# ESSAYS IN MARKET DESIGN

By

Pallavi Pal

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2020

**ABSTRACT**

ESSAYS IN MARKET DESIGN

By

Pallavi Pal

**Chapter 1: Quality Differentiation and Optimal Pricing Strategy in Multi-sided Markets**

This paper analyzes the generalized quality differentiation model in multi-sided markets with positive externalities, which leads to new insights into the optimal pricing structure of the firm. We find that quality differentiation for users on one side leads to a decrease in the price charged to users on the other side, thereby affecting the pricing structure of multi-sided firms. In addition, quality differentiation affects the strategic relationships among the choice variables for the platform, so that the platform strategically uses quality differentiation to raise its profits.

**Chapter 2: Dynamic Game with Multidimensional Type: The Case of Carbon-Credit Market**

A significant problem with the carbon credit market that has become apparent in recent years is that the market price has been far more volatile than originally envisioned. The underlying problem is the ill-understood pricing anomalies in a repeated period dynamic setting. In this paper, we drive the equilibrium price path in a dynamic setting and suggests ways to overcome price instability. The model setup allows the firms to differ in terms of their value for the carbon credit as well as the urgency of obtaining it. For example, a firm with an early deadline for obtaining the carbon credits will have a higher demand urgency. We find that the equilibrium price is affected by future demand and supply expectations. The findings show that the cap or the supply limit for each period can be used to decrease price instability. Currently, the government or the carbon credit seller decides a per period

limit on the supply, which decreases over time. However, this paper suggests that to curb price fluctuation the per period supply should be a function of expected future demand. We show that correlating supply rate with expected future demand leads to a more stable price.

**Chapter 3: Revenue-Maximizing Number of Ads per Page in the Presence of Market Externalities**

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.

# ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisors Kyoo il Kim and Thomas D. Jeitschko for the continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Besides my advisors, I would like to thank the rest of my thesis committee: Hanzhe Zhang , Jay Pil Choi and Chenhui Guo, for their insightful comments and encouragement. A special thanks to Mike Conlin, Ashish Agarwal and David Croson who provided me an opportunity to discuss and give a fresh prospective to my research.

I would like to acknowledge financial support from the Department of Economics, Graduate School, College of Social Sciences, and Council of Graduate Students, along with the numerous fellowships received through the duration of these PhD years that have been instrumental in successful completion of this work. The administrative help and support received from Lori Jean Nichols and Jay Feight has made navigating this PhD program easier than it would otherwise have been.

I thank my fellow graduate students Riju Joshi, Sambojyoti Biswas, Haoyang Li, Bhavna Rai, Katie Harris, Alyssa Carlson, Udita Sanga, Katie Bollman, Akanksha Negi, Tram Hoang and Ruonan Xu for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we had over the last few years. My warmest gratitude goes to my family – Mom, Dad, Neha, Sid, and Parikshit as well as my friends back in India – Ojal, Deeksha, Devangana and Ameya for inspiring me to follow my dreams and supporting me spiritually throughout writing this thesis and my life in general. Finally, I would like to thank my best friend and life partner Dhruv Gupta, with whom I shared the joys and challenges of the PhD life and who motivated me to pursue PhD.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1

# QUALITY DIFFERENTIATION AND OPTIMAL PRICING STRATEGY IN MULTI-SIDED MARKETS†

## 1.1 Introduction

There has been an increasing shift toward multi-sided firms in many industries. The early pioneering models of multi-sided platforms were introduced by Armstrong (2006), Caillaud and Jullien (2003), Parker and Van Alstyne (2005), and Rochet and Tirole (2003). Importantly, many real-world organizations determine how close or how far they are from a multi-sided economic model based on changing industrial parameters. In this paper, we examine how platforms in multi-sided markets optimally use quality differentiation as a business strategy.

As many previous studies emphasize, it is important to establish how multi-sided platforms differ from typical one-sided firms and how such differences lead to new business implications. Many markets that have traditionally featured one-sided firms now feature more two-sided firms due to advanced technology; for example, the taxi industry only had one-sided firms before Uber appeared. Firms typically begin with an one-sided model and switch to a multi-sided model as they become more established. Doing so allows potential platforms to overcome the "chicken-and-egg" problem by first providing complementary goods themselves. For instance, Amazon started off as a pure retailer but has moved closer to a two-sided model over time by enabling third-party sellers to trade directly with consumers on its website. We can also find many other examples in which a firm faces a strategic choice of how many sides to pursue. For instance, in the personal computer market, Apple produces its own hardware, whereas Microsoft leaves this to independent manufacturers. As a result, Apple manages only a two-sided platform with consumers and software providers, while Mi-

†This work is joint with Soo Jin Kim and is unpublished.

crosoft manages a three-sided platform with consumers, software providers, and hardware providers. Indeed, there is an increasing shift from traditional business structures in which the firm focuses only on one side of users to platforms where the firm serves more than one side of users.

