design.

Let us look at an example to understand the consequences of externality in this market. Consider Walmart as an advertiser that has an ad appearing in two consumer search queries. In the first case, the ad appears adjacent to one from Amazon (as seen in figure 1), and in the other, it appears as an exclusive ad. In this example, the consumer who sees only Walmart's ad is more likely to buy from them, compared to the consumer who sees both Amazon and Walmart's ad. The presence of Amazon's ad exerts an externality on Walmart's benefit from the ad, which translates into Walmart having a higher willingness to pay for the exclusive display. Notice that the externality from the other ad is dependent on the number as well as the quality of the ads, where the quality measures the consumers' preference for an advertiser. For instance, an adjacent ad from a high-quality advertiser, say Amazon, exerts a higher externality than an ad from a low-quality advertiser. Thus, I create an externality index that captures the quality as well as the number of other ads on the page.

Although the externality estimates are interesting by themselves, this paper aims to compute the revenue-maximizing number of ads for which we also need to accommodate other components that affect the equilibrium price. The selection and pricing of the ads is done through an auction referred to as Generalized Second Price (GSP) auction.[4] Additionally, the advertisers pay per click and not for the ad's display; thus, the advertiser submits a per click bid. The winning ads are decided using the "weighted bid", where the weight captures an ad's click probability. Note that a single auction is used to allot all the ad positions on a given page. The positions are assigned in

[4] A new auction is held for consumer search.

order of the weighted bids, meaning an ad with a higher weighted bid is allotted a higher position. The equilibrium condition shows that the ad price depends on the advertiser's willingness to pay, the consumer's click probability, and the information available to the advertiser.

To evaluate this empirically, I use a data set provided by the Yahoo! research lab. 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. 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.[6] Notice that the data does not provide the quality of an ad, I solve this limitation by estimating the quality of an ad in terms of the advertiser's effect on consumer's click probability.

For estimation, I use a discrete choice method to model consumers' click decisions and a partial identification method to bound the advertisers' willingness to pay. The estimation is performed in three steps. In the first step, I estimate the parameters that affect a consumer's click decision using a weighted logit model.[7] In the second step, I set up a hedonic regression to estimate the effect of the externality on an advertiser's bid. The final step estimates bounds on the distribution of the advertiser's unobserved willingness to pay for which I follow the Haile and Tamer (2003) methodology for partial identification of distribution.

The findings show that externality has a negative and non-linear effect on the advertiser's willingness to pay. For a one-percent increase in externality leads to 0.34, 0.6, 1.3 and 0.14 percent decrease in the advertiser's willingness to pay for categories car insurance, laptop, cable tv and coins respectively.[8] The effect of including one more ad is dependent on the quality of the additional ad as well the number of ads already on the page. In the laptop category following the Walmart example earlier, this would imply that addition of a second ad decreases willingness to pay by twenty five percent if the new ad is from Amazon, but only by two percent if it is from a local

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

[6] the number of ads on the first page is assumed to be seven ads unless observed less than 7 ad positions. This is a common assumption made for papers using this data set from yahoo such s Agarwal and Mukhopadhyan (2016) [1]

[7] The reason for using a binary logit model instead of a multinomial logit is for two reasons. First, the choice set of the displayed ads varies depending on the consumers' preferences. Thus, the choice set is endogenous. Second, the choice set does not have a fixed number of choices.

[8] Note that cruise did not show a negative or significant effect of externality

retailer.[9]

These estimated primitives and equilibrium price condition are used in the Monte Carlo simulation to derive and measure gains from implementing the revenue-maximizing number of ads. To derive the revenue-maximizing number I calculate the expected revenue for different values of the number of ads per page and then select the one that has the highest expected revenue. The results show that the revenue-maximizing number is 5 ads per page, on average. For three out of five categories the revenue-maximizing number of ads are different than the currently display quantity of ads per page (i.e. seven ads). Implementing the suggested number of ads per page would lead to, on average, a 4.5 percent gain in revenue, which translates into 5.2 billion dollars in revenue.[10] One of the reason for a difference in the result in this paper and current practices is that the paper suggests optimizing number of ads separately for each category, however the search engines currently optimize the number of ads jointly for all categories.

These results provide evidence on how search engines can increase revenue by changing the design of the ad space. Furthermore, these suggestions extend the recent changes in the online advertising market; for example, Microsoft has introduced a service (RAIS) that provides advertisers with an option of an exclusive ad display.

This paper contributes to several different strands of literature.[11] It contributes to the studies that look at the effect of externality in the online advertising market. To the best of my knowledge, this is the first paper that estimates the effect of externality on advertiser's willingness to pay. Additionally, this paper also contributes to the studies on equilibrium price in auction design, where recent papers have looked at solving the equilibrium price under a more realistic assumption on the information available to the advertisers (for example, Athey and Nekipelov (2010) look at entry uncertainty). In this paper, I extend this further by looking at the incomplete information case and provide for equilibrium conditions that can be estimated empirically.[12] Lastly, this paper

[9] assuming local retailer is of low quality and amazon is of high quality

[10] calculated using google's advertising revenue in the second quarter of 2019 - see here for details https://www.statista.com/statistics/266249/advertising-revenue-of-google/

[11] Please refer to the next section for a full literature review.

[12] The closest paper to this analysis is Gomes (2014), which solves for the incomplete information case in the GSP auction. Here, I do a nontrivial extension of their work by solving the incomplete information case in the weighted auction. Furthermore, I contribute to this literature by providing equilibrium bounds that can be estimated empirically.

contributes to the relatively new literature on ad display design; the closest related paper is Jerath and Sayedi (2015). They look at introducing exclusive ad display options, whereas this paper looks at the more general case of the revenue-maximizing number of ads, in which an exclusive ad is a special case.

The remainder of the paper is structured as follows: section 2 gives an overview of the literature, section 3 gives an overview of the market and data, section 4 presents the theoretical model, section 5 discusses the identification strategy, section 6 specifies the econometric method, section 7 gives the results, section 8 presents the counterfactual analysis. Finally, section 9 summarizes the findings and discusses the broader consequences of this paper.

2 Literature review

This research is related to a few different strands of literature. It contributes to the literature on externality in the online advertising market. The empirical studies on externality have focused on the effect of externality on the consumer's decision to click on an ad, such as Jeziorski and Segal (2015) and Narayanan and Kalyanam (2015).[13] Despite the growth of literature on online ad externality, little effort has been made to empirically estimate the indirect effect of externality on an advertiser's behavior. This paper focuses on these previously unexplored issues: the effect of online ad externality on an advertiser's willingness to pay for an ad and, consequently, on ad platform revenue.

This paper is also related to the work on estimating the unobserved advertisers' willingness to pay 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. Borgers et al. (2013) analyze Yahoo! data to estimate 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

[13] For instance, Jeziorski and Segal (2015) [36] show that consumers click on multiple ads and that the click probability is affected by the presence of other ads. Narayanan and Kalyanam (2015)[45] show that, for large firms, higher ad positions lead to smaller click probability improvements.

as Google or Microsoft).[14] A key contribution of this paper is that it looks at the equilibrium behavior under weaker information assumptions. The empirical literature in sponsored search auctions has looked at variants of full information, with few looking at uncertainty in the market. This paper relaxes the full information assumption and examines the optimal bidding behavior under incomplete information. Gomes and Sweeney (2014) solve for the incomplete information case in a non-weighted Generalized Second Price (GSP) auction. This paper extends their work by looking at the incomplete information case in weighted GSP auction. The extension is nontrivial as the weight introduces a multidimensional type of the bidder. Additionally, the equilibrium bid does not have a closed-form. 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 willingness to pay. To the best of my knowledge, this is the first paper that proposes how to estimate the willingness to pay under incomplete information.

This paper is also related to the literature exploring multi-ad display settings in sponsored search ads. For instance, few papers look at giving the advertisers the option of bidding for both multi-ad and exclusive-ad option, such as Jerath and Sayedi, (2015), Deng and Pekec (2013), and Ghosh and Sayedi (2010). The change in the auction design makes the advertisers strategically change their bid, which has led to advocacy for changing the auction design to VickreyClarkeGroves (VCG) auction, for example, in Sayedi, Kinshuk, Baghaie (2018). In this paper, the counterfactual suggestion is on showing a fixed number of ads that differ across the ad product category. The advantage of the suggested fixed number of ads is that it does not require a change of setup for the advertisers. Thus, it is easier to execute. Another contribution of this paper is that, unlike the previous papers that look at this question from a theoretical point of view, this paper estimates the market parameters in the empirical section. Thus, the simulations provide a more realistic magnitude of gain through the proposed new method.

Lastly, 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

[14] Specifically, they accommodate uncertainty in bidders' perceptions (due to randomness in a bidder's quality score over time, as well as in the set of competitor bidding in the auction at any time).

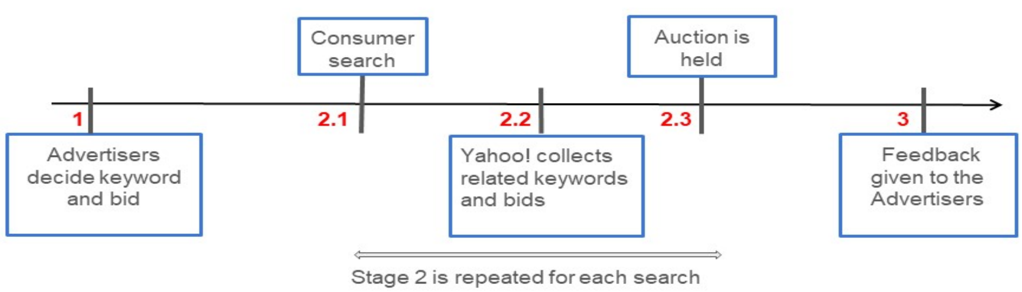
extend it and show how to apply the method in an online auction, i.e., a Generalized Second-Price auction.

3 Market Environment and Data Analysis

3.1 Overview of the search advertising market

In this section, I discuss the search ad market from the advertiser's point of view. Sponsored search ads are paid search links that appear alongside search results (as shown in fig(1)). These ad links are purchased by advertisers in order to have their website appear higher in the search results page. Multiple ad position slots are allotted for each search result page through an auction. The ad display process has three stages, as shown in the timeline (in fig(2)).

Figure 2: Timeline



The figure shows the stages within each period. The advertiser decides the bid and keywords at the start of the period, and the bid enters all auctions held in the period. A separate auction is held for each search query. Keywords are words specified by the advertiser that describe the ad and are used by Yahoo! to match it to the search queries.

Stage 1. Advertisers select their bid and keywords.

In the first stage, the advertiser has to decide the bid per click as well as the ad-related words (referred to as *keywords*). Following the earlier example, this means Walmart needs to decide the bid as well as a set of keywords for an ad on a gaming laptop. The advertisers can specify multiple keywords for an ad. For instance, in this example, Walmart specifies keywords 'gaming laptop' and 'gaming laptop cheap'. The multiple keywords help in reaching consumers with diverse search queries.

Stage 2. Auction is held, and winners are decided.

In the second stage, for every search query, Yahoo! collects all related ads, which is done by matching the search query words to the ad's keywords. The matched ads enter an auction to decide the winning ads that will be displayed to the consumers. Yahoo!, similar to other search engines such as Google, uses a special auction for deciding on the winning ad called the Generalized Second Price (GSP henceforth) auction. In this auction, apart from deciding the winning ad, the auction also decides the ad position of each ad. The winning ads are allotted ad positions in order of the ranked per click bid, implying the highest bidder gets the highest position, second highest gets the second-highest position. The ad payment is equal to the bid of the next highest rank. This auction was further modified in 2006 (first by Google then followed by Yahoo! and Bing) to include weights, called 'quality scores' assigned to each ad. These scores were initially based on the estimated click-through rate the bidder would attain if it were in the first position.[15] The logic behind this design is straightforward: allocating an advertisement to a given slot yield expected revenue equal to the product of the price charged per click, and the click-through rate.

Stage 3. Yahoo! sends feedback on ad's performance and ad price to the advertisers.

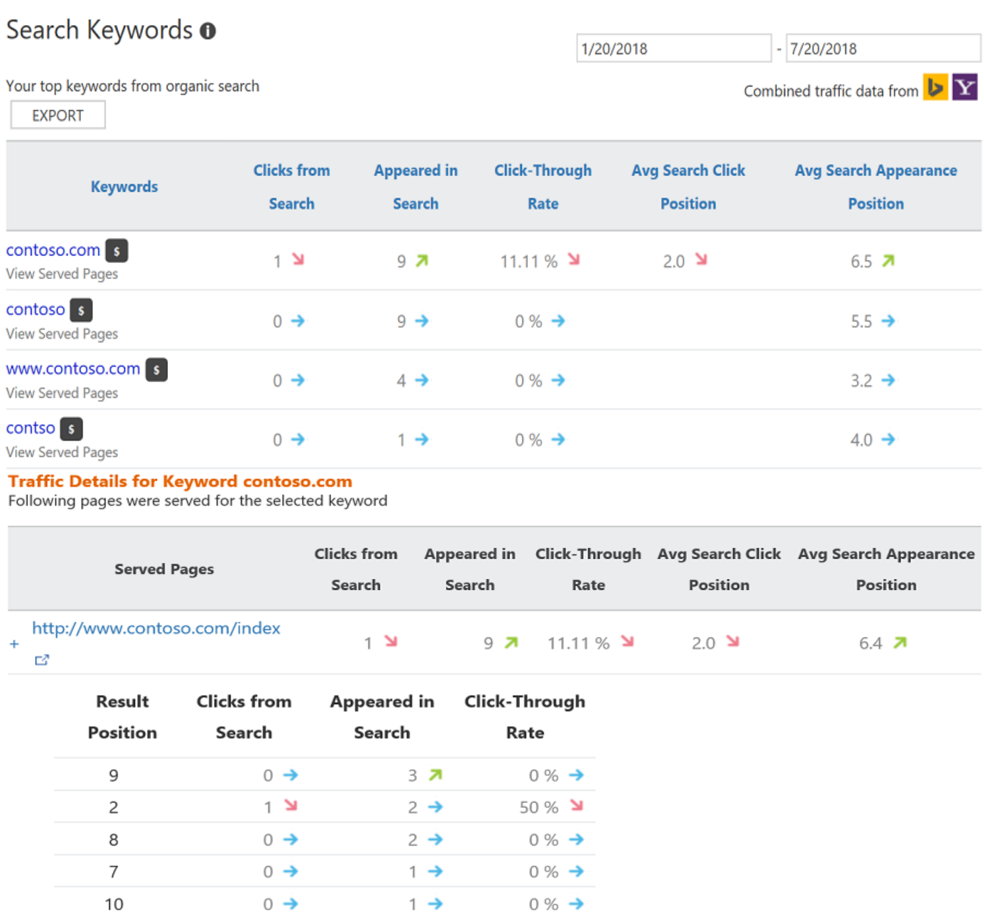
The third and final stage compromises of a feedback report that goes to the advertiser about the ad performance. figure 3 shows a snapshot of an advertiser's account on Yahoo!. The variables in the data set are similar to the feedback information available to the advertiser. The advertiser gets detailed feedback on the ad, which is aggregated for each specified keyword. Following the earlier example, this means for Walmart's ad on 'cheap gaming laptop', Walmart gets periodical information[16] on the two keywords, namely 'gaming laptop' and 'cheap gaming laptop'. For each keyword, they get the display and click frequency in each ad position. Apart from the consumer response, they also get information on the price paid per click for each winning position. For

[15] Later, Google introduced additional variables into the determination of the score, including measures of the match between the advertisement and the query. However, the time period of our data is before this when the quality score used was primarily based on click-through rate.

[16] I assume that the advertiser uses the day as a given period to change/revise their bid. Note that the information provided to the advertiser can be more detailed than the daily aggregation that I assume here. However, previous papers have noted that the change in a bid does not change much within a day; see Borger (2013)[11].

example, feedback for Walmart's ad shows that the search queries that matched keyword 'cheap gaming laptop' had their ad display 100 times in the 1^{st} position, and that translated into five clicks, with 0.2 cents as the price per click. The data set used in this paper contains similar information, where, on the consumer side, it reports the display and click frequency for each keyword-ad position observation aggregated on a daily basis, and on the advertisers' side, it reports the advertiser's bid and winning ad positions. One variable that is known to advertisers and the auctioneer is the advertiser's own quality score. This is known to the players but is not observable to the econometrician. I will solve this problem by estimating the 'click-through rate' on the consumer's side using consumer response data.

Figure 3



Search Keywords

1/20/2018 - 7/20/2018

Your top keywords from organic search

EXPORT

Combined traffic data from

Keywords	Clicks from Search	Appeared in Search	Click-Through Rate	Avg Search Click Position	Avg Search Appearance Position
contoso.com $ View Served Pages	1	9	11.11 %	2.0	6.5
contoso $ View Served Pages	0	9	0 %		5.5
www.contoso.com $ View Served Pages	0	4	0 %		3.2
contso $ View Served Pages	0	1	0 %		4.0

Traffic Details for Keyword contoso.com

Following pages were served for the selected keyword

Served Pages	Clicks from Search	Appeared in Search	Click-Through Rate	Avg Search Click Position	Avg Search Appearance Position
+ http://www.contoso.com/index	1	9	11.11 %	2.0	6.4

Result Position	Clicks from Search	Appeared in Search	Click-Through Rate
9	0	3	0 %
2	1	2	50 %
8	0	2	0 %
7	0	1	0 %
10	0	1	0 %

The figure shows a sample snapshot of information available to advertisers. The first one shows aggregated data for each keyword. The second one shows the details for each keyword. Click through rate refers to the ratio of the number of clicks times the number of displays. Refer to this link for more information.

Apart from this, a key point to notice is that the advertisers do not get information on the exact identity of other ads on the page. Thus, to capture externality, the following assumption is made about advertisers' belief about the presence of other ads on the page.

Assumption 1. *Advertisers can calculate the expected quality of ads in each position.*

Advertisers measure externality in the market by using the expected quality in each position that is they know by assumption(1). This assumption is realistic as the advertisers have multiple period experience on the search engine and can also do a search related to their ads to see which other ads are appearing on the search result page.

3.2 Data

The data set is provided by Yahoo! as part of the *Yahoo! Research Alliance Webscope program.*[17] It is a four-month period of data covering search queries from January 2008 to April 2008. The sample covers all search ads[18] in 5 categories, namely *Laptop*, *TV Cable, Cruise, Collectible Coins* and *Car Insurance*. Each category is treated as a separate data set, and the results are obtained separately for each of them. The advantage of data from multiple industries is that after the estimation, we can compare the results across industries to see whether the results are sensitive to industry characteristics. The data set has two parts. The first part has consumer side information, and the second part has advertiser bid and auction outcome information. For this the analysis, I limit my sample to ads on the first page of the search result.[19] This is mainly done because 90% of the clicks in the data are from first page ads, as can be seen by figure A. This finding is also consistent with the observed pattern in the industry, which shows that the consumers do not go beyond the first page.[20]

Consumers side data: On the consumer side, I have information about ad display and consumers' click response for each position-advertiser-keyword combination. Continuing the Walmart example, the data will report that Walmart's keyword 'cheap gaming laptop' got 100 displays in

[17]The data set I analyzed was part of the Advertising & Markets Data and, more specifically A3. Yahoo! Search Marketing Advertiser Bid-Impression-Click data on competing Keywords

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

[19]A similar restriction was followed in Athey and Nekipelov (2010)[7], who use Bing data

[20]Various articles show that, apart from Yahoo! consumers on Google and Bing so not go beyond page 1: https://www.conversionguru.co.za/2017/05/29/90-people-dont-go-past-page-1-google-search-results-searching/, https://www.theleverageway.com/blog/how-far-down-the-search-engine-results-page-will-most-people-go/

the 1^{st} position, and that translated into five clicks. Table(1) gives the list of variables used and table(3) gives the summary statistics. Notice that this is similar to feedback given to an advertiser in figure 3. Thus, this data is useful in analyzing how advertisers decide their equilibrium profit-maximizing bid.

Apart from the clicks and display information, a few other important measures can be obtained from the data. First, we can deduce the click rate of each ad,[21] which is measured as the ratio of the number of clicks over the number of displays. The summary statistic show an average of 1% click rate, implying that about 1% of the ads get clicked. Additionally, the keyword (matched words between ad and search) gives an approximation on the type of search. Therefore, the number of words in the keyword referred to as keylength can be used as an approximation for the length of the search query. Previous papers in the literature[22] have noted that longer search queries are typically associated with a more focused search intent and can thus be more valuable for the advertisers. The maximum number of words is 10, with an average keylength of 3 words. Another variable used is the popularity of the keyword, measured by the relative number of searches. This controls for the possible effect of the popularity of the search.

Advertisers side: On the advertisers' side, the data is likewise aggregated on a day level. For each ad,[23] I have information on the bid for the ad, the number of times the ad won an auction, the winning ad position and the total number of advertisers shown in a day.[24] Table 2 lists the variables and table 4 provides the summary. 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 the lowest value, Thus, the bid can be taken as the lower bound on the actual bid. The average bid is 0.6 cents. I limit my analysis to the ads on the first page considering the top 7 ads.[25] These variables help in identifying the auction outcome for each advertiser. Through the data, I can measure how many times an advertiser had a winning ad in the auction, which position he won and what was

[21] ad is defined as the set of keywords for which the advertiser had the same bid on a given day

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

[23] I define ad as the set of keywords-advertiser combination in a day for which advertiser had the same bid

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

[25] On average, seven ads were shown on the page. I assume the number of are seven ads unless the number of positions observed was less than seven.

his corresponding bid.[26] Apart from the given variables, I can also measure the ad specificity in terms of the number of keywords specified for an ad. This gives me an approximation of whether the ad was made for a broader search or a specific search. Additionally, I also use the popularity of the keywords measure by the display frequency.

A note on grouping ads: As the data set has no information on consumer queries, it is hard to know which ads enter the same auction. Nonetheless, the rule through which Yahoo! decides which ads enter an auction can provide useful insights on how ads were matched together. Recall that here, each consumer query is a separate auction. For each query, Yahoo! pulls out relevant ads by matching the ads' keywords with the consumer query. In effect, keywords related to each other enter the same auction. Thus, the paper creates markets that are sub-groups of keywords that are related to each other. I assume the sub-groups are created such that only the ones within a market compete with each other.[27] I assume that the same set of advertisers enter all auctions within a market; this way, the markets can be treated as a proxy for auctions.

Differences across ad product category: Table(5) shows mean values for different variables across categories.[28] The car-insurance category stands out with a high bid level, with an average 4.36 cents. It also has a relatively concentrated market with 20 advertisers per day. The high bid level makes this an important market to analyze from Yahoo!'s perspective. Apart from car insurance, the laptop market is also important to analyze from Yahoo!'s perspective as it is the most popular search category with an average 540.7 search per day. It also seems to have a high level of competition as there are, on average, 45 advertisers per day. Lastly, the TV cable category has the highest high click rate and the second highest per click bid making it yet another profitable market for Yahoo!. Thus, the data provides categories that have differentiating characters and can, therefore, help us check how the results vary across the category.

[26] The real identity of keywords and advertisers are kept confidential by de-identifying the data. This is done to avoid revealing any proprietary information. Also, all the bids are scaled by an unknown amount in order to avoid revealing information about the total revenue of the platform. Even though the data is de-identified, I can still track the same keywords as the same de-identified number is used for all observations.

[27] I give more details on the process of making markets in the estimation section.

[28] Note that the data was masked, so the actual keywords were converted into random numbers and alphabet. I know the name five different categories, which I match to the masked categories using characteristics of the market. Please refer to the appendix(D.3) for more details

3.3 Characterizing features of the data:

This section provides evidence on how the assumptions of the current model fit the observed data. I also look at the variation in the data that might be helpful to identify the parameters of interest.

1. **Winning bid statistics:** A distinct feature of this auction is that each auction has more than one winner. Thus, unlike standard auction designs, where the researcher can only use a single winning bid for each auction, I can use multiple winning bids for each auction. This is especially useful in the last step of the analysis, where I have to use bids to estimate the distribution of advertisers' willingness to pay. Figure (A) shows the mean bids for different winning positions in the five categories.[29]. It seems that in some categories, such as laptops, the bidders are willing to bid substantially higher to get the first position. However, in other categories, such as cable, advertisers are not particularly willing to pay a higher price for a better rank. This difference might be due to the varying sensitivity of consumers to the position of the ad.

2. **Variation in attributes of the ad:**

   To better understand the determinant of how advertisers and consumers value each other. I look at how consumers interact with the ad. One of the attributes of ads is the position of the ad. In general, consumers tend to focus on higher positioned ads than the ones below. To see whether I can disentangle the effect of an ad's characteristics from the effect of an ad position on the consumer's behavior, I need variation in the position allotted to an ad. I observe many ads in the data which are placed at different positions on different days. Using this, in figure 10, I plot the relative click of the same ad in different positions. To compare this across ads, I measure the clicks for each ad relative to the clicks the ad received in the first position. The measure of clicks used here is click-through rate, which is the probability of an ad getting a click. The click-through rate is calculated as the ratio of clicks to impressions. In the estimation section, this is derived from the consumers' choice model.

3. **Externality:** An important observation from the data, is the preliminary evidence of exter-

[29]To consistently graph all the categories in one graph, I have scaled the bid in car insurance category to match the bid range of other categories (only for this graph)

nality. The externality is evident by examining the dependence of bid for a given advertisement on the relevance of other ads displayed in the keyword category. For example, figure 11 shows the linear relation between the average bid and the percentage of high-quality ads in the market.[30] This shows how markets with high-quality ads have a net negative effect on bids. This result gives preliminary evidence that the advertiser chooses to bid less when other high-quality ads appear in the search result page, indicating they take into account the effect of externality on consumer's response to their ad.[31]

4 Industry Model

In this section, I present a model of the advertiser's equilibrium bid and the consumer's click decision using features of weighted Generalized Second Price auction (GSP^w) and discrete choice model. I first specify the click behavior using a discrete choice model. Next, I present a model of advertiser's equilibrium bid, in which the advertisers choose the bid that maximizes their profit from advertising. This step further recovers two parts of the ad value; one externality dependent effect and one advertiser dependent effect (referred to as v_j). The model setup is used to estimate and recover policy invariant parameters such as position-dependent click probability and an advertiser's value for an ad. These can subsequently be used to predict advertisers' behavior as the number of ads per page varies in the counterfactual scenario. In the following sections, advertisers are denoted by $j \epsilon \mathscr{J} := \{1, ...J\}$ and consumers are denoted by $i \epsilon \mathscr{I} := \{1, ...I\}$.

4.1 Consumers' side:

Each consumer i enters the market with 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 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

[30] the the result is obtained regressing bid on the dummy variable for high-quality ads, after controlling for position, category, popularity, keyword effects.

[31] The quality is measured as the ratio of clicks to impressions. In the model, it will be formally derived from the consumer side.

derived click benefit by visible characteristics of the ad. Along with the visible ad characteristics, the consumer also uses the belief that ads at a higher position are of higher quality and relevance. This belief stems from the fact that the search engine's algorithm assigns a higher position to ads with higher quality score, ceteris paribus. Thus, I also add the position as a variable that predicts ad benefit. Another consideration is that in an online space, each click requires the consumer to spend a considerable amount of time on it, which can be thought of as a search cost or time cost.

The expected utility of consumer i receives from clicking in ad j in market m is given as:

$$U_{i,j,m} = U(x_j, k_j, x_i; \eta^j, \eta^i) + \epsilon_{i,j,m} \tag{1}$$

Where $\{\eta^j, \eta^i\}$ are coefficients that reflect how intensely the cost and benefit variables affect the utility. The variables $\{x_j, x_{k_j}\}$ capture the benefit from a click on ad j at position k_j; this includes advertiser specific and position-specific fixed effects, along with ad's popularity measure[32]. Apart from these variables, $\{x_i\}$ are consumer specific variables that help capture heterogeneous search cost; this includes variables such as how detailed is the search (captured by the length of search query[33]). The term $\epsilon_{i,j}$ is the idiosyncratic shock to the consumer's benefit from clicking on the link; it represents a part of the utility which is observed by consumer i, but not by the researcher. I assume $\epsilon_{i,j,m}$ is independently and identically distributed according to type 1 extreme value distribution. Additionally, if the consumer does not click on the ad, she uses her time for an outside good, leading to a normalized utility of $U_{i,0,m} \equiv 0$. Given the utility function in the above equation, I now define the equilibrium click behavior of the consumers.

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.[34] Let $y^*_{i,j}$ denote the binary variable capturing consumer i's equilibrium click decision for ad j , with $y^*_{i,j} = 1$ if consumer decides to click on the ad. Then the click decision

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

[33] This is measured as the number of words in the matched keyword, since I do not have the queries

[34] although most of the literature assume single click per page, this is more realistic situation in this market, as can be seen by a new move by Bing to give 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

can be written as follows:

$$y_{i,j}^{*} = \begin{cases} 1 & \text{if } U(x_j, k_j, x_i; \eta^j, \eta^i) + \epsilon_{i,j,m} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The above equation is used in the empirical section to estimate the probability of click aggregated over all consumers. The estimation gives us the predicted probability of a click for an ad j in position k, represented by $ctr_{j,k}$. Assuming that the ad and position effects are separable[35] the click probability can be rewritten as:

$$ctr_{j,k} = s_j c_k \qquad\qquad \forall\ j\ \epsilon \mathscr{J}\ \&\ \mathrm{k}\ \epsilon\ \{1, 2, ...K\} \tag{3}$$

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 click probability in equation(3) is used in advertiser's maximizing problem as described in the next section.

4.2 Advertising model:

In this section, I first specify the auction design. I then model the advertiser's maximization problem and derive the equilibrium bid.

Auction Setup

Recall that in this market, the advertiser only pays if a consumer clicks on his ad, and does not pay for the ad display. Thus, the auction considers the per click bid and price. For each consumer query, the search engine sells K ad-positions to N advertisers. The K positions are allotted through a weighted Generalized Second Price auction (referred to as GSP^w from now on). Each advertiser is assigned a quality score, s_j, which measures the ad's relevance to the consumers.[36] The advertisers

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

[36] The score can also be thought of as the quality of the ad, as higher consumer relevance is an indication that the consumer perceived it to be of a higher quality.

are ranked in terms of their weighted bids, i.e.

$$b_j^w = b_j \times s_j$$

The auction assigns position in decreasing order of weighted bids, essentially allotting k^{th} ad position to advertiser with k^{th} highest weighted bid. For example, the top position goes to advertiser with the highest weighted bid, second position goes to second highest and so on. The price paid is equal to the bid of advertiser in the next slot weighted by their relative score:[37]

$$\text{Ad position } k \text{ alloted to } j \text{ if } b_j^w = b_w^{[k]}$$
$$\text{Per click price is equal to } p_k = \frac{b_w^{[k+1]}}{s^{[k]}}$$

Here $b_w^{[k]}$ denotes the order statistic of weighted bids such that $b_w^{[k]}$ is the k^{th} highest weighted bid and $s^{[k]}$ denotes the score of advertiser with k^{th} highest weighted bid. Thus, the k^{th} highest position goes to the advertiser with k^{th} highest weighted bid, and he pays the price of the bid below him weighted by their relative score.

Advertisers maximization problem

Note that on the advertiser's side, we are interested in two separate parameters. First, I want to see the effect of externality on advertisers' willingness to pay (i.e. the ad value), and second, I want to derive the exclusive ad value for the advertisers. To do this, I began by establishing a framework of how the externality effect can be separated from the advertisers' effect on ad value. To do this, I use the assumption(2) stated below. I first use this assumption to separate the externality effect on the bid. I then use the externality free bid to derive the exclusive ad value.

Each advertiser j has a per click ad value of V_j. The per click ad value is dependent on the advertiser j as well as the number of other ads next to ad j captured by K_j. I now make a simplifying assumption to separate the effect of externality and the advertiser's effect.

Assumption 2. *Advertiser j's per click ad value is separable in externality effect and advertiser*

[37] I assume no reserve price for simplicity.

effect.

$$V_{j,K_j} = v_j(EXT_j)^{\beta_1} \tag{4}$$

where $v_j \, \epsilon \, \mathbb{R}_+$ is dependent only on the advertiser j. It is independently drawn from the cdf $F(v)$ with support on $[\underline{v}, \bar{v}]$.[38] The second part captures the externality which is depends on other ads EXT_j is a function of K_j, which denotes the set of ads present next to ad j. Note β_1 capture the intensity of externality's affect on ad value. In the empirical section I well estimate β_1 (refer to step(6.2) of the estimation section for more details).

Apart from the above assumption, the following assumptions are also required to solve for the advertisers' equilibrium bid:

1. Advertisers know the distribution of the weighted value, denoted by $\omega_j \equiv s_j \times v_j$, $\omega \sim F_w(.)$.
2. Advertisers know their quality score.[39]
3. Advertisers know the number of ads per page, the average ad quality for ads in each position as well as the total number of advertisers participating in the ad category.

4.2.1 Deriving externality effect on the bid:

The goal of this step is to recover the externality impact on the bid and in turn, on the advertiser's ad value. Using assumption(2), I can derive the relation between externality and the advertisers' bid; the next proposition formally states the result.