Given this shift toward multi-sided business, the main strategic choice that we analyze in this paper is the choice of the quality of interaction between users on two different sides. We focus on platforms that enable interaction between two sides, for example, buyers and sellers. The term "interaction" covers various traditional interactions, such as those observed in auction houses and those on internet sites in person-to-business transactions, for example, Amazon. This also includes interactions or exchanges between application developers and application users on software platforms such as computers (e.g., Apple, Microsoft); mobile devices (e.g., iPhone, Samsung); and video games (e.g., Sony PlayStation, Xbox). These markets usually have more than one type of quality access on the buyers' side. For example, Amazon offers buyers two types of quality access; basic (low-quality) access is free, while premium access (*Prime membership*, which is high-quality) is the paid service. It is easy to see that the interactions or exchanges through the high-quality access have better quality—for example, *Prime* provides two-day shipping. Many ride-sharing services offer differentiated quality tiers for customers. Uber provides riders multiple types of quality services. Although each region may have different availability, there are basically low-quality service at a cheaper price (e.g. *Uber X, Uber Pool*) as well as high-quality service at a little more expensive price (e.g. *Uber XL, Uber Select, Uber Black*). For instance, whereas trips with *Uber Pool* may take longer time to finish by sharing the same car with other strangers, *Uber Select* or *Black* is a private car sharing service with more luxurious car. Other ride-sharing services offer similar quality differentiation: Didi in China provides three quality tiers with *Express, Premier*, and *Luxe*.

These are examples of quality differentiation by firms serving two sides. Ultimately, this paper provides business implications of quality differentiation in the multi-sided platform market by building a simple model of two-sided monopolists in a market with interactions between buyers and sellers. The firm chooses the price and the quality of interaction for the buyers' side.

The first main finding of the paper is that a firm provides higher quality per dollar to buyers as it serves more sides. That is, when a firm expands its business from one-sided service (serving buyers/riders only) to multi-sided service as a platform (serving both buyers/riders and sellers/drivers), it provides better quality to buyers. Intuitively, if a firm offers a better quality price menu to buyers, it attracts more buyers. When it serves the sellers' side at the same time, more demand from the buyers' side makes sellers on the other side earn more revenue. The platform can extract those additional revenues on the sellers' side, which incentivizes the platform to offer a better quality price ratio to buyers.

We also find that the quality differentiation on the buyers' side decreases the price charged on the sellers' side compared to the price charged by a single-quality, two-sided firm. The intuition behind this finding is related to product differentiation and demand elasticity. If the platform provides multiple quality options for buyers, it means that it provides more differentiated products, which makes buyers more inelastic to price changes because the platform can extract extra surplus from high-quality buyers without deterring buyer participation. As the buyers' side becomes more inelastic, it creates an incentive for the platform to extract a higher rent from the buyers' side and subsidize the sellers' side by lowering the price faced by sellers.

By providing more quality choices, even if the quality differentiation is small, the platform is able to obtain more profits. Given that buyers are heterogeneous in their valuation of product quality, there is an incentive for the platform to offer different levels of product quality at different prices to extract more rents from buyers. Following quality differentiation, the platform needs to charge a lower price for low-quality buyers. However, it is able to earn higher extra markup from the high-valuation group of buyers (who will buy a high-quality product) if the quality gap is widened. Given that the number of high-quality buyers is sufficiently large, the platform's ultimate profit is greater with quality differentiation. Based on the model predictions, we discuss some business implications concerning how quality differentiation in one side of the market helps the platform raise its profit.

**Related literature** There is a broad literature on the corresponding problem of a monopolistic firm seeking to maximize profits by offering quality-differentiated products in one-sided standard markets. The seminal papers are Spence (1977), Mussa and Rosen (1978), Maskin and Riley (1984), and M. Itoh (1983). We generalize this problem by varying the number of sides served by the firm.

This paper is also related to the literature on pricing structure in markets with multi-sided firms (e.g., Rochet and Tirole (2003), Armstrong (2006b), and Reisinger (2010)). Aside from being broadly related to the literature on pricing in multi-sided markets, papers on skewed pricing in multi-sided markets are closely related to our paper in terms of our theoretical implications. Suarez and Cusumano (2008) discuss the platform's subsidy pricing strategy to attract greater user adoption, although they do not set up an economic model to confirm this strategy. Bolt and Tieman (2008), Schmalensee (2011), and Dou and Wu (2018) study skewed pricing strategies in two-sided markets, i.e., the subsidy and money sides. However, those papers do not consider forms of product differentiation, such as the quality differentiation examined in our paper, as a means of skewing prices.