Proposition 2. *The additive separability of externality and individual effect on valuation leads to additive separable bid parameter (after controlling for various factors).*[40]

$$Log(\underbrace{b(v_j, s_j; \alpha_{K,n}, Ext_j)}_{\textit{observed bid}}) = Log(\underbrace{b(v_j, s_j; \alpha_{K,n})}_{\textit{externality free bid}}) + \underbrace{\beta_1 Log(Ext_j)}_{\textit{externality effect on bid}} \tag{5}$$

Where the bid $b(v_j, s_j; \alpha_{K,n}, Ext_j)$ gives the bid in the presence of externality, which is dependent on advertiser specific ad value (v_j) and score (s_j) as well as the auction parameters $\alpha_{K,n}$, where

[38] There can be various factors that determine advertiser's value for an ad. Firstly advertiser's attributes effect the value of a ad-click, such as per unit price of the product/service, profit margin for each unit sold.
[39] Search engine reports the quality score to the advertisers, so it is a realistic assumption
[40] proof in the appendix(C)

K denotes the number of ads and n denotes the number of advertisers. The proposition shows that the log of bid in the presence of externality is separable in externality effect $\beta_1 Log(Ext_{j,K_j})$ and the unobserved hypothetical bid if there was no externality in the market $Log(b(v_j, s_j; \alpha_{K,n}))$.

The above proposition will help to estimate the externality in the empirical section, where I estimate β_1 that gives the elasticity of ad value to externality. Another insight from this proposition is that once we know the $\beta_1 Log(Ext_j)$ we can estimate the externality-free bid, i.e. $b(v_j, s_j; n)$. In the next section, I look at the equilibrium bid in case of no externality. The next section helps derive the externality free ad value, i.e. v_j, which can also be thought of as the exclusive ad value.

4.2.2 Advertisers equilibrium bid in the no externality case:

Let us now look at the advertiser's decision of optimal bid in the case of no externality. The advertiser chooses the equilibrium bid that maximizes the expected profit from the ad. The profit from the ad depends on the winning position. Let us look at the ad profit from winning position k, which is equal to the click probability in position k times the per click profit. The per click profit is equal to per ad value minus the per click price at for position k, i.e. $P(k)$. The profit from winning position k is given as follows

$$\left[v_j - \mathbb{E}\left(P(k)\middle|b_j^w = b^{[k]}\right)\right]$$

The per click profit for position k is attained only if advertiser wins position k, which is equal to the weighted bid of advertiser j being the k^{th} highest weighted bid, i.e. $Prob(b_j^w = b^{[k]})$. Thus, the total expected profit sums over the profit from each position weighted by the probability of winning the position. The profit function at the equilibrium bid b is given as follows:

$$\Pi(b; v_j, s_j) = \max_b \sum_{k=1}^{K} Prob(b_j^w = b^{[k]}) * c_{k,j} \left[v_j - \mathbb{E}\left(P(k)\middle|b_j^w = b^{[k]}\right)\right] \tag{6}$$

The profit function shows that the advertiser faces a trade-off, a higher bid will lead to an increase in the probability of getting a click, but it will increase the per click price. We will see that in equilibrium, this trade-off will make advertisers shade their bid below their per click value, i.e. $b_j^* < v_j$. The following proposition solves for the equilibrium bid.

Proposition 3. *There exists a unique symmetric Bayesian Nash equilibrium of the weighted GSP auction.*

The equilibrium bid is given in equation (23). It is excluded here to minimize the introduction of notations. Please see the appendix(B) for more details on the equilibrium bid analysis. The equilibrium bid in equation (23) has some issues regarding empirical identification. First, the structural elements can not be identified using observed parameters in the data; specifically, the equation uses the unknown latent distribution $F(.)$. Secondly, it does not have a closed-form solution.

Bounds of exclusive ad value:

As the equilibrium bid does not have a closed-form, I will instead use bounds on the equilibrium bid to derive bounds on the exclusive ad value. The next proposition bounds the bid in the Generalized Second Price auction in between bid from two more well know auctions that has a closed and easily tractable equilibrium bid. Specifically, I use VickreyClarkeGroves auction (VCG) auction and a Weighted Generalized First Price auction (GFP). The next proposition bounds the bid shading in these three auction designs. Bid shading is the amount by which the advertisers shade their bid below their ad value—$(v_j - b_j)$.

Proposition 4. *The bid shading in GSP^w auction can be bounded between the bid in analog VCG and generalized first price (GFP) auctions.*[41]

$$(v_j - b_j)_{VCG} \leq \underbrace{(v_j - b_j)_{GSP^w}}_{\textit{price less than bid}} \leq \underbrace{(v_j - b_j)_{GFP^w}}_{\textit{price equal to bid}} \qquad \forall j \epsilon \daleth \tag{7}$$

Shows that the bid shading is more than the truthful bidding in VCG; however, it is less than that of the first-price auction. This is because even though the GSP auction bid affects price, the effect on price is still less than that of a first-price auction. To understand this, recall that in GSP, your price is equal to the bid below your bid, whereas, in the analog first-price auction, you pay your bid. Using the above, I can bound the value as follows:

Proposition 5. *If the proposition(4) holds, advertiser's value can be bounded in terms of observed*

[41]proof of proposition (4) and (5) are given in appendix(C)

variables:

$$b_j \leq v_j \leq b_j + \Phi(g_b(.), c_k, b_j) \qquad\qquad \forall j \epsilon \daleth \tag{8}$$

The lower bound is a well known observation in the auction literature that a rational advertiser will not bid more than the value. The upper bound is the bid plus the bid shade amount in case of the GFP auction. The upper bound can be elaborated to be written as follows:

$$b_j + \Phi(g(.), c_k, b_j) = b_j + \frac{\sum_{k=1}^{K} c_k Prob(b_{-j,w}^{[k]} \leq bs_j \leq b_{-j,w}^{[k-1]})}{\frac{d(\sum_{k=1}^{K} c_k Prob(b_{-j,w}^{[k]} \leq bs_j \leq b_{-j,w}^{[k-1]}))}{d(b_j)}} \tag{9}$$

Here $\Phi(.)$ is equal to the bid shading in generalized first price auction, which is equal to the total probability of getting a click over it's derivative. This can be evaluated to be written in terms of the observed bid, the bid distribution and the click through rate. Therefore, the above inequality can be estimated empirically.

5 Identification

In this section, I provide intuition for the identification of the advertiser's externality-free value distribution and the parameters for externality. I also characterize the machine learning algorithm used for creating markets. Note that for the theoretical model I suppressed the market subscript. However for next few sections, I add the subscript m, which represents a market. Note, market is the ad-group and day combination.

5.1 Advertiser's value distribution:

The approach followed to derive upper and lower bounds on advertiser's value distribution uses a combination of two standard method in the auction literature, as discussed in paper (Guerre, Perrigne, Voung, 2000) and (Haile and Tamer 2003). The inequalities derived at the equilibrium bid (Proposition (8) imply for each advertiser i, we have the following bounds:

$$\hat{b}_{j,m}^{ext} \leq v_{j,m} \leq \hat{b}_{j,m}^{ext} + \Phi(\hat{g}(.|\hat{b}), \hat{c}_{k,m}, \hat{b}_{j,m}^{ext}, \hat{s}_{j,m}) \qquad\qquad \forall \quad j, m \tag{10}$$

This gives an upper and lower bound for each advertiser in the sample. Thus, the distribution of advertisers' valuation, $F_v(.)$ can be recovered if I know (1) the estimates of reduced(i.e. externality free) bids (i.e. $\hat{b}^{ext}_{j,m}$); and (2) the bid shading amount, $\Phi(\hat{g}(.|\hat{b}), \hat{c}_{k,m}, \hat{b}^{ext}_{j,m}, \hat{s}_{j,m})$. These parameters are estimated in the initial steps of the estimation process and are then used to derive the distribution bounds in the last step. For now lets assume we can estimate them.

From the bounds on value in equation(10), I derive bounds on its distribution by using stochastic dominance, which recall implies that if $x \leq y \; \forall \; x \; y$, then $F_x() \geq F_y()$. Let $H_b()$ be the distribution of $\hat{b}^{ext}_{j,m}$ and $H_\phi(.)$ be the distribution of $\hat{b}^{ext}_{j,m} + \Phi(\hat{g}(.|\hat{b}), \hat{c}_{k,m}, \hat{b}^{ext}_{j,m}, \hat{s}_{j,m})$. Using stochastic dominance in equation(10) gives the following result:

$$H_\phi(.) \leq F_v(.) \leq H_b(.) \tag{11}$$

Consistent and asymptotically normal estimates of the pointwise upper and lower bounds can be obtained by taking the idealized sample analogs of these endpoints. This is a standard case of nonparametric estimation of a CDF using kernel estimation, which gives us $\hat{H}_\phi(.)$ and $\hat{H}_b(.)$. Further details are given in the estimation section.

Now getting back to the estimates used above, i.e. $(\hat{c}_{k,m}, \hat{s}_{j,m}, \hat{b}^{ext}_{j,m}, \hat{g}(.|b))$. These are discussed in the next two subsections. The first two $(\hat{c}_{k,m}, \hat{s}_{j,m})$ are estimated using consumer's click data in a discrete choice model. The other two $(\hat{b}^{ext}_{j,m}, \hat{g}(.|b))$ are estimated advertisers bid data.

5.2 Consumers click behavior:

I use the utility function of consumers in equation (2) to estimate the predicted probabilities of a consumer clicking on the advertisement. The weighted logit model is used to separately estimate the effect of ad characteristic (i.e. $\hat{s}_{j,m}$) as well as the position (i.e. $\hat{c}_{k,m}$). The data identifies the position and ad characteristic for each ad click by the consumer. I also observe the impressions, which signify the number of times an ad was displayed per period. The variables, impressions, and clicks together help identify the number of clicks and no-click observations per day. Further detail is given in the method section.

5.3 Externality estimation

Externality is defined as the quality weighted sum of number of ads per page, which can be written as $Ext_{j,m} = \sqrt{\sum_{k \neq K} \hat{s}^2_{-j,m}}$, where $\hat{s}$ is the average quality of an ad in position k for all ads except ad j.[42] The quality estimate, $\hat{s}$, is derived from the consumer side analysis[43]. The quality is denoted on 0 to 1 range with 1 being the highest quality. Fig(15) shows the distribution of externality variable across different categories. I use variation in average quality of ads across markets and periods to identify the externality co-efficient.[44] The estimates are then used to calculate the externality-free bid, $b^{ext-free}_{j,m,t} = b_{j,m,t}(Ext_{j,m})^{-\beta_1}$. There is a concern that externality might be endogenous. Even though the quality of rival's ads does not directly impact the advertiser $j's$ bid, there is possibly an indirect effect as the relative quality affects the winning probability. Thus, the variable externality is treated endogenous, and the externality in other markets are used as instruments. The instrument variable is independent of any supply-side effect from the advertisers, and at the same time, it correlates with the demand side i.e. consumer behavior. Thus, it is correlated to the externality but is independent of any effect from the bid decision.[45]

5.4 Creating market using unsupervised machine learning:

Each advertiser specifies a set of keywords that describe his ad. As mentioned earlier, Yahoo! matches these keywords with consumer's query in order to pull out relevant ads. In effect, keywords related to each other enter the same auction. Thus, the paper creates markets that are sub-groups of keywords such that only the ones within a market compete with each other. In this section, I give details of how the markets are created.

The first step of grouping keywords is to find a way to calculate the similarity between different keywords. A keyword is composed of multiple words. These words are used as attributes that take positive value for keywords that contain the word. The similarity measure then assigns a positive score for each matched word and a negative score for each non-matched word.[46] This score is used

[42]this term is weighted by different positions the advertiser gets in a day

[43]This is under the assumption that the quality of an ad can be captured by the ad dependent affect on consumer's click probability. Note that this measure is widely used as quality measure by search engines such as google and yahoo.

[44]The number of ads in most cases is 7 ads for first page

[45]This is similar to the standard practice of using prices in other markets as an instrument for price in IO literature.

[46]I normalize each word count variable by subtracting the mean and dividing by the standard deviation. The

in the machine-learning clustering algorithm to create markets within each category.

Fig(21) shows the visualization of the distance between keywords for the five main categories. The distance measure used is cosine similarity. A key insight that will help increase precision for estimating markets is to use the fact we know the main categories. We will use them to see what is the best method to cluster the markets. Specifically, we will focus on two aspects that are needed for the algorithm. Firstly, the distance metric used between keywords. There are several distance measures, such as Euclidean distance and cosine similarity. I use different distance measures to sort the keywords in the five main categories and see which one works the best. For details, see fig(22), the graph plots the distance between keywords for various distance measures[47]. It turns out cosine similarity measure works the best, coincidentally this is also one the most commonly used distance measure to match text documents in the machine learning literature. Thus, cosine similarity will be the selected choice. The main categories also help in the choice of the clustering method (k-means, Agglomerative Hierarchical Clustering, DBSCAN). All of them perform well, although k-means performs a bit better than others on few parameters, as discussed in the appendix.

The advantage of using unsupervised machine learning is that except the similarity measure and features(words in this case), no other variable needs to be specified. Furthermore, we don't need a measure of actual markets to get a prediction. Remember that here we are creating subgroups(or markets) within a base category, for example, one of the base categories is Laptop. Thus the algorithm might have a tendency to group all of them in one big group as 'laptop' will be a common word for all of them. To solve this, we delete the single word 'laptop' as a matching feature[48]. This means two keywords 'Business Laptop,' and 'Student Laptop' will have zero similarity score as the only word common is 'laptop', which is disregarded in case of making markets. The rest of this section gives details of the k-means algorithm.

After the keywords are processed and vectorized, the markets are defined using k-means clustering. Given a set of keywords $X = \{x_1, x_2, ...x_{N_m}\}$ and an exogenously determined number of group, i.e. M, the algorithm assigns each keyword $x \epsilon X$ to one of the m groups. For each group

procedure puts more emphasis on the overall lower frequent word being matched, under the assumption that such words are a more valuable signal for differentiating a particular set of keywords than words that frequently appear across all keywords.

[47] compressed in two dimensional space

[48] also advertisers specifying 'laptop' as a keyword are grouped according to other keywords specified by them

has a centroid, which is one of the elements from X. A keyword x_l is in group m if and only if the similarity measure is the highest for other $x's$ in m than to those in other groups. Let $\lambda_{l,m}$ denote the allocation variable such that:

$$\lambda_{l,m} = \begin{cases} 1, & \text{if } x_l \text{ is in cluster } m \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Let θ_m be the chosen centroid of group m. The cluster algorithm decides the allocations (λ) and centroids(θ) by maximizing the mean squared distance between points within the group, as shown below:

$$\min_{\theta,\lambda} \left(1 - \sum_{m=1}^{m} \lambda_{l,m} \frac{X.\Theta_m}{||X||||\Theta_m||} \right) \tag{13}$$

The k-means algortihm used in this paper uses cosine distance to calculate similarity between points[49]. Equation(13) is solved recursively. It is a two steps process which is repeated till stable groups are reached. In the first step it optimally selects $\lambda_{l,m}$ for each keyword i given $\theta = \{\theta_1, ...\theta_M\}$ and in second step optimal θ is picked given λ. The algorithm is given as:

$$\begin{aligned} &\text{Step:1} \quad \min_{\lambda} \left(1 - \sum_{m=1}^{M} \lambda_{l,m} \frac{X.\Theta_m}{||X||||\Theta_m||} \right) && \forall i \\ &\text{Step:2} \quad \sum_{l=1}^{N_m} \lambda_{l,m}(x_l - \theta_m) = 0 && \forall m \end{aligned}$$

The steps are repeated until convergence of λ and θ. The optimal number of clusters (i.e. markets). The optimal number of clusters m is determined by repeating the algorithm across a different number of clusters and then use the silhouette score, which provides an average measure of how well each keyword matches with the allotted market, compared to how it matches with other markets.

6 Econometric Method:

In this section, I describe the estimation method. The estimation steps are as follows:

[49]There are several other distance measures that can be used in k-means algorithm

- Step 1: Estimate consumers click probability.
- Step 2: Estimate the externality effect on advertisers bid.
- Step 3: Estimate lower and upper bounds on the advertisers exclusive ad value distribution.

6.1 Step 1: Estimate consumer click probability

In this step, I model the consumers click choice. 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 position.[50] The advertisers' partial effect on click probability is used to create the measure for ad quality. To estimate these two effect I consider a click/no-click binary choice setting using a weighted logit model,[51] where the weights are on the frequency of clicks and no click. As shown in the theory section the consumers i's click choice on ad j is given as follows:

$$y_{i,j} = \begin{cases} 1 & \text{consumer } i \text{ clicks on ad } j \text{ if } U_{i,j} > 0 \\ 0 & \text{otherwise } U_{i,j} \leq 0 \end{cases}$$

The above equation shows that the click decision is captured by the binary variable $y_{i,j} = 0, 1$, where $y = 1$ is consumer decides to click on the ad. It also shows that consumer click on the ad whenever the utility from the click, represented by $U_{i,j}$ is greater than zero. We can use this variable to set up a logit model. The utility from a click can be further elaborated to depend on observable as follows:

$$U_{i,j,m} = \phi_0 + \Phi_1 X_{\text{position dummy}} + \Phi_2 X_{\text{advertiser dummy}} + \Phi_2 X_{\text{controls}} + \epsilon_{i,j,m} \tag{14}$$

where the position dummy captures the ad position effect and advertiser dummy captures the ad effect. Apart from this the control variables include consumer specific variables such as search popularity measure[52] as well as the keylength which captures how detailed is the search. Lastly the

[50] click rate is the expected percentage of clicks received in each position

[51] similar set up is considered in [7]

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

term $\epsilon_{i,j}$ is the idiosyncratic shock which is independently and identically distributed according to type 1 extreme value distribution. The logit model gives us the probability that consumer i chooses to click ad j in market m is:

$$\hat{P}(y_{i,j,m} = 1|x) = \frac{exp(\Phi x_{i,j,m})}{1 + exp(\Phi x_{i,j,m})}$$

where $y_{i,j,m}$ is a binary variable, which equals to 1 if consumer j chooses to click on ad i in market m, and $x_{i,m}$ denotes the set of variables considered in equation(14).