Regarding markets with multi-sided firms, few papers have focused on firms choosing the quality of interaction on their platform. These papers study markets with negative network externalities and do not endogenize the quality choice. Crampes and Haritchabalet (2009) examine the choice of offering a pay ads regime and no pay ads packages. Peitz and Valletti (2004) compare the advertising intensity when media operators offer free services and when the subscription price is positive. Viecens (2006) is an exception because she studies a setup with endogenous quality differentiation on two-sided platforms. However, the quality differentiation in her model takes a different form from ours in that she focuses on the quality provided by users on one side and not on the quality provided by the platform itself.

Therefore, her results do not provide any implications for the platform's dynamic pricing structure, such as subsidizing one side at the expense of the other. Another quality-related aspect explored in the context of two-sided markets is the case in which users care about the quality of the other users with whom they interact. These papers are relevant for matching markets such as dating sites. Jeon et al. (2016) examine this problem in a platform setting. Renato and Pavan (2016) consider this problem in a matching setup. Hagiu (2012) studies a model in which users value the average quality of other users. The setup in these papers, however, is different from that in ours in that we focus on the quality of interaction and not on the quality of users.

Our analyses are also related to the literature on product differentiation in multi-sided markets, in that quality differentiation is one form of product differentiation. Smet and Cayseele (2010) focus on product differentiation in platform markets, which still differs from our paper in that they do not account for its consequences for optimal pricing strategies.

The literature has not focused on endogenizing the choice of network quality in multi-sided markets with positive externalities. Therefore, our results on how multi-sided markets are combined with quality differentiation on a platform provide new insights into the related business.

## 1.2 Model setup

We model the interaction between buyers and sellers. Economic value is created through the interaction between these two sides. We consider the case of a single firm providing a platform for interactions between buyers and sellers. For instance, Amazon acts as a transaction-based platform for interactions between sellers and buyers. In addition to charging an access fee to use the platform, the firm can also control the quality of interaction. Figure(A.1) displays the structure of the two-sided firm with quality differentiation on the buyers' side using Amazon as an example. The two types of access quality offered on the buyers' side are Amazon Basic and Prime—Amazon Prime is the higher-quality access, as it provides premium services including free two-day shipping.

Specifically, we model a two-sided monopolist firm that price discriminates on the buyers' side by offering two different types of access quality: $q_k \in \{q_l, q_h\}$, where $k$ denotes the quality provision, either *l*ow or *h*igh. In this model, the users' gain from the platform comes through interaction between users on two sides: buyers and sellers. Both buyers and sellers obtain utility from interacting with each other. The quality variable controls the gain from each such interaction. We assume that both buyers and sellers are heterogeneous with respect to the per-interaction or usage benefit. The usage or per-interaction benefits are $b_i^b(q_k)$ for the buyers' side (for individual buyer $i$) and $b_j^s$ for the sellers' side (for individual seller $j$), where the superscript $b$ ($s$) denotes buyers (sellers). For simplicity, we assume that $b_i^b(q_k) = B\alpha_i^b q_k$, where $B$ represents the basic benefit for every buyer. This benefit is dependent on the quality of interaction; thus, the monopolist can control the benefit by choosing the quality of platform access $q_k$ on the buyers' side. The term $\alpha_i^b$ denotes the heterogeneity among buyers; it follows a distribution function $F^b$ with support on $[0, 1]$ and density $f^b$. Additionally, each buyer pays a price for using the platform, which is denoted $p_k^b$ as a usage or per-transaction fee. For a buyer $i$, the utility function is given by the following:

$$\begin{aligned} U_{ik}^b &= [b_i^b(q_k) - p_k^b]N^s, \quad \text{where } k \in \{l, h\} \\ \Leftrightarrow U_{ik}^b &= (B\alpha_i^b q_k - p_k^b)N^s. \end{aligned} \tag{1.1}$$

The buyers' utility is the net benefit from each interaction with the other side, i.e., $(B\alpha_i^b q_k - p_k^b)$, multiplied by the number of interactions, which is denoted $N^s$.[1] From the utility specification, it can be shown that the buyer with the highest benefit is that with $\alpha_i^b = 1$, i.e., $B\alpha_i^b = B$.

[1] We assume here that every buyer interacts with every seller and that every seller interacts with every buyer. We can easily extend this to a model where the number of interactions is a function of the total number of sellers, such as $g(N^s)$. The results are robust to such extensions.

We assume that the access fee $p_k^b$ is charged per-transaction for simplicity. Indeed, certain kinds of platforms charge for premium service based on usage: for Uber, a basic service (Uber X) is cheaper than a premium service (Uber Select) for the same trip. We view that the per-usage fee for the basic service (low-quality) is zero (i.e., $p_l^b = 0$ in this example) whereas that for the premium service (high-quality) is the difference between two riding costs (i.e., $p_h^b > 0$). Nevertheless, it is worth noting that many platforms charge an one-time fixed access fee: for instance, Amazon Prime costs either nothing for Basic but \$119 per year (or \$12.99 per month) for Prime service. Note that the qualitative results hold under the model with fixed fees, as shown in the Appendix(B). Such one-time fixed fee structure can also be considered as a variation of per-transaction fee by considering the fixed fee as continuous: if Amazon Prime members make hundred transactions per year on average, the average per-transaction cost is \$1.19 (i.e., $p_l^b = 0$ and $p_h^b = \$1.19$).

The sellers' side is not affected by quality; therefore, there is no quality component. We assume that $b_j^s = S\alpha_j^s$, where $\alpha_j^s$ represents seller $j$'s heterogeneity with respect to the per-usage benefit, which is also distributed by a distribution function $F^s$ with support on [0,1] and density $f^s$. For seller $j$, the utility function is given by the following:

$$\begin{aligned} U_j^s &= (b_j^s - p^s)N^b \\ \Leftrightarrow U_j^s &= (S\alpha_j^s - p^s)N^b. \end{aligned} \tag{1.2}$$

The sellers' utility is the net benefit from each interaction with the other side, i.e., $S\alpha_j^s - p^s$, multiplied by the number of interactions, which is denoted $N^b$.[2] All sellers are charged $p^s$ per interaction, so the price discrimination is only on the buyers' side—for example, an individual seller on Amazon pays \$0.99 for each sale. The total number of interactions is $N^b N^s$.[3]

[2] In this model, we assume that the sellers' side is affected only by the total number of buyers, not the type of buyer with which a seller interacts. Thus, the total number of interactions for each seller is $N^b$, and the total number of interactions by sellers is not affected by the type of buyer with which they interact.

[3] Note that the qualitative results still hold under a generalized setup with any function of $N^b N^s$.

Next, the cost of the two-sided monopoly firm depends on the quality provided. The total cost of a transaction is given by $c(q_l) \geq 0$ for a transaction between a low-quality buyer and a seller and $c(q_h) \geq 0$ for a transaction between a high-quality buyer and a seller. In Amazon example, such cost differentiation captures that two-day shipping for Prime members is costlier than standard shipping for Basic members. We normalize the cost for sellers $c^s = 0$. The cost function is assumed to be increasing and convex in quality ($c'(q) > 0$ , $c''(q) > 0$). We analyze the nontrivial case in which $q_h > q_l$.

The demand on the buyers' and sellers' sides is represented by $D^b$ and $D^s$, respectively. In equilibrium, demand will be equal to the number of participants on each side, which means $D^b_k = N^b_k$ and$D^s = N^s$, where $k \in \{l, h\}$. Given the equilibrium demands, we turn to the monopolistic platform's problem. The monopoly platform's problem can be written as follows:

- If one type of quality is offered:

$$\max_{(p^s, p^b, q)} \Pi = [p^b + p^s - c(q)]D^b D^s. \tag{1.3}$$

- If two types of quality are offered:

$$\max_{(p^s, p^b_l, p^b_h, q_l, q_h)} \Pi = [p^b_l + p^s - c(q_l)]D^b_l D^s + [p^b_h + p^s - c(q_h)]D^b_h D^s. \tag{1.4}$$

Note that if the monopolist serves only one side of the market, say the buyers' side, its profit function is given by the following:

$$\max_{(p^b, q)} \Pi_{\text{one-sided monopolist}} = [p^b - c(q)]D^b, \tag{1.5}$$

in the case of one quality. Throughout the paper, we make the following assumptions.

**Assumption 1.** The cost is increasing and convex in quality: $c'(q) > 0$ , $c''(q) > 0$.

This suggests that it becomes increasingly costly to provide higher-quality service, which is a standard assumption.

**Assumption 2.** The distribution functions for buyers and sellers are increasing in type: $(f^b)' \geq 0$ and $(f^s)' \geq 0$, which implies that $f^b(1) > 0$ and the inverse hazard rate is non-increasing in user type: $\frac{\partial \frac{1-F^k(\theta)}{f^k(\theta)}}{\partial \theta} \leq 0, k = \{b, s\}$.