Parameters estimated in step 1:

- Click rate of position k in market m ($\hat{c}_{k,m}$): this is measured as the predicted probability of a click in position k in market m.

- Quality measure ($\hat{s}_{j,m}$): The quality measure used the predicted probability of a click for advertiser j in market m, denoted by $s_{i,m}$. This measure is then scaled to be between $[0, 1]$ by dividing it by the highest value.[53]

- Externality measure: Once we have the quality measure, we can derive the externality measure. Externality for ad j is equal to the weighted sum of the number of other ads, where the weight is equal to the square of the average quality of other ads in each position for market m. Thus, externality in market m for ad j is equal to $\hat{Ext}_{j,m} = \sqrt{\sum_{k \neq K_j} \hat{s}^2_{-j,m}}$[54], where $\hat{s}$ is the average quality of an ad in position k for all ads except ad j in market m.

6.2 Step 2: Estimate the externality effect on advertiser's bid

The goal of this step is to recover the externality impact on the bid. Proposition (2) in the theoretical model provides the equation showing the relationship between externality and equilibrium bid. Using this, I now analyze the relationship empirically. I use a hedonic regression approach suggested by [27] for english auction (this paper shows how to apply for GSP auction setting). It presents equilibrium bid (denoted as $b_{j,m}$) as a function of externality (denoted as $Ext_{m,j}$), the number of bidders, quality of the ad and a vector of observable market characteristics. Thus, using

[53]Note the top 0.01% of the values are considered the highest quality, thus are given value 1 in the 0-1 scale

[54]note that here the subscript k_j is suppressed for simplicity, and the subscript j captures the dependence on ad j

the bid data, the underlying bid for advertiser j in market m can be written as:

$$
\begin{aligned}
log(b_{m,j}) &= \beta_0 + \beta_1 log(Ext_{m,j}) + Z_m\beta_2 + \beta_3\hat{s}_j + \delta_{N,K} + u_{m,j} \qquad (15)\\
\hat{Ext}_{m,j} &= \text{estimated externality in the market}\\
Z_m &= \text{Market characteristics}\\
\delta_{N,K} &= \text{Dummy variables for number of advertisers, ads and ad-position}\\
\hat{s}_j &= \text{estimated ad quality}
\end{aligned}
$$

In above equation, β_1 measures the effect of externality on the advertiser's bid, interpreted as percentage change in bid for 1 more additional ad on the page. I control for market characteristics such as popularity of the keywords of the ads, the keylength as well as the specificity of the ad(measure whether the ad is made for a specific or broad category[55]). Lastly, I control for number of advertisers and number of ads effect by using dummy variables for each one of them.

As noted in the identification section, even though quality of rival's ads does not directly impact the advertiser $j's$ bid, there is possibly an indirect affect as the relative quality affects the winning probability. The resulting correlation between externality and of competitor's weighted bid induces a positive bias in the OLS estimate of β_1. Therefore I estimate externality via instrumental variables, focusing on correlation between externality in different markets for a given period, which gives potential exogenous variation by capturing the demand side variation but being independent of the supply side. My first stage regression in the 2SLS methods is:

$$Ext_{m,j} = \gamma_1 Ext_{m',j} + \Gamma_2 Z_m + \gamma_2\hat{s}_j + \delta_{N,K} + \mu_{j,m} \qquad (16)$$

Parameters estimated

- β_1: This measures the impact of externality on bid. As externality oncreases by 1%, bid will decrease by β_1 percent.
- Externality free bid - the externality-free bid is residual when the externality effect is removed from the observed bid $b_{j,m}^{ext-free}$.

[55]measure by the number of keywords specified for an ad

- The distribution of externality free bid: I can also calculate the distribution of the weighted externality free bid, as follows:

$$\hat{g}(b) = \frac{1}{\delta_g} \sum_j \sum_m 1\{\hat{b}_{j,m}^{[w]ext-free} \leq b\} K\left(\frac{\hat{b}_{j,m}^{[w]ext-free} - b}{\delta_b}\right) \tag{17}$$

6.3 Step 4: Estimate lower and upper bounds on advertiser's externality-free value distribution

This step involves estimating the distribution of advertisers value, $F_v(.)$. Note, that this step uses a nonparametric estimation method since in this step the goal is to estimate a distribution and not a parameter. Additionally, since the distribution is partially identified, meaning that the only the upper and lower bound on the distribution is identified, I would estimate a lower and upper distribution that bound F_v

This step used equation(8) from the theory section,as reproduced below:

$$\hat{b}_{j;m,t}^{ext-free} \leq v_j \leq \hat{b}_{j;m,t}^{ext-free} + \Phi(\hat{g}(.), \hat{c}_k, \hat{b}_j^{ext-free}, \hat{s}_j) \qquad \forall \quad j \quad \epsilon \quad \{1, 2, ...n_{m,t}\} \tag{18}$$

From the bounds on value in equation(10), I derive bounds on its distribution by using stochastic dominance, which recall implies that if $x \leq y \; \forall \; x \; y$, then $F_x() \geq F_y()$. Let $H_b()$ be the distribution of $\hat{b}_{j,m}^{ext}$ and $H_\phi(.)$ be the distribution of $\hat{b}_{j,m}^{ext} + \Phi(\hat{g}(.|\hat{b}), \hat{c}_{k,m}, \hat{b}_{j,m}^{ext}, \hat{s}_{j,m})$. Using stochastic dominance in equation(10) gives the following result:

$$H_\phi(.) \leq F_v(.) \leq H_b(.) \tag{19}$$

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

$$\hat{H}_b(h) = \frac{1}{\delta_h} \sum_j \sum_m 1\{\hat{b}_{j,m} \leq h\} K\left(\frac{\hat{b}_{j,m} - h}{\delta_h}\right) \tag{20}$$

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

$$\hat{H}_\phi \leq F_v \leq \hat{H}_b \tag{21}$$

7 Results

In this section, I discuss the results that are reported in list of tables A.1. The discussion on results is presented separately for each estimation step: step(1) consumer side - click behavior step(2) Externality effect on advertiser's bid (3) Deriving of Advertisers' parameters - advertiser's value for an ad and bid markdown.

7.1 Consumer side :

On the consumer side, I derive consumer's click probability using a weighted logit model with keylength and popularity of search as controls. The main parameters of interest are the effect of position and ad on click probability. To capture the effect of each position and ad I include a dummy variable for position and advertiser.[56]

Position effect on click probability: Table(6) shows results of average partial effect(APE). Fig(12) provides the predicted click probability (referred to as CTR) for different ad positions. As expected, the results show that higher positions have a higher probability of a click. Categories cable, car insurance, and coins see a higher probability of click with an average click rate of around 3%. In graph(13), I plot the ratio of predicted click probabilities (or CTR) for adjacent positions, i.e. $\frac{\text{predicted clicks at position } k}{\text{predicted clicks at position } k+1}$. This ratio helps in understanding the proportional increase in a click when you switch to one position above your current one. It shows that the effect of a switch to the position above you is heterogeneous and dependent on which position you are switching to. For category Car Insurance, it seems the biggest gain is in being in the top 5, as the jump to the 5^{th} position has the highest gain. For laptop top 4 positions have the highest gain, whereas for Cable and Cruise first position holds the most importance. This heterogeneity gain across position is important as it impacts advertisers' bid behavior; an advertiser will be willing to pay more for a higher position if there is a large enough increase in the click rate.

Ad Quality: Measured as the effect of the ad on click probability:- The ad's quality is identified as the predicted click probability for different advertisers. This is obtained by using the predicted click probability for each market and advertiser combination. To compare the quality across advertisers

[56]I include a fixed effect for the market as well

I normalize it to a 0-1 range, with 1 being the highest quality ad.[57] As shown in graph(14), the quality estimates differ across product categories, where some categories are more skewed towards low quality than others. Car insurance seems to have the most skewed quality distribution with the mean of only 0.05, which means that the average ad quality is 5% of the magnitude of highest quality ad. On the other hand, cable and laptop seem to have relatively well distributed ad quality, with an average of 0.3 and 0.2 respectively.

Externality Index: As stated earlier externality variable is the quality weighted sum of the number of other ads on the page. I define the externality index to give it similar properties to the HHI index. Thus, the externality index is calculated using weights equal to the square of quality. This intuitively has an advantage over the linear index as in the current index a single high-quality ad would have a higher effect than two average quality ads. Fig(15) shows the distribution of externality variable across different categories.

7.2 Externality effect:

Table(8) shows results of the impact of externality on advertiser's bid using Eqn(15). The model is estimated to look at non-linear effects by looking at the log of externality effect on the log of the bid. Fig(16) plots the estimated β_1 co-efficient that captures the percentage change in bid when externality is increased by 1% . The results show that 1% increase in externality leads to a decrease of 0.39%, 0.59%, 1.37%, 0.08% in categories car insurance, laptop, cable and coins respectively. The cruise does not show any significant affect of externality. The result in percentage term is difficult to interpret. Thus I will now look at increase in externality in terms of addition of an ad next to an advertiser's own ad. For example, including one more high quality ad when 7 ads are already present on the page decreases bid by 0.9% percent for car insurance, 1.5% in laptop, 3.56% in cable, 0.2% in coins. I find evidence that the externality has a non-linear effect on bid as it depends on the quality of the ad as well as the number of ads already present on the page. The next two paragraphs elaborate on it further.

[57] The top 0.01% are given a score of 1

7.3 Non-linear effect of externality:

- *Effect of quality of the ad:* The externality imposed on advertisers' willingness to pay by the addition of one more ad on the page is influenced by the quality of the additional ad. Graph(17) shows the percentage decrease in the bid when the display goes from an exclusive to two-ad display.[58] The graph shows how the affect of the additional ad depends on its quality.[59] The graph is plotted for the laptop category. A similar pattern is observed in all the other categories as well. Following our earlier example this means, when an ad is included next to Walmart's ad, Walmart decreases their bid by 25% if its a high quality ad from Amazon's, and only 1% if it is a low quality ad.

- *Effect of the number of ads included on the page:* The effect of externality is diminishing in the number of ads added to the page. This means that the addition of the first ad has the highest negative effect. This effect decreases as more ads are included next to an ad. This diminishing effect can be thought of as saturation of the market, where after a certain number, more ads by competitors do not impact the advertiser. The graph(18) plots the decrease in advertisers' willingness to pay due to the addition of an ad for laptop category. For example, the effect of the second ad decrease advertisers willingness to pay by 25%, however, the effect of the seventh ad decrease the advertisers' willingness to pay by only 2%. Thus, this shows that the effect of an additional ad is more when there are fewer ads present on the page.

7.4 Advertiser willingness to pay for an ad and profit margin:

In this section, I analyze the advertiser's benefit from an ad. I deduct the externality effect on the bid from stage 2 results and use the residual of step 2 as the externality-free bid for the analysis in this section. This step derives the distribution for advertisers exclusive ad value.

Advertiser's ad value: Using equation(10), 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

[58] Note that the theoretical setup proves that the percentage decrease in advertisers willingness to pay due externality is same as the percentage decrease in bid. Thus, here I can talk about this interchangeably

[59] Note that with an exclusive ad display there is no externality in the market. Thus, to measure the increase in the externality, I assume the base externality without any ads on the page is one. This is done to show a more realistic impact of externality, because otherwise any amount of externality from the second ad (say high quality or low quality ad) would show a hundred percent increase in externality

shown in the graph(20). The bounds are tight for three out of 5 categories, implying that inequality is sufficient for inference. The graph(19) 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.[60]

An interesting finding in stage 3 results is that the decision of whether to use GSP (Generalized Second-Price auction) or GFP(Generalized First Price auction) does not matter much when the number of bidders is very high. This is usually the case with online markets. This is important as the choice of the auction is discussed intensively in the academic literature as well as the industry. For example, Google switched to GSP from GFP auction, and facebook uses VCG instead of GSP. In this paper, I show I can bound the GSP bid between the VCG bid and the GFP bid. The results show that the bounds are very close to each other, implying the three auctions might give a similar bid. In this case, one might favor the first-price auction as you are paying your bid, thus increasing revenue. I further explore the design of auction by looking at which one gives higher payoff when we include a prior stage of selecting the optimal number of ads.

8 Counterfactual analysis: revenue-maximizing number of ads

In this section, I examine a method to calculate the revenue-maximizing number of ads. I allow the search engine, Yahoo!, to vary the number of ads per page, with the range being from one to seven ads per page.[61] For each case of the number of ads page, I derive the equilibrium bid and thus, the expected revenue for Yahoo!.[62] The revenue-maximizing number of ads is selected as the one that has the highest expected revenue. This counterfactual is done separately for each product category so the selected number of ads can vary across category. This helps us see whether it is a good strategy to set the number of ads different for each category.

The steps below give details of the algorithm used to determine the revenue-maximizing number

[60] The estimated distributions use a boundary condition, which is that the value have to be less than the highest observed bid

[61] The upper bound of the range is selected to be seven as the data showed seven ads per page. Thus, this counterfactual is trying to determine whether seven was optimal or a smaller number of ads was better.

[62] Note, as the equilibrium bid does not have a closed-form I use the upper bound on the bid, which is given by the bid for generalized first-price auction as shown in proposition(4). Additionally these simulations are done assuming everyone has average quality, however this can be extended to look at advertisers with varying score quality. The externality is assumed to be one for base case of exclusive ad.

of ads. For each simulation round, the following steps are executed:

1. Draw N independent values from the empirical distribution:

$$v_j \sim \hat{H}_U(\hat{\phi})$$

2. Solve for equilibrium bid using the empirical estimates of average quality ($\bar{s}$), average click rate ($\hat{c}_k$) and externality co-efficient ($\hat{\beta}_1$).

3. Calculate the revenue (i.e. TR) for each case of number of ads per page:

$$TR(K) \qquad\qquad \forall K = \{1, 2...7\}$$

4. Pick the revenue maximizing number of ads (K^*)

$$K^* = \underset{K}{\text{Argmax}} \quad Mean[TR(K)]$$

These steps are repeated for 1000 iterations. At the end of the iterations, I calculate the mean revenue from each quantity of ads per page. The revenue-maximizing number of ads is the one that has the highest average revenue $K^* = \text{Argmax}_K \quad \frac{\sum_{it=1}^{1000} TR_{it}(K)}{1000}$. The percentage increase in revenue is calculated by taking the difference in the revenue between the selected optimal one and the current number of ads (i.e. seven ads).

8.1 Results

The results from the counterfactual simulation are presented in table(9). The table shows that for three out of the five categories, the number of ads suggested by this algorithm is less than the current number of ads displayed (i.e. less than seven). I also calculate the potential gain in revenue by comparing the revenue from the number of ads suggested by this method to the currently used number of ads (i.e. seven ads). The results show that the largest gain is for the cable TV category of about [15.6%, 22.7%] increase in revenue, with the selected number being two-three ads per page. It seems that the categories that had a high externality co-efficient had the highest gain.

However, apart from the externality effect, the gains also depend on the mean quality of the ads in the category. For example, laptop category, which has a medium level of externality effect and a medium level of average quality, show a higher gain of about [1.014%,2.985%] compared to the car insurance category that had medium externality effect but low average quality and showed a gain of around [0.6942%, 1.698%].

The difference in the revenue-maximizing number of ads across ad product categories show that the features of the ad market might influence the choice of the search engine. Features that determine the bid as well as the externality index play a key role in determining the number of ads per page. This paper suggests that the markets that have a high average quality and more homogeneous products should show lesser number of ads per page compared to other categories.

9 Conclusion

This research looks at the externality generated by the multi-ad display setting on search engines such as Google, Yahoo!, and Bing. The externality in this market is defined as the external cost on an advertiser's willingness to pay for an ad due to the presence of competitors ads on the same page. In particular, I estimate the impact of the number of ads per page on the ad price and then use the estimate to simulate the revenue-maximizing number of ads. I begin by developing a theoretical model to derive the equilibrium ad price in the market. The equilibrium conditions show that the ad price is dependent on the externality effect and the auction design The primary empirical contribution is twofold. The first contribution is the estimation of the impact of externality on the ad price. The findings show that the advertiser's willingness to pay decreases by an average of 18.48% when the display changes from an exclusive to a two-ad display, where the average is across different product categories.[63] This decrease is more if a high-quality ad is displayed compared to a low-quality one. Following the Walmart example from the introduction, this means that the Walmart's willingness to pay decreases by 25% when the display changes from an exclusive to a two-ad display, where the new ad is by Amazon. Furthermore, the decrease is only 2% if the ad next to Walmart is by a low-quality advertiser rather than Amazon. Additionally, the effect on advertisers' willingness to pay is more for the first few ads added next to their ad, implying that

[63]next to a high-quality ad

the effect of an additional ad becomes minimal when there are already five to seven ads on the page.