This implies that we consider the case with a positive mass of consumers. Additionally, it guarantees that as users value the per-usage benefit more, there are more platform users than nonusers.

**Assumption 3.** $\frac{\partial U^b(\alpha_i^b, p^b, q)}{\partial p^b} < 0$, $\frac{\partial U^b(\alpha_i^b, p^b, q)}{\partial q} > 0$, $\frac{\partial U^s(\alpha_i^s, p^s)}{\partial p^s} < 0$.

**Assumption 4.** The single crossing property holds: $\frac{\partial^2 U^b(\alpha_i, q(\alpha_i))}{\partial \alpha_i \partial q} > 0$.

This means that we can always distinguish high-type buyers from low-type buyers based on their $\alpha_i$.

## 1.3 Model

### 1.3.1 Model with one quality case

As a benchmark, we begin by considering a model that omits the practice of quality differentiation and then modify the model to allow for its practice in Section 1.3.2. The two-sided monopolist offers a single quality of access in this case. Buyers have the following utility:

$$U_i^b = (B\alpha_i^b q - p^b)N^s, \tag{1.6}$$

where $q \in [0, 1]$ denotes the quality index. The seller has the following utility:

$$U_j^s = (S\alpha_j^s - p^s)N^b. \tag{1.7}$$

We first analyze the user side (buyers and sellers) to identify the equilibrium demand. The equilibrium demand functions are derived from the participation constraint:

$$\begin{aligned} D^b &= Prob\left(U_i^b \geq 0\right) \Leftrightarrow D^b = 1 - F^b\left(\frac{p^b}{Bq}\right). \\ D^s &= Prob\left(U_j^s \geq 0\right) \Leftrightarrow D^s = 1 - F^s\left(\frac{p^s}{S}\right). \end{aligned} \tag{1.8}$$

The monopoly problem can be written as follows:

$$\max_{(p^s, p^b, q)} \quad \Pi = [p^b + p^s - c(q)] D^b D^s, \tag{1.9}$$

where we have normalized the marginal cost on the seller's side to zero, i.e., $c^s = 0$. We can solve for the optimal solutions for the buyers' side and sellers' side as follows:

$$\frac{p^b - c(q) + p^s}{p^b} = \frac{1}{\varepsilon^b}; \quad \frac{p^s - c(q) + p^b}{p^s} = \frac{1}{\varepsilon^s}. \tag{1.10}$$

Equation (1.10) is similar to the standard Lerner pricing formula, which states that the markup on price is equal to the inverse of the price elasticity of demand. The price elasticity of demand for buyers (or sellers) is represented by $\varepsilon^b$ (or $\varepsilon^s$). The added term in the case of a two-sided firm is the extra opportunity cost of increasing the price on the buyers' (or sellers') side, which is the marginal loss on the sellers' (or buyers') side, equal to $p^s$ (or $p^b - c(q)$).

Finally, the optimal condition for quality is given as follows:

$$c'(q) = \frac{[p^b - c(q) + p^s]}{D^b} \frac{\partial D^b}{\partial q}. \tag{1.11}$$

The optimal quality equates the marginal cost to the marginal revenue of increasing quality. The marginal revenue is the change in buyers' demand,[4] which is represented by $\frac{\partial D^b}{\partial q}$ times the per-transaction profit. From the conditions given above, the equilibrium price-quality structure is given by Theorem 1.3.1.

**Theorem 1.3.1.** *The equilibrium price and quality variables are given by the following equation:*

$$\frac{qc'(q)}{\nu^b} = \frac{p^b}{\varepsilon^b} = \frac{p^s}{\varepsilon^s}, \tag{1.12}$$

*where $\nu^b$ is the quality elasticity of demand on the buyers' side, and $\varepsilon^b$ or $\varepsilon^s$ is the price elasticity of buyer or seller side, respectively.*

[4] In the optimal quality equation, the term for the number of sellers ($D^s$) is canceled out, as it is present on both sides of the equation.

Equation (1.12) is similar to the equilibrium of two-sided market with quality decision by the platform as in Rochet and Tirole (2003). However, our equilibrium condition still differs from that in Rochet and Tirole (2003) insofar as we focus on a three-way trade-off which includes the quality variable as well, instead of a trade off between prices on the two sides. By placing a main focus on price variables only, Rochet and Tirole (2003) find that price on the buyers' side would be lower if sellers are more price elastic because the platform wants to balance the sizes of two sides. Theorem 1.3.1 further says that the higher quality the platform offers, the higher price it charges to buyers, while such upward pressure on prices diminishes as buyers becomes less quality sensitive.