The second empirical contribution of this paper is that it estimates the advertiser's externality-free willingness to pay (also referred as exclusive ad value). Using the equilibrium conditions, I can estimate bounds on the distribution of the exclusive 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 do a randomized controlled trail 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 mechanism is usually slower. Additionally, frequent changes in the environment can make advertisers leave the ad platform due to 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 the expected revenue for different quantities of ads per page and derive the revenue-maximizing number of ads. The counterfactual analysis shows that the revenue-maximizing number of ads will differ across ad product categories according to the market concentration and the product differentiation. I find that three out of the five categories provided in the data show sub-optimal number of ads. Furthermore, using the suggested number of ads leads to, on average, a 4.5 percent increase in revenue., which translates into a revenue gain of 5.2 billion dollars in revenue.[64]

These counterfactual results can be further combined with other design improvements to increase the expected gain. For example, the restriction on the number of ads can be implemented with an increase in the ad size. Additionally, this research has broader implications as these results can be applied to any online advertising platform, which shows multiple ads on the same ad space or to the same consumer. A few examples are Amazon and Facebook that show multiple ads to

[64] calculated using 2018 Google's ad revenue of 116.3 billion US dollars. - see here for details https://www.statista.com/statistics/266249/advertising-revenue-of-google/

the consumer.[65]

While this paper provides a possible method for determining the optimal number of ads for the search engine, I believe the differentiating features of the ad product category are the key to further analyzing a more intricate revenue-maximizing number of ads per page.

[65] Note that the application is limited to the case of the adjacent ads being from the same ad product category.

References

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

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

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

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

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

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

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

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

[9] Bajari, P., Houghton, S., and Tadelis, S. Bidding for incomplete contracts.

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

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

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

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

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

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

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

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

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

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

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

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

[22] Goldman, M., and Rao, J. Position auctions in practice.

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

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

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

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

[27] Haile, P. A., Hong, H., and Shum, M. Nonparametric tests for common values at first-price sealed-bid auctions. Tech. rep., National Bureau of Economic Research, 2003.

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

[29] Heish, Y., Shum, M., and Yang, S. To score or not to score? structural estimation of sponsored search auctions. Tech. rep., mimeo, 2014.

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

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

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

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

[34] Izmalkov, S., Khakimova, D., and Romanyuk, G. Position auctions with endogenous supply.

[35] Jehiel, P., and Moldovanu, B. Allocative and informational externalities in auctions and related mechanisms.

[36] Jerath, K., and Sayedi, A. Exclusive display in sponsored search advertising.

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

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

[39] Krasnokutskaya, E. Identification and estimation in highway procurement auctions under unobserved auction heterogeneity.

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

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

[42] Lu, Z., and Riis, C. Bayes-nash equilibria in generalized second price auctions with allocative externalities.

[43] Milgrom, P. Simplified mechanisms with applications to sponsored search and package bidding. Tech. rep., Working paper, Stanford University, 2007.

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

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

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

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

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

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

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

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

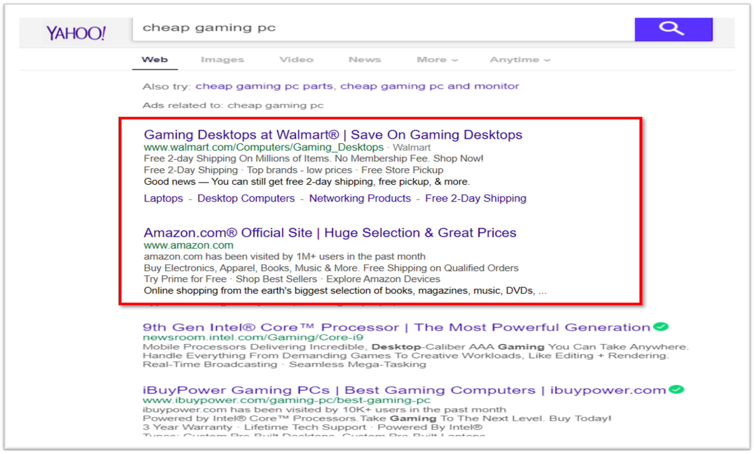
[52] Varma, G. D. Standard auctions with identity-dependent externalities. RAND Journal of Economics (2002), 689–708.

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

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

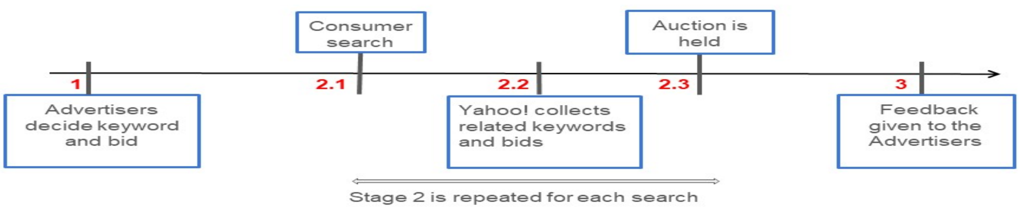
A List of tables

Figure 4



Sample snapshot of ads on Yahoo!, the links in the red box are search ads.

Figure 5: Timeline



The figure shows the stages within each period. The advertiser's decides the bid and keywords at the start of the period, and the bid enters all auctions held in the period. A separate auction is held for each search query. Keywords are words specified by the advertiser that describe the ad and are used by Yahoo! to match it to the search queries.

Figure 6

Search Keywords

1/20/2018 - 7/20/2018

Your top keywords from organic search

EXPORT

Combined traffic data from

Keywords	Clicks from Search	Appeared in Search	Click-Through Rate	Avg Search Click Position	Avg Search Appearance Position
contoso.com $ View Served Pages	1 ↘	9 ↗	11.11 % ↘	2.0 ↘	6.5 ↗
contoso $ View Served Pages	0 →	9 →	0 % →		5.5 →
www.contoso.com $ View Served Pages	0 →	4 →	0 % →		3.2 →
contso $ View Served Pages	0 →	1 →	0 % →		4.0 →

Traffic Details for Keyword contoso.com

Following pages were served for the selected keyword

Served Pages	Clicks from Search	Appeared in Search	Click-Through Rate	Avg Search Click Position	Avg Search Appearance Position
+ http://www.contoso.com/index	1 ↘	9 ↗	11.11 % ↘	2.0 ↘	6.4 ↗

Result Position	Clicks from Search	Appeared in Search	Click-Through Rate
9	0 →	3 ↗	0 % →
2	1 ↘	2 →	50 % ↘
8	0 →	2 →	0 % →
7	0 →	1 →	0 % →
10	0 →	1 →	0 % →

The figure shows a sample snapshot of information available to advertisers. The first one shows aggregated data for each keyword. The other one shows the details for each keyword. Click through rate refers to the ratio of the number of clicks times the number of displays. Refer to this link for more information.

Table 1: **List of variables available for consumer's side data**

Variable	**Description**
Day	The day of the month
Advertiser ID	This gives the masked id for each advertiser, Thus the advertiser can be tracked in this data
Clicks	This gives the number of clicks received by a advertisement in a day at a particular position
Ad Displays	This gives the number of times a advertisement was displayed to a consumer within a day at a particular position
Keyword	The keyword gives the specified words matched between search and the ad
Keylength	The keylength gives the number of words specified in the keyword
Ad position	This gives the position of the advertisement on the search result page

Ad definition id-keyword combination for the same bid gives the advertisement

Table 3: Summary Statistics for consumers' side data

Variable	Mean/Range	Std. Dev/Max
Consumer's side variables:		
Keywords (ad description & search common words)	3174 (count)	-
Keylength (no. of words in keyword)	3.06	10(max)
CTR: Click Through Rate (click percentage)	.98%	9.86
# of Search per Day	1925.6	996.3
Ad-position	1-7	-

[2] Ad Display Frequency and number of ads are on per day level

Consumer side data: Aggregated for each day-advertiser- keyword - ad position observation

Total number of observations: 131524 Restricting data to ads on first page

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

Variable	**Description**
Day	The day of the month
Advertiser ID	This gives the masked id for each advertiser, Thus the advertiser can be tracked in this data
Bid	The bid provided is averaged over day-position-keyword
Ad Displays	This gives the number of times a advertisement was displayed to a consumer within a day at a particular position
Keyword	The keyword gives the specified words matched between search and the ad
Keylength	The keylength gives the number of words specified in the keyword
Ad position	This gives the position of the advertisement on the search result page

Ad definition id-keyword combination for the same bid gives the advertisement

Table 4: Summary Statistics for advertisers' side data

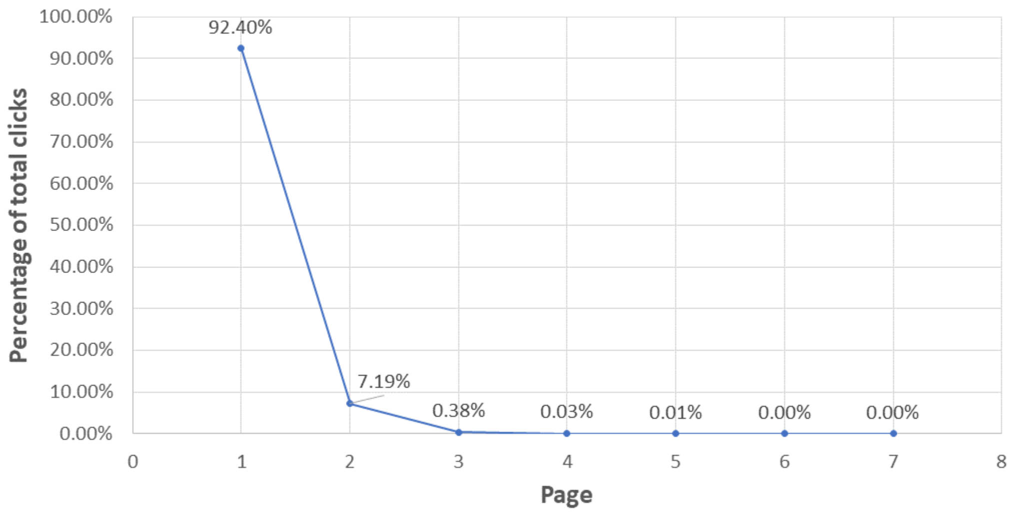
Variable	Mean/Range	Std. Dev/Max
Advertiser's side variables:		
Per Click Bid[1]	.632	2.0
Number of ads per page	6.95	.27
Keywords (ad description & search common words)	3174 (count)	-
Keylength (no. of words in keyword)	3.06	10(max)
Number of advertisers	26.4	10.04
Ad-position	1-7	-

[1] Bid measured in cents

[2] Ad Display Frequency and number of ads are on per day level Advertiser side data:Aggregated for each day-ad position observation

Total number of observations: 21,599 Restricting data to ads on first page

Figure 7: percentage of clicks across pages



The graph plots the percentage of the clicks received across different pages. This also shows why the paper restricts the analysis to first page. As we need to model consumers click behavior limiting the ads to first position captures most of the click decision. The later pages will show less variation in click -no click behavior

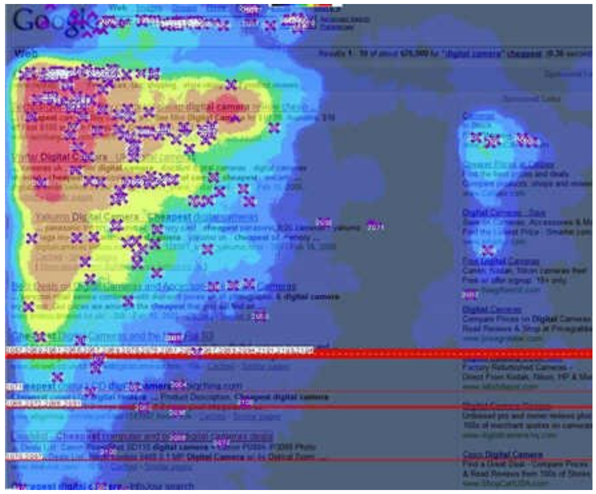
Table 5: Features of Different Categories

Category	Click Rate	Bid	Advertisers	Searches/Day
Car Insurance	2.1%	4.359	20	386.24
Laptop	1.6%	.233	45	540.73
TV Cable	2.4%	.600	25	277.26
Cruise	1.1 %	.371	22	533.84
Coins	1.7%	.174	23	103.52

Table reports mean value of the variable

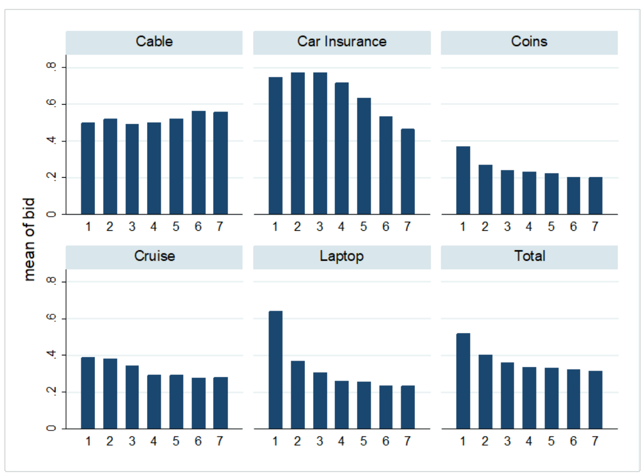
Bid Value in dollar terms: the bid is in cents

Figure 9: Eye Tracking Pattern



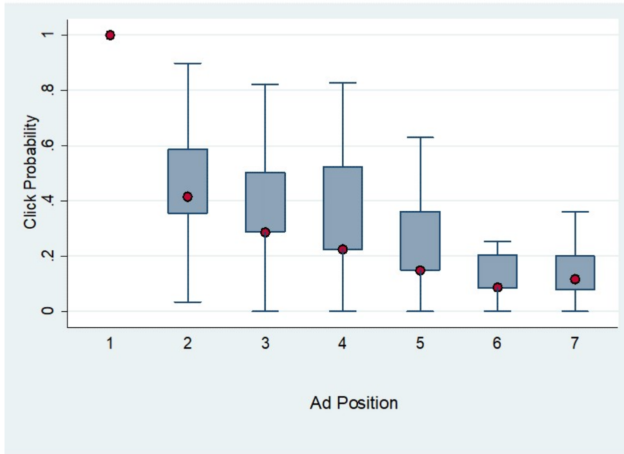
This figure plot the results of a eye tracking research. It shows the eye movement of the consumer on the page. The red region is the on scanned the most followed by yellow and then blue. The picture reconfirms that consumers do a top to bottom scan. Thus, making ads on higher position more valuable.

Figure 8: Mean of bids at different winning positions



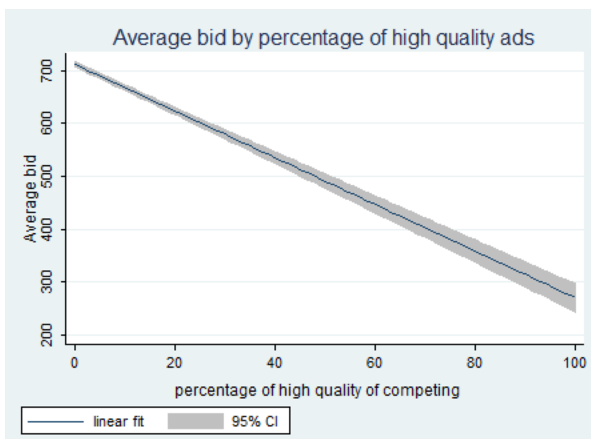
The bid is measured in cents. Due to the large difference in bids for car-insurance and other categories, I have scaled the bid in car-insurance to 1/10th of a cent. (This is only done for this graph and not in the data)

Figure 10: Mean click probability for each ad across position



- The plot shows the relative click of a single ad in different positions.
- The click probability is measured relative to the clicks the ad received in the first position.

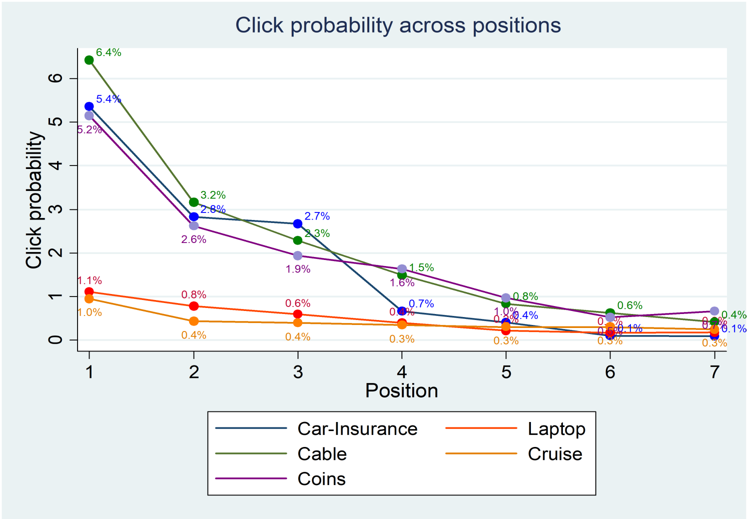
Figure 11: Effect on the bid of competing for high-quality ads



- The plot graphs linear regression of bid on the presence of other high-quality ads on the page
- The regression includes controls for auction and consumer heterogeneity.
- The quality is measured as click probability, which is calculated as the ratio of click to impressions (note in actual estimation this is estimated on consumer side model)

A.1 List of results table:

Figure 12



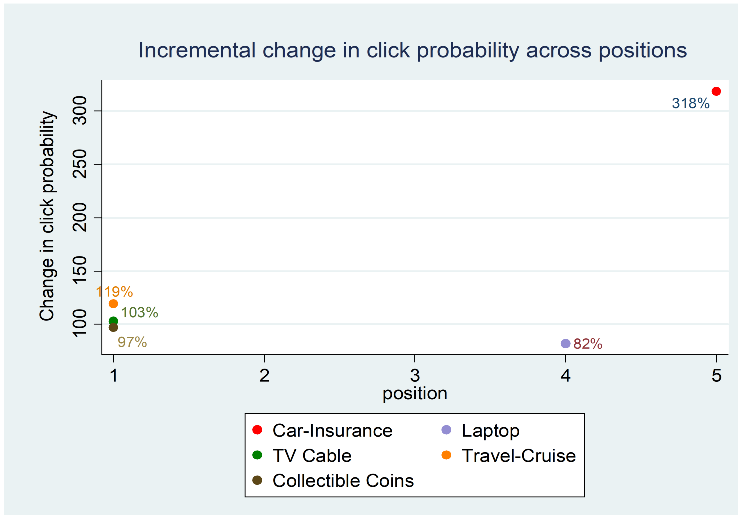
- The plot shows the predicted click probability across positions. - The click probability is the predicted value from the weighted logit model - The values are shown in percentage, for example an ad in position 1 in Cable category gets a predicted click 6.4% of the times it is displayed

Table 6: **Estimated Average partial effect of position on Click probability**

Variable	Laptop	Car-insurance	Cable	Cruise	Coins
2nd Position	-0.00110***	-0.0116***	-0.0228***	-0.00367***	-0.00831***
	(0.0002)	(0.0010)	(0.0015)	(0.0003)	(0.0015)
3rd Position	-0.00180***	-0.0112***	-0.0288***	-0.00399***	-0.0176***
	(0.0002)	(0.0011)	(0.0015)	(0.0003)	(0.0014)
4th Position	-0.00235***	-0.0202***	-0.0343***	-0.00431***	-0.0193***
	(0.0002)	(0.0012)	(0.0015)	(0.0003)	(0.0015)
5th Position	-0.00288***	-0.0214***	-0.0390***	-0.00459***	-0.0233***
	(0.0002)	(0.0012)	(0.0015)	(0.0003)	(0.0015)
6th Position	-0.00301***	-0.0247***	-0.0407***	-0.00455***	-0.0258***
	(0.0002)	(0.0012)	(0.0015)	(0.0003)	(0.0015)
7th Position	-0.00300***	-0.0248***	-0.0422***	-0.00493***	-0.0249***
	(0.0002)	(0.0012)	(0.0015)	(0.0003)	(0.0015)

Aggregation in data: Data aggregated at day-position - keyword level
Advertisement definition id-keyword combination for the same average bid gives the advertisement
Bid Value in dollar terms: the bid is 1/1000 cents, also to protect the sensitive data all bids have an unknown added amount, so we can not say much about the actual amount
Average partial effect on click probability relative to 1st position
Fixed effects for market and advertiser included
.

Figure 13



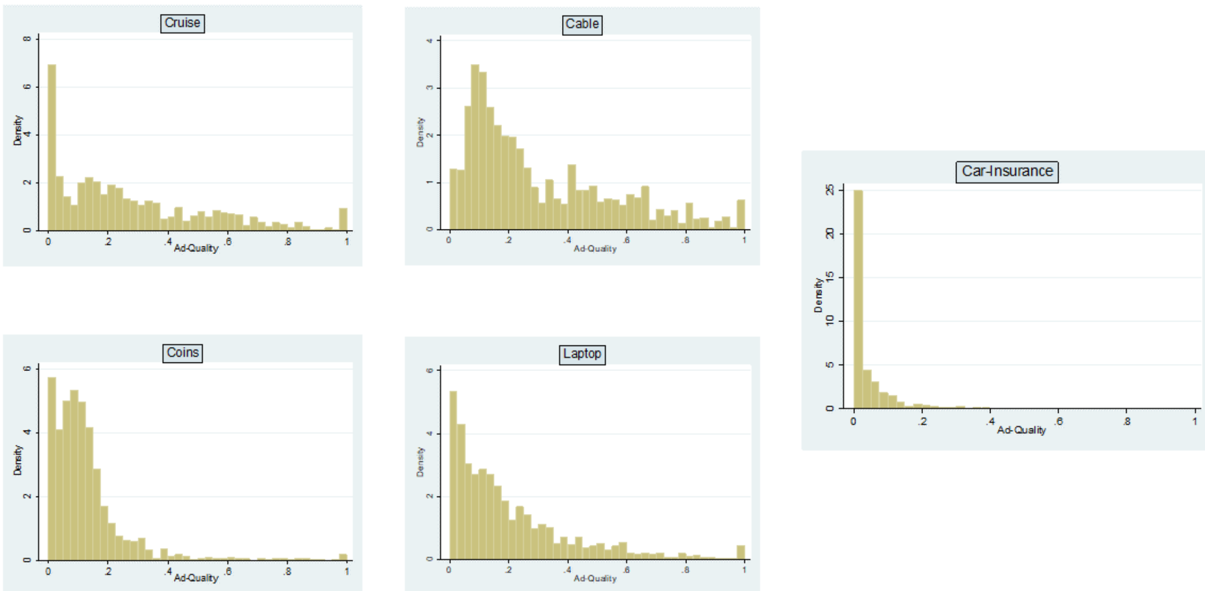
- The plot shows the jump in click probability when the ad moves one position up i.e. from position k to $k-1$. for example, the value in position 1 is the percentage increase in the click probability when you move from 2^{nd} position to 1^{st}.

Table 7: 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

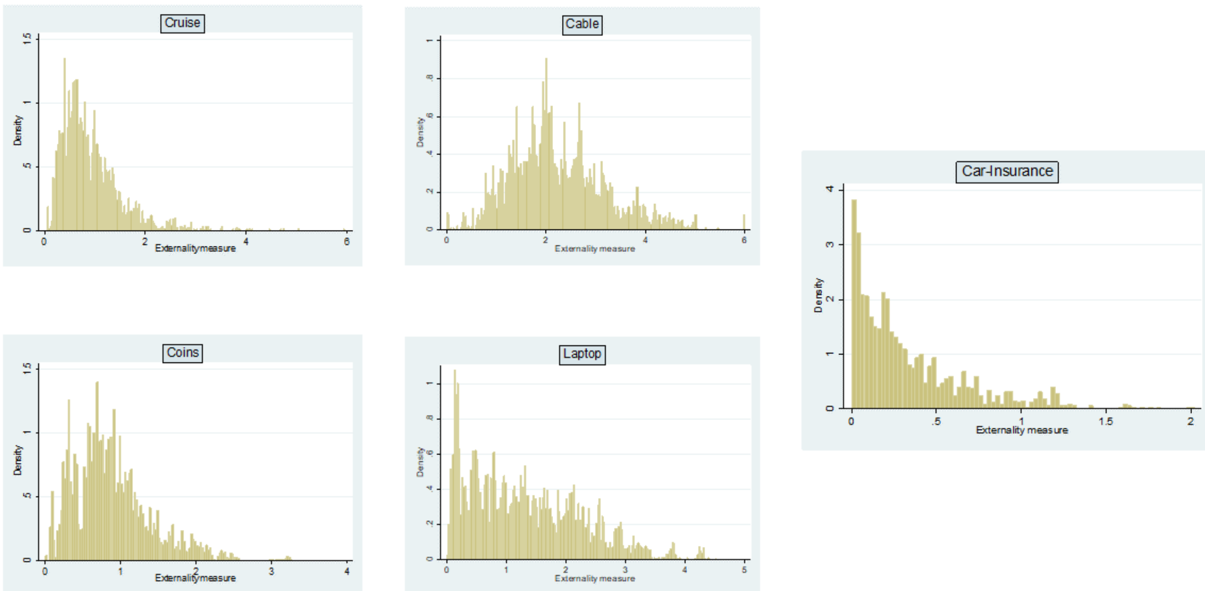
Quality score The score is the predicted click probability of advertisers in each market. It is estimated on the consumer side

Figure 14: Ad-Quality



The plot shows distribution of ad quality across the five categories. The quality score is measured as the advertiser's effect on click rate. It is estimated on the consumer side

Figure 15: Externality

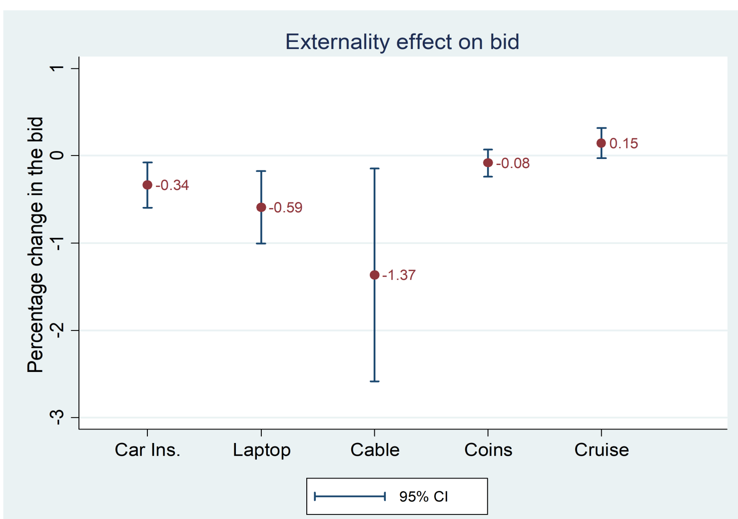


- The plot shows externality index created for different advertisers and market. It plots the distribution for all the five product categories present in the data

Table 8: externality co-efficient across categories

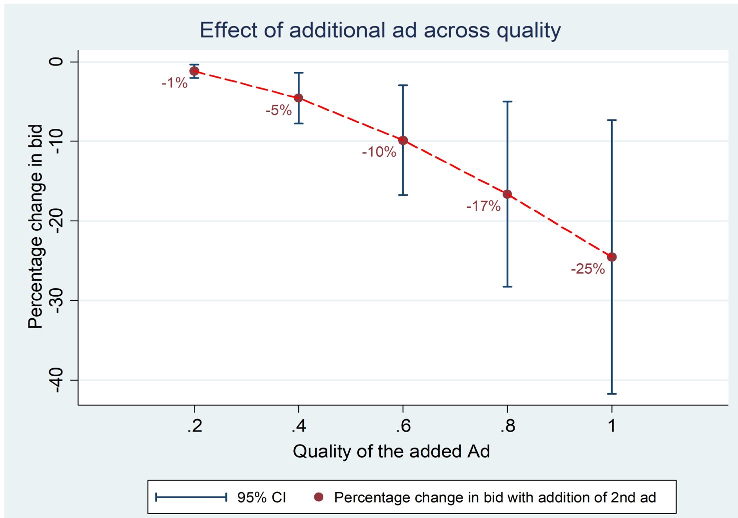
	Log Bid				
	Car Ins	Laptop	Cable TV	Coins	Cruise
Log Externality	-0.336	-0.592	-1.365	-0.0836	0.146
	(0.133)	(0.211)	(0.621)	(0.079)	(0.0879)
Quality	-1.9	-0.618	1.5882	1.071	-1.322
	(0.454)	(0.341)	(0.4865)	(0.1967)	(0.0667)
Ad Popularity	0.0000725	0.000753	0.0019	0.0001	-0.00004
	(0.0000469)	(0.00014)	(0.0002)	(0.00014)	(0.000019)
Keylength	0.0046342	0.006	0.00675	0.0103521	0.007206
	(0.00142)	(0.003)	(0.0019)	(0.0010219)	(0.00043)
Ad Specificity	-0.064	-0.124	-0.076	-0.0945421	-0.0319
	(0.007)	(0.026)	(0.01145)	(0.00537)	(0.00309)
No. of adv FE	X	X	X	X	X
Sub-category FE	X	X	X	X	X
No. of ads FE	X	X	X	X	X
N	729	6977	3219	3589	5041

Figure 16



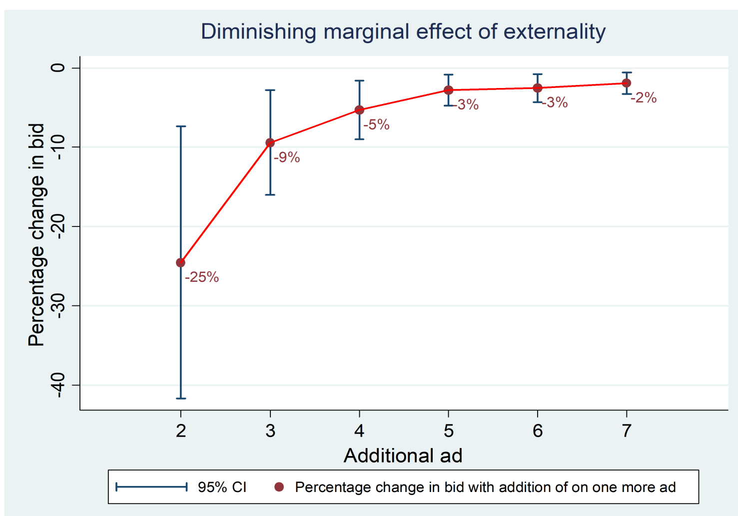
- In this graph, I plot the co-efficient of externality result. The results are from 2SLS log-log regression. - The externality index is created similar to the HHI index, such that it is the square root of the weighted sum of the number of ads, where the weight is the square of average quality.

Figure 17: Effect of an additional ad on advertisers bid across different ad quality



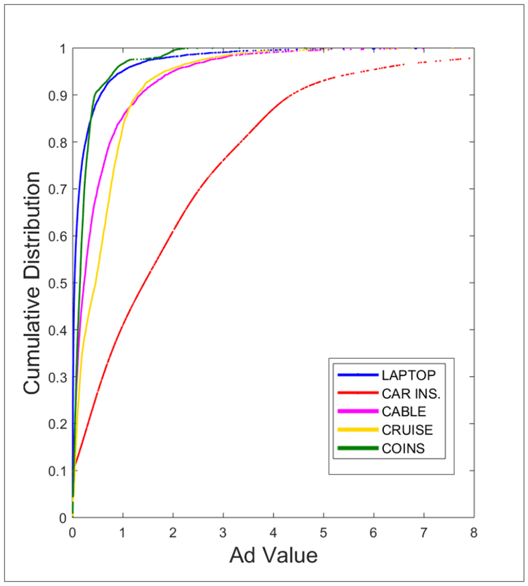
Notes: In this graph, I plot the percentage decrease in a bid for each additional ad. The x-axis shows the quality of the added ad. The graph plots result in the laptop product category. Similar results hold for others as well.

Figure 18: Diminishing effect of an additional ad on advertisers bid



The graph plots results for laptop product category. Similar result holds for others as well.

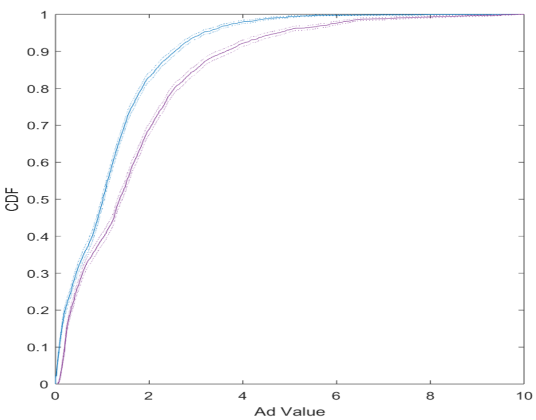
Figure 19: Upper bound for the cumulative distribution function of the ad value



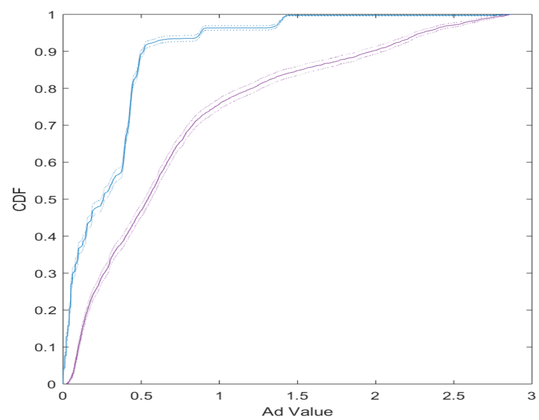
1 The plot shows the upper bound of the estimated distribution of advertisers' ad value.