Given that many platforms consider their quality of service as another important decision variable, our finding broadens the necessary considerations for the platform: it needs to set optimal prices not only by considering the price structures in terms of the ratio of price elasticities, as pointed out in Rochet and Tirole (2003), but also by taking quality elasticity into account, as we additionally find. For instance, if buyers on Amazon become more elastic to quality increase, Amazon is able to increase the buyers' side demand substantially by improving its quality at a minimal level. Accordingly, if Amazon changes its shipping policy for Prime members to three-day shipping, instead of two-day shipping, and lowers the fee charged to Prime members (i.e., $p^b$ is lower than \$119 per year, which is currently set for two-day shipping benefits), this makes Amazon better off because more Basic members will sign up for Prime members: the demand increasing effect arising from quality improvement outweighs price decreasing effect due to large quality elasticity.

Allowing the platform to consider how high or low quality its service is offered, or how quality tiers it provides are differentiated, opens more doors for optimal business strategies, insofar as quality choice gives another level of flexibility in terms of profit maximization to the platform. Indeed, as we will show, the platform's optimal pricing decisions depend on its quality choice.

### 1.3.2 Model with two qualities case

We start by analyzing the users' side (buyers and sellers) to identify the equilibrium demand. First, the buyers have two choices for accessing the platform. They can join the platform through either low-quality access or high-quality access. Given the two types of quality, high and low, the number of participants joining with low-quality access is determined by the number of buyers who satisfy the following two conditions:

1. (IR constraint) The buyers' utility from low-quality access is greater than zero: $Pr(U_l^b \geq 0)$.
2. (IC constraint) Buyers for whom the utility derived from low-quality access exceeds that from high-quality access: $Pr(U_l^b \geq U_h^b)$.

The two conditions jointly determine the proportion of low-type buyers.

$$D_l^b = \Pr\left(\frac{p_h^b - p_l^b}{B(q_h - q_l)} \geq \alpha_i^b \geq \frac{p_l^b}{Bq_l}\right) = F^b\left(\frac{p_h^b - p_l^b}{B(q_h - q_l)}\right) - F^b\left(\frac{p_l^b}{Bq_l}\right), \tag{1.13}$$

where $D_l^b \equiv D^b(p_h^b, p_l^b, q_h, q_l)$. Similarly, the number of participants joining the high-quality service is given by the number of buyers who satisfy the following two conditions:

1. (IR constraint) The buyers' utility from the high-quality good is greater than zero: $Pr(U_h^b \geq 0)$.
2. (IC constraint) Buyers for whom the utility derived from the high-quality good exceeds that from the low-quality good: $Pr(U_h^b \geq U_l^b)$.

The IR condition is satisfied when the IC constraint of the high type and IR constraint of the low type holds.[5] Thus, the proportion of high-type buyers is given by:

$$\begin{aligned} & D_h^b = \Pr(U_h^b \geq U_l^b) \\ \Leftrightarrow\ & D_h^b = \Pr\left(\alpha_i^b \geq \frac{p_h^b - p_l^b}{B(q_h - q_l)}\right) = 1 - F^b\left(\frac{p_h^b - p_l^b}{B(q_h - q_l)}\right), \end{aligned} \tag{1.14}$$

[5]We maintain the standard single crossing condition, which implies that higher types have greater willingness to pay (WTP) for quality at any price or that consumers may be ordered by their type.

where $D^b_h \equiv D^b(p^b_h, p^b_l, q_h, q_l)$. Given the utility function for buyers, the total number of buyers joining the platform is given by the following:

$$\begin{aligned} &D^b = \Pr(U^b_l \geq 0) = \Pr(b^b_l \geq p^b_l) \\ \Leftrightarrow\ &D^b = \Pr(B\alpha^b_i q_l \geq p^b_l) = \Pr(\alpha^b_i \geq \frac{p^b_l}{Bq_l}) = 1 - F^b\left(\frac{p^b_l}{Bq_l}\right), \end{aligned} \tag{1.15}$$

where $D^b \equiv D(p^b_l, q_l)$. Equation (1.15) shows us how the number of participants on the buyers' side depends only on the price and quality of the low-quality good. Although there are network externalities in the total utility derived from the platform or the gross transaction utility, the per-unit transaction demand is not dependent on the participation rate on the other side.[6] This is because the participation constraint (IR constraint) for the high-quality buyers is slack. This means that the participation of low-quality buyers guarantees the participation of high-type buyers. In other words, the buyers on the margin of joining the platform are low-quality buyers.