Figure 20: Cumulative distribution of advertiser's ad value

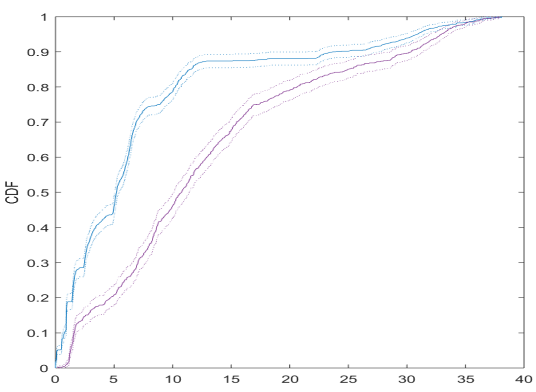
(a) Cable TV



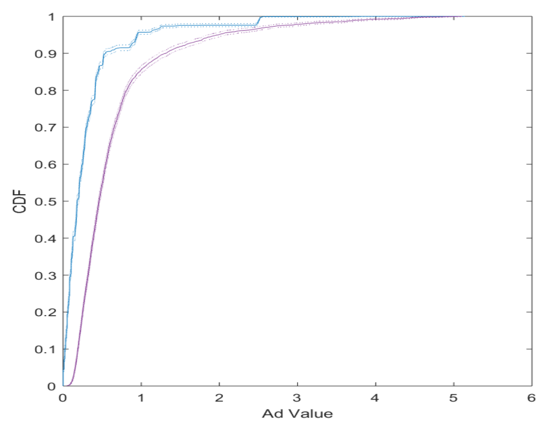
(b) Cruise



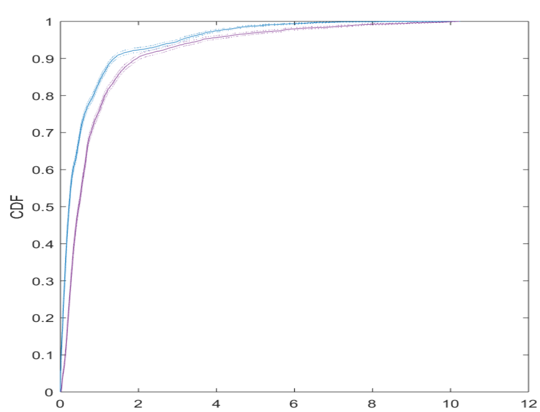
(c) Car Insurance



(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 cumilative distribution at each point. Here ad value captures the advertisers' externality free value that can be thought of as their value for an exclusive ad

A.1.1 counterfactual results

Table 9: Results for the revenue-maximizing number of ads per page

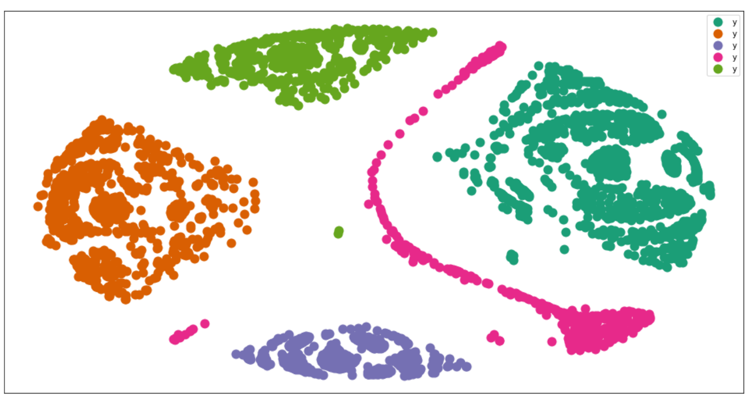
	Predicted number of ads per page		Predicted Change in profit		Externality effect	Quality (average)
	lower bound	upper bound	lower bound	upper bound	(high, medium, low)	(high, medium, low)
Car Insurance	5	3	0.6942 % (0.6940%, .6943%)	1.698% (1.697%, 1.699%)	Medium	Low
Laptop	5	4	1.014 % (1.0139%, 1.0154%)	2.985 % (2.9849%, 2.9868%)	Medium	Medium
TV Cable	3	2	15.575 % (15.5735%, 15.5766%)	22.7 % (22.697%, 22.7036%)	High	High
Coins	7	7	0	0	Low	Low
Cruise	7	7	0	0	Nill	Medium
No. of Simulation rounds	1000	1000	1000	1000	1000	1000

Please refer to section(8) for more details.

Note: The table shows the results from the counterfactual analysis of deriving the revenue-maximizing number of ads. The results are calculated separately for each category and within each category for each upper and lower bound of the ad-value distribution. The table also shows the percentage increase in revenue from using the suggested number of ads as compared to the current number of ads, which is seven. The externality estimate and quality variables are labeled low, medium and high in order to make it simpler to compare categories. The top level among the five categories is given the label high, the next two are given medium and the next two are given low index.

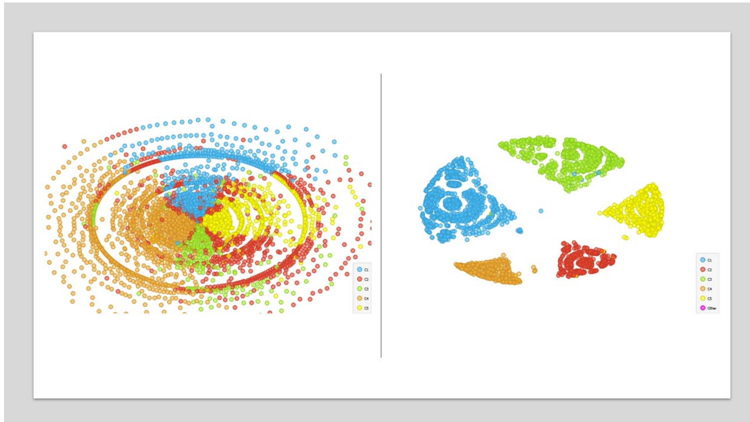
A.1.2 Machine learning (k-means) clustering algorithm

Figure 21: 2-D projection of Keywords across categories



The plot shows keywords in the five main categories. The distance is condensed to represent the two dimensional view. Each scatter point is a keyword. The closer keywords are more related to each other. The five categories are color coded

Figure 22: Keywords for Cosine and Euclidean distance



The plot shows keywords in the five main categories. It shows the relative performance of using different distance measure. the one on the left uses Euclidean distance and the one on the right uses Cosine distance. For the analysis cosine distance is used.

B Equilibrium bid analysis

This model looks at the incomplete information case. Each advertiser does not know bids or value of other advertisers but knows the primitive distributions; namely the distribution of value F_v, the distribution of score F_s as well as the the distribution of a variable later defined as weighted value ($\omega_i \equiv s_i \times v_i$), $\omega \sim F_w(.)$. Additionally the average quality of ads in each position[66], the number of advertisers and ads per page are common knowledge.

For each position, the profit is equal to probability of a click, i.e $c_{i,k}$, times the per click profit, i.e. $\left(v_i - \mathbb{E}(\frac{\beta_w^{[k+1]}}{s_i})\right)$. let $\beta^{[k]}_{-i,w}$ denoted the k^{th} highest weighted bid among all other advertisers except i. In other words , $\beta^{[k]}_{-i,w}$ is the k^{th} highest order statistic among $\beta^1_{-i,w}, \beta^2_{-i,w} .....\beta^n_{-i,w}$. Advertiser i wins position k if $\beta^{[k]}_{-i,w} \leq bs_i \leq \beta^{[k-1]}_{-i,w}$. Given the setup, advertiser i maximizes the sum of per position profit weighted by the probability of winning the position. Thus the equilibrium bid maximizes the following objective function:

$$\beta(v_i, s_i) = Arg\max_{\hat{b}} \sum_{k=1}^{K} Prob(\beta^{[k]}_{-i,w} \leq bs_i \leq \beta^{[k-1]}_{-i,w}) c_{i,k} \underbrace{\left[v_i - \mathbb{E}\left(\frac{\beta_w^{[k+1]}}{s_i} \middle| \beta_w^{[k]} = s_i \hat{b} \right) \right]}_{\text{per click profit at position k}} \tag{22}$$

Firstly notice that the bid function is dependent on the value v_i as well as the score s_i. This paper simplifies this in the next lemma:

Lemma 1. *The weighted GSP auction equilibrium bid function* $b^{GSP^w}(v_i, s_i) \to \mathscr{R}_+$ *is equivalent to*

$$\beta^{GSPM} : (\omega_i) \to \mathscr{R}_+, \ \textit{where } \omega_i = s_i \times v_i$$

[66]The assumption of average quality of the ad in each position is only needed for externality , the BNE equilibrium in incomplete information case does not need this assumption

Also the weighted bid is equivalent to the following function:

$$\beta_w : (\omega_i) \to \mathscr{R}_+, \ \textit{where } \omega_i = s_i \times v_i$$

Let $\omega_i = s_i \times v_i$ be referred as the weighted value. The lemma shows that the bidding function is equivalent to a function that is depend only on one-dimensional advertiser's type ω_i. Additionally it shows that at equilibrium we can rewrite the weighted bid $s_i\beta_i$ as a function of weighted value , i.e. $\beta_w(\omega_i)$. This simplification comes in handy for proof of bounds and equilibrium as now inverse of bid function is one dimension.

Proposition 6. *The unique symmetric Bayesian Nash equilibrium of the weighted GSP auction is given by the following bidding strategy:*

$$\beta^*(\omega) = \omega - \Gamma(\omega) - \sum_{n=1}^{\infty} \int_0^{\omega} M_n(\omega, t)\phi(t)dt \qquad \forall \omega \sim F_w(.) \tag{23}$$

where

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

The above proposition shows the equilibrium bid[67]. Proof : From lemma(1) the advertiser type given by can v_i, s_i can be rewritten as ω_i. Thus, this problem becomes the non-weighted gsp auction with type ω_i instead of v_i, the proof of the equilibrium then follows using paper- Gomes and Sweeney (2014) [25].

C Theoretical Proofs

C.1 Proof for proposition(2)

The profit maximizing objective function for the case of no externality is given as below :

$$\Pi(b; v_j, s_j) = \max_b \sum_{k=1}^{K} Prob(b_j^w = b^{[k]}) * c_{k,j} \left[v_j - \mathbb{E}\left(P(k) \middle| b_j^w = b^{[k]} \right) \right] \tag{24}$$

Let b^* be the equilibrium bid of the no externality case, then I show that $b^*(EXT_{K_j})^{\beta_1}$ is the equilibrium bid of externality case. Now in case of externality the ad value is $V_{j,K_j} = v_j(EXT_{K_j})^{\beta_1}$ instead of v_j. Substituting this in the equation above , I get the profit maximizing objective function in the presence of

[67] 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 [25]

externality.

$$\Pi(b; v_j, s_j) = \max_b \sum_{k=1}^{K} Prob(b_j^w = b^{[k]}) * c_{k,j}\left[v_j(EXT_{K_j})^{\beta_1} - \mathbb{E}\left(P(k)\middle|b_j^w = b^{[k]}\right)\right] \tag{25}$$

Note that the above transformation does not affect the ranking of the ad values and thus they also do not impact the ranking of the bid.[68]. Thus, This can be further solved to:

$$\Pi(b; v_j, s_j) = (EXT_{K_j})^{\beta_1}\left[\max_b \sum_{k=1}^{K} Prob(b_j^w = b^{[k]}) * c_{k,j}\left[v_j - \mathbb{E}\left(P(k)\middle|b_j^w = b^{[k]}\right)\right]\right]$$

Thus, the solution to the above objective function is $b^* * (EXT_{K_j})^{\beta_1}$, which is the equilibrium bid in the presence of externality. Thus, the externality bid given by $b(v_j, s_j; \alpha_{K,n}, ext)$ is equal to $\underbrace{b(v_j, s_j; n, \alpha_{K,n})}_{\text{externality free bid}} * (EXT_{K_j})^{\beta_1}$.

Taking log this implies the following:

$$Log(\underbrace{b(v_j, s_j; \alpha_{K,n}, ext)}_{\text{observed bid}}) = Log(\underbrace{b(v_j, s_j; n, \alpha_{K,n})}_{\text{externality free bid}}) + \underbrace{\beta_1 Log(Ext_{j,K_j})}_{\text{externality effect on bid}} \tag{26}$$

□

C.2 lemma(1) proof:

Recall that in the weighted GSP auction, the advertisers report per click bid. The equilibrium bid β^{GSP} for advertiser i is given as :

$$\beta^{GSP} = Arg \max_{\hat{b}} \Pi(\hat{b}|v_i, s_i) = \max_{\hat{b}} \sum_{k=1}^{K} \gamma_i c_k \left[v_i - \frac{\mathbb{E}\left(\beta_w^{GSP,[k+1]}\middle|b_w^{[k]} = \hat{b} * s_i\right)}{s_i}\right]$$
$$\times Prob(\beta_w^{GSP,[k+1]} \leq \hat{b} * s_i \leq \beta_w^{GSP,[k+1]}) \tag{27}$$

Consider an alternative auction, which I refer GSP modified (GSPM). In this auction the advertisers report bid for s_i number of clicks, where the number of clicks is equivalent to the advertiser's quality score. The equilibrium bid β^{GSPM} for advertiser i is given as :

$$\beta^{GSPM} = Arg \max_{\tilde{b}_w} \Pi(b_w|v_i, s_i) = \max_{\tilde{b}_w} \sum_{k=1}^{K} \gamma_i c_k \left[v_i - \frac{\mathbb{E}\left(\beta_w^{GSPM,[k+1]}\middle|b_w^{[k]} = \tilde{b}_w\right)}{s_i}\right]$$
$$\times Prob(\beta_w^{GSP,[k+1]} \leq \hat{b} * s_i \leq \beta_w^{GSP,[k+1]}) \tag{28}$$

[68]assuming monotonic bid

$\tilde{b} = \frac{\hat{b}_w}{s_i}$, then we can rewrite the optimizing problem as

$$\beta^{GSPM} = Arg \max_{\tilde{b}} \Pi(b_w|v_i, s_i) = \max_{\tilde{b}} \sum_{k=1}^{K} \gamma_i c_k \left[v_i - \frac{\mathbb{E}\left(\beta_w^{GSPM,[k+1]} \middle| b_w^{[k]} = \tilde{b} \right)}{s_i} \right] \times Prob(\beta_w^{GSP,[k+1]} \leq \hat{b} * s_i \leq \beta_w^{GSP,[k+1]}) \tag{29}$$

Now I will use the information that s_i is know to advertiser i and the auctioneer. Thus, the above problem can be rewritten to maximize $\check{b} = \tilde{b}/s_i$ so that the optimal bid per click looks like :

$$\beta^{GSPM} = Arg \max_{\check{b}} \Pi(\check{b}|v_i, s_i) = \max_{\check{b}} \sum_{k=1}^{K} \gamma_i c_k \left[v_i - \frac{\mathbb{E}\left(\beta_w^{GSPM,[k+1]} \middle| b_w^{[k]} = \check{b} * s_i \right)}{s_i} \right] \times Prob(\beta_w^{GSPM,[k+1]} \leq \check{b} * s_i \leq \beta_w^{GSPM,[k-1]}) \tag{30}$$

The optimization problem in equation(28) and (30) are equivalent, Thus, if all other advertisers $j \neq i$ have $\beta_j^{GSP} = \frac{\beta_j^{GSPM}}{s_j}$, then advertiser will also have $\beta_i^{GSP} = \frac{\beta^c GSPM_i}{s_i}$. This shows that the GSPM equilibrium is one of the equilibrium for GSP, but since GSP has unique equilibrium, it implies the two auctions give the same equilibrium bid.