Next, the total number of sellers who join the platform is given by the following:

$$D^s = \Pr(U^s \geq 0) = 1 - F^S\left(\frac{p^s}{S}\right), \tag{1.16}$$

where $D^s \equiv D^s(p^s)$. Given the total number of buyers and sellers, the equilibrium level of participation is the following:

$$\begin{aligned} D^b &= D(p^b_l, q_l) = 1 - F^b\left(\frac{p^b_l}{Bq_l}\right) \\ D^b_l &= F^b\left(\frac{p^b_h - p^b_l}{B(q_h - q_l)}\right) - F^b\left(\frac{p^b_l}{Bq_l}\right), \\ D^b_h &= D^b(p^b_h, p^b_l, q_h, q_l) = 1 - F^b\left(\frac{p^b_h - p^b_l}{B(q_h - q_l)}\right) \\ D^s &= D^s(p^s) = 1 - F^s\left(\frac{p^s}{S}\right). \end{aligned} \tag{1.17}$$

Given quality differentiation, the monopoly problem can be written as follows:

$$\max_{(p^s, p^b_l, p^b_h, q_l, q_h)} \quad \Pi = [p^b_l + p^s - c(q_l)]D^b_l D^s + [p^b_h + p^s - c(q_h)]D^b_h D^s. \tag{1.18}$$

[6]This setup has one restriction that we need to impose, which is that the proportion of low-type buyers has to be nonnegative: $D^b_l \geq 0$.

The following is the breakdown of the equilibrium prices and quality for buyers and sellers:[7]

#### 1.3.2.1 Price of low-quality access on the buyer side

We first examine the price of low-quality access on the buyer side as follows.

$$\frac{p_l^b - c(q_l) + p^s}{p_l^b} = \frac{1}{\varepsilon^b}, \text{ after using equilibrium value of } p_h^b, \tag{1.19}$$

where the price elasticity of demand for low-quality access is represented by $\varepsilon^b$. Note that the equilibrium conditions for the price elasticity with respect to the lower price is denoted as those for the full demand: that is why we have $\varepsilon^b$, not $\varepsilon_l^b$.

#### 1.3.2.2 Price of high-quality access on the buyers' side

The price of high-quality access on the buyers' side can be obtained as follows:

$$p_h^b = p_l^b + \underbrace{[c(q_h) - c(q_l)] + \frac{p_h^b}{\varepsilon_h^b}}_{\text{additional cost plus extra market power}}, \tag{1.20}$$

where the price elasticity of demand for high-quality access is represented by $\varepsilon_h^b$. The optimal price for high-quality access is equal to the price for low-quality access and the additional cost, i.e., $c(q_h) - c(q_l)$, plus an additional markup, i.e., $\frac{p_h^b}{\varepsilon_h^b}$.

#### 1.3.2.3 Price for sellers

We now turn our attention to the price for sellers.

$$p_l^b + p^s - c(q_l) = \frac{p^s}{\varepsilon^s} - \frac{p_h^b}{\varepsilon_h^b}\frac{D_h^b}{D^b}. \tag{1.21}$$

Again, Equation (1.21) is similar to the standard Lerner pricing formula, which states that the markup on price is equal to the inverse of the price elasticity of demand. The added term in the case of a two-sided firm is the extra opportunity cost of increasing the price on

[7] The details are in the Appendix.

the sellers' side, which is the marginal loss on the buyers' side, and is equal to the average per-interaction profit on the buyers' side.[8]

#### 1.3.2.4 Low-quality service for buyers

Given prices, the monopolistic platform solves the profit maximization problem separately for the low- and high-quality services for buyers. First, for the low-quality service, the first order condition can be derived as follows:

$$p_l^b + p^s - c(q_l) = \frac{q_l}{\nu^b} \frac{D_l^b c'(q_l) + D_h^b c'(q_h)}{D^b}, \tag{1.22}$$

where $\nu^b$ denotes the quality elasticity of demand for low-quality access. The cost of low quality service equates the marginal cost to the marginal revenue of increasing quality.

#### 1.3.2.5 High-quality service for buyers

Here, the first order condition for the high-quality service for buyers is given by the following:

$$c'(q_h) \frac{q_h}{\nu_h^b} = \frac{p_h^b}{\varepsilon_h^b}, \tag{1.23}$$

where $\nu_h^b$ denotes the quality elasticity of demand for high-quality access. The marginal cost $c'(q_h)$ should be equal to the marginal revenue, which is the product of increased high-quality buyers, i.e., $\frac{\nu_h^b}{q_h}$, and the extra markup generated from the increase in high-quality buyers, i.e., $\frac{p_h^b}{\varepsilon_h^b}$.