C.3 Proof of proposition (4):

Suppose bidder i bids $\tilde{b}$ and everyone else is playing according to the equilibrium increasing bidding strategy $b(v_i, s_i)$. This equivalently means that bidder i has weighted bid $\tilde{b}_w$ and everyone else bids the equilibrium weighted bid $b_w(w_i)$ (refer to lemma(1) for more details). Recall that $F_w(.)$ is the distribution of the weighted value ω and $G_w(.)$ is the distribution of the equilibrium weighted bid b_w. To solve for the equilibrium we first elaborate on the probability of getting a position. The probability of winning position k can be written as :

$$z_k(\tilde{b}_w) = \sum_{k=1}^{K} \frac{(N-1)}{(K-1)} (1 - G(\tilde{b}_w))^{k-1} (G(\tilde{b}_w))^{N-k} \tag{31}$$

$$\tag{32}$$

as the equilibrium weighted bid is an increasing function of weighted value w_i, the above is equivalent to

$$z_k(b_w^{-1}(\tilde{b}_w)) = \sum_{k=1}^{K} \frac{(N-1)}{(K-1)} (1 - F(b_w^{-1}(\tilde{b}_w)))^{k-1} (F(b_w^{-1}(\tilde{b}_w)))^{N-k} \tag{33}$$

$$\tag{34}$$

Consider an efficient equilibrium , then the profit function is given as :

$$\sum_{k=1}^{K} c_k z_k(b_w^{-1}(\tilde{b}_w))\Big[v_i - \mathbb{E}\Big(\frac{b_w^{[k+1]}}{s_i}\Big| b_w^{[k]} = \tilde{b}_w\Big)\Big]$$

Substituting value of $z_k(b_w^{-1}(\tilde{b}_w))$ from equation(33) and then differentiating we get the above equation we get :

$$\begin{aligned}
&\Big(v_i - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\Big)\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(b_w^{-1}(\tilde{b}_w))^{k-1})F(b_w^{-1}(\tilde{b}_w))^{N-k-1} f(b_w^{-1}(\tilde{b}_w))b^{'}(b_w^{-1}(\tilde{b}_w)) \\
&+ c_k \frac{(N-1)}{(K-1)}(k-1)(1-F(b_w^{-1}(\tilde{b}_w)))^{k-2}F(b_w^{-1}(\tilde{b}_w))^{N-k} f(b_w^{-1}(\tilde{b}_w))b^{'}(b_w^{-1}(\tilde{b}_w))\Big] \\
&- \sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(b_w^{-1}(\tilde{b}_w)))b^{'}(b_w^{-1}(\tilde{b}_w))^{k-1}(F(b_w^{-1}(\tilde{b}_w)))^{N-k} \frac{d\Big(\mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\Big)}{d(b)} = 0
\end{aligned}$$

I focus on the symmetric equilibrium where $b_w() = \tilde{b}_w()$ thus, the above can be rewritten as (35)

$$\begin{aligned}
&\Big(v_i - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\Big)\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_i)^{k-1})F(\omega_i)^{N-k-1} f(\omega_i)b^{'}(\omega_i) \\
&+ c_k \frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_i)^{k-2})^{N-k} f(\omega_i)b^{'}(\omega_i)\Big] \\
&- \sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_i))b^{'}(\omega_i)^{k-1}(F(\omega_i))^{N-k} \frac{d\Big(\mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\Big)}{d(b)} = 0 \qquad (36)
\end{aligned}$$

To further solve the last term lets first open up the expected price:

$$\begin{aligned}
\mathbb{E}\Big(\frac{b_w^{[k+1]}}{s_i}\Big| b_w^{[k]} = \tilde{b}_w\Big) &= \mathbb{E}(\frac{b_w^{k+1:N}}{s_i}| b^{k+1:N} \leq \tilde{b}_w \leq b^{k-1:N}) \\
&= \mathbb{E}(\frac{b_w^{1:N-k}}{s_i}| b^{1:N-k} \leq \tilde{b}_w) \\
&= \int_0^{\tilde{b}_w} \frac{x}{s_i} \underbrace{\frac{(N-k)F^{N-k-1}(b^{-1}(x))g(b^{-1}(x))}{F^{N-k}(b^{-1}(\tilde{b}_w))}}_{\text{conditional distribution}} dx
\end{aligned}$$

Using integration by parts, we get :

$$\begin{aligned}
&= \frac{\tilde{b}_w}{s_i}\int_0^{\tilde{b}_w} \frac{(N-k)F^{N-k-1}(b^{-1}(x))f(b^{-1}(x))}{F^{N-k}(b^{-1}(x))}dx - \int_0^{\tilde{b}_w} \frac{F^{N-k}(b^{-1}(x))}{s_i F^{N-k}(b^{-1}(\tilde{b}_w))}dx \\
\Rightarrow \mathbb{E}\Big(\frac{b_w^{[k+1]}}{s_i}\Big| b_w^{[k]} = \tilde{b}_w\Big) &= \frac{\tilde{b}_w}{s_i} - \Gamma(\tilde{b}_w, f(.)) \qquad (37)
\end{aligned}$$

As can be expected the price in generalized second price auction is less than the bid and the decrease in the bid is defined by $\Gamma(\hat{b}, s_i, g(.))$ which is equal to $\int_0^{\hat{b}s_i} \frac{G^{N-k}(x)}{s_i G^{N-k}(b)}dx$.substituting the expected price in equation(36) we get:

$$\left(v_i - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)b^{'}(\omega_i)$$

$$\times f(\omega_i)\Bigg[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_i)^{k-1})F(\omega_i)^{N-k-1} + c_k\frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_i))^{k-2}F(\omega_i)^{N-k}\Bigg]$$

$$-\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_i))^{k-1}(F(\omega_i))^{N-k}\frac{d\left(\frac{\tilde{b}_w}{s_i} - \Gamma(\tilde{b}_w, f(.))\right)}{d(b)} = 0$$

$$\Rightarrow\left(v_i - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)$$

$$\times b^{'}(\omega_i)f(\omega_i)\Bigg[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_i)^{k-1})F(\omega_i)^{N-k-1} + c_k\frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_i)^{k-2})F(\omega_i)^{N-k}\Bigg]$$

$$-\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_i)^{k-1}(F(\omega_i))^{N-k}\left(\frac{1}{s_i} - \frac{d(\Gamma(\tilde{b}_w, f(.)))}{d(b)}\right) = 0$$

$$\Rightarrow v_i = \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})$$

$$+\frac{\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_i))^{k-1}(F(\omega_i))^{N-k}\left(1-\frac{d(\Gamma(\tilde{b}_w,f(.)))}{d(b)}\right)}{s_i b^{'}(\omega_i)f(\omega_i)\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_i))^{k-1}F(\omega_i)^{N-k-1} + c_k\frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_i))^{k-2}F(\omega_i)^{N-k}\Big]}$$

using $1 - \frac{d(\Gamma(\tilde{b}_w,f(.)))}{d(b)} < 1$ and $atequilibriumb^{'}(\omega_i) = \frac{1}{b^{'}(\omega_i)}$ I get

$$\leq \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})$$

$$+\frac{\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_i))^{k-1}(F(\omega_i))^{N-k}}{s_i\frac{f(\omega_i)}{b^{'}(\omega_i)}\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_i)^{k-1})F(\omega_i)^{N-k-1} + c_k\frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_i)^{k-2}F(\omega_i)^{N-k})\Big]}$$

using the auction property that bid is always greater than the price, i.e. $\mathbb{E}(\frac{b_w^{[k+1]}}{s_i}) < b_w$

$$\Rightarrow v_i \leq \frac{b_w}{s_i} + \frac{\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_i))^{k-1}(F(\omega_i))^{N-k}}{s_i\frac{f(\omega_i)}{b^{'}(\omega_i)}\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_i)^{k-1}F(\omega_i)^{N-k-1} + c_k\frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_i)^{k-2}F(\omega_i)^{N-k}\Big]} \tag{38}$$

Now to use this inequality in estimation part I will need to substitute the bid distribution in place of latent distribution using the following equality conditions :

$$G(\beta_w) = F(\omega|N)$$
$$g(\beta_w) = \frac{f(\omega|N)}{\beta'(\omega)}$$

Thus, equation(38) can be written as :

$$v_i \leq \frac{b_w}{s_i} + \frac{\sum_{k=1}^{K} c_k \frac{\binom{N-1}{}}{\binom{K-1}{}} (1 - G(\beta_w))^{k-1} (G(\beta_w))^{N-k}}{s_i g(\beta_w) \Big[\sum_{k=1}^{K} c_k \frac{\binom{N-1}{}}{\binom{K-1}{}} (N-k)(1 - G(\beta_w))^{k-1} G(\beta_w)^{N-k-1} + c_k \frac{\binom{N-1}{}}{\binom{K-1}{}} (k-1)(1 - G(\beta_w))^{k-2} G(\beta_w)^{N-k} \Big]}$$

where Δ is defined as $\frac{\sum_{k=1}^{K} c_k \frac{\binom{N-1}{}}{\binom{K-1}{}} (1-G(b_w(v_i,s_i)))^{k-1} (G(b_w(v_i,s_i)))^{N-k-1}}{s_i \Big[\sum_{k=1}^{K} c_k \frac{\binom{N-1}{}}{\binom{K-1}{}} (N-k)(1-G(b_i^*))^{k-1} G(b_i^*)^{N-k-1} + c_k \frac{\binom{N-1}{}}{\binom{K-1}{}} (k-1)(1-G(b_i^*))^{k-2} G(b_i^*)^{N-k} \Big]}$

$\square$

Next I also show that the upper bound is equal to the equilibrium bid of the analog generalized first price auction. The generalized first price auction (GFP) will have agents pay their bid for different positions, since in this case auctioneer considers a score, the agents pay their weighted bid. THe maximizing profit function would be

$$b^{GFP}(v_i, s_i) = Arg \max_{\hat{b}s_i} \sum_{k=1}^{K} c_k z_k(\hat{b}s_i) \Big[s_i v_i - b_i s_i \Big]$$

the equilibrium bid will then be :

$$\sum_{k=1}^{K} c_k \frac{d(z_k(\hat{b}s_i))}{d(b)} \Big[s_i v_i - b_i s_i \Big] - \sum_{k=1}^{K} c_k z_k(\hat{b}s_i) = 0$$
$$\rightarrow b^{GFP} = v_i - \frac{\sum_{k=1}^{K} c_k z_k(\hat{b}s_i)}{\sum_{k=1}^{K} c_k \frac{d(z_k(\hat{b}s_i))}{d(b)}}$$
$$\rightarrow b^{GFP} = v_i - \frac{G(b_w)}{\Delta g(b_w)}$$

It can be shown that in case of one position , this equilibrium bid is equal to the equilibrium bid of first price auction.

D Revenue-maximizing number of ads: Looking at flexible number of ads

The empirical exercise has shown that the advertiser's account for the externality imposed by other advertisers' presence on the ad space. The bids of the advertiser decrease with an increase in the number of ads on the page. The natural question then is what is the optimal number of advertisements on the ad space. In this section, I propose an addition to the auction mechanism, that can derive the optimal number of ads on the page. I propose a new mechanism that adds a pre-auction stage where the optimal number of ads are decided. following are the steps of involved in the mechanism :

- Bidder:
  1. Bidders submit a 2-dimension bid. [69]
  2. The bid is composed of the bid for an exclusive ad option and a specified percentage decrease in the bid for each additional advertisement on the same page. [70]
- auction mechanism
  1. step 1: Optimal number of ads is decided
  2. step 2: GSP auction is held

The proposed auction can derive the optimal number every time an auction is held; in other words, every time the search is entered in Yahoo!. Using the estimates of the advertisers' value and externality estimate, we can compare the old and new pricing mechanisms using simulation. This has been left for future research.

D.1 Simulation details:

I now specify the steps involved in the simulation for deciding the revenue-maximizing number of ads. The steps are repeated for 100 simulation rounds, and the simulation is done separately for each category.

1. Draw N independent values from the empirical distribution:

$$v_j \sim \hat{H}_U(\hat{\phi})$$

2. Solve for equilibrium bid using:
   Quality ($\bar{s}$), click rate ($\hat{c}_k$) and externality co-efficient (β_1).

3. Pick the revenue-maximizing number of ads:

$$N^* = \underset{N}{\text{Argmax}} \quad TR(N)$$

[69] Here, I am assuming the externality, i.e./ the percentage decrease in the bid due to each additional bid, is not known to the auctioneer. An the alternative specification would be the advertisers only report the bid, and the auctioneer applies the pre-specified percentage decrease

[70] the percentage the decrease is equal to the externality calculated in the step 1

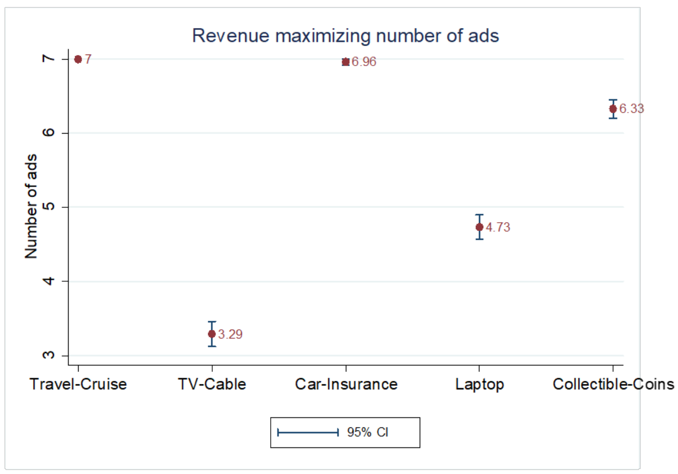
4. Compare Yahoo!'s revenue in revenue-maximizing N^* and the average seven ads:

$$\Delta(gain) = TR(N^*) - TR(7)$$

D.2 Results

- *The revenue-maximizing number of ads:* The revenue-maximizing number of ads changes with each category. The graph(23) plots the average number of ads allotted in the new auction design, along with the confidence interval. The product categories cable TV and laptop are shown to have the highest decrease in the number of ads shown compared to the current norm of showing an average of 7 ads.

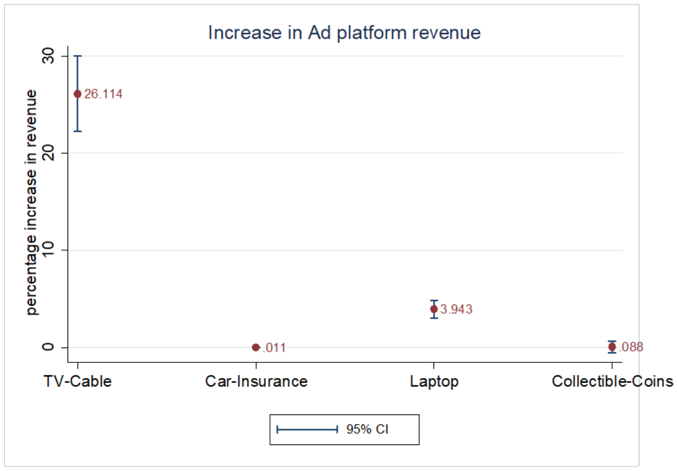
Figure 23: revenue-maximizing number of ads across categories



1 The plot shows the average number of ads that are displayed in the new auction design that selects the number of ads as well as the winning ads.

- *Revenue gain across categories:* The new auction design leads to higher revenue for the search engine. Graph(24) plots the average gain in revenue across categories. The highest gain is observed in the cable TV, followed by the laptop. Notice that although laptop and car insurance had similar co-efficient for the externality effect, the gains are different. This is because the gain also depends on the quality of the ad. The average quality is higher for laptop category than car insurance.

Figure 24: Gain in revenue from using the new method



D.3 Within product category variation :

Recall that this data has varied product categories, namely 'laptop', 'cable', 'cruise', 'coin' and 'car insurance'. Although we know the different product categories, the keywords are declassified and thus the product categories are also declassified.[71] 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(10) gives a summary of how variables differ among categories. Additionally the table(10) shows the corresponding mean value for all features for different categories.

Let us first look at features of category 0. This category is characterized by above average bid and small number of competitors relative 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.[72] 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'.

The next category that stands out is category 2, which is characterized by high number of competition, high number of ads per search 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.[73]
Apart from this the other category that is easy to identify is category 4. 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'. Now lets try and identify the last two categories, these ones are very similar and harder to identify. Thus it is first important to analyze characteristic of the category left in the data, which are 'Cable' and 'Cruise'. 'Cruise' is a more popular search category and has more detailed search that is higher keylength. By analyzing the data it seems category 3 fits 'Cruise' and category 2 fits 'Cable'. 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 category and interpret how and why the results might differ for categories with different features.

[71] the categories are identified through specificity of keyword

[72] 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

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

Table 10: **Feature of Different Categories**

Variable	Description	Clicks	Bid	# advertisers	search
Car Insurance	High price per click & highly concentrated market : 'Car Insurance'	2.1%	4.359	20	386.24
Laptop	Popular & high competition : 'Laptop'	1.6%	.233	45	540.73
Cable	Less popular & above average price : 'Cable'	2.4%	.600	25	277.26
Cruise	Relatively popular & detailed search : 'Cruise'	1.1 %	.371	22	533.84
Coins	Low value across variables: 'Coins'	1.7%	.174	23	103.52

Showing mean value for each category
Aggregation in data: Data aggregated at day-position - keyword level
Advertisement definition id-keyword combination for the same average bid gives the advertisement
Bid Value in dollar terms: the bid is 1/100 cents, also to protect the sensitive data all bids have an unknown added amount, so we can not say much about the actual amount
Total number of observations: 207982