## 1.4 Equilibrium results

We derive several important implications from the model. Before we examine the diverse effects of quality differentiation on the platform's strategies, we first address that the monopolistic firm provides higher quality to buyers for each dollar that they pay when it serves more sides of the market. As the monopolist opens more sides to serve, say from a one-sided

[8] Note that we have normalized the cost on the seller's side to zero, so $c^s = 0$.

firm serving buyers only to a two-sided firm as a platform serving both buyers and sellers, such openness increases the quality per dollar offered to buyers. Mathematically, the quality per dollar is denoted as $\frac{q}{p^b}$. This finding is summarized in Proposition 1.4.1.

**Proposition 1.4.1.** *As a firm serves more sides of the market, it provides higher quality per dollar offered to buyers:* $\frac{q}{p^b}^{\textit{One-sided firm}} \leq \frac{q}{p^b}^{\textit{Two-sided firm}}$.

Intuitively, the monopolist serving multiple sides has more incentive to offer a better quality price ratio to buyers because it now obtains more profit from the seller's side. If the two-sided firm offers a better quality price menu to buyers, it attracts more buyers. When it serves the sellers' side at the same time, more demand from the buyers' side makes sellers on the other side earn more revenue. The platform can extract those additional revenues on the sellers' side, which incentivizes the platform to offer a better quality price ratio to buyers.

In other words, the openness of the platforms from serving one side of users only to serving both sides increases the quality per dollar, which ultimately attracts more buyers than a one-sided firm. Both traditional one-sided firm and two-sided platforms offer differentiated quality tiers for buyers. When those two types of firms engage in quality differentiation on buyers' side, we show that the platform is more willing to offer a better deal to low type members by giving a high quality per dollar in the basic level than a one-sided firm. For example, the average quality of services provided by taxi companies, as traditional one-sided firms, is known to be lower than that provided by ride sharing platforms, such as Uber and Lyft. As Liu et al. (2019) shows, taxi drivers are more likely to detour with non-local customers, which results in longer travel time. Such empirical evidence supports our finding that the platform's basic quality provision is better than a one-sided firm's. The reason the platform provides a higher quality for basic service is because it is able to exploit such quality improvement in one side to encourage more participation from the other side, thereby raising its profits. As in the ride sharing platforms example, when such platforms provide higher basic quality for riders, which attracts more riders, it ultimately gives a stronger incentive for drivers to join the platforms, thereby permitting them to extract more rents from drivers: the underlying incentives for the platforms to provide higher service provision than a one-sided firm arise from this cross-subsidization motive. This finding is summarized in Corollary 1.4.2.

**Corollary 1.4.2.** *The platform in two- or multi-sided market is more likely to offer higher quality of basic service or product than a one-sided firm.*

Now, we will show how the quality provision decided by the platform affects its optimal use of the interaction between two sides, such as cross-subsidization, in the following propositions below. Returning to the multi-sided platform case with two qualities provision, we first find that quality differentiation reduces the other side's price level. The Appendix contains a proof of this result.

**Proposition 1.4.3.** *Quality differentiation on the buyers' side decreases the price charged on the sellers' side relative to the price charged by the platform that offers one quality.*

The effect of offering low- and high-quality access to the buyers makes the buyers more price inelastic because the platform can extract additional surplus from high-quality buyers without deterring buyers' participation. Now, recall that for multi-sided firms, the optimal pricing scheme is to subsidize the more elastic side of the market and extract rents from the other, more inelastic side. As the buyers' side becomes more inelastic, it creates an incentive for the platform to extract higher rent from the buyers' side and subsidize the sellers' side by lowering the price charged to sellers. This can be explained by comparing the equilibrium pricing structures of the two cases:

$$\begin{aligned} &\textit{one-quality case:} \quad \frac{p^s}{\varepsilon^s} = \frac{p^b}{\varepsilon^b} = \frac{qc'(q_l)}{\nu^b}. \\ &\textit{two-quality case:} \quad \frac{p^s}{\varepsilon^s} = \frac{p^b_l}{\varepsilon^b} + \frac{D^b_h}{D^b}\frac{p^b_h}{\varepsilon^b_h} = \frac{q_l}{\nu^b}\frac{D^b_l c'(q_l) + D^b_h c'(q_h)}{D^b} + \frac{q_h}{\nu^b_h}\frac{D^b_h c'(q_h)}{D^b} \quad . \end{aligned} \tag{1.24}$$

Similar to the single-quality case, the platform faces a trade-off of whether to charge a higher price on the buyers' or sellers' side in cases with two quality types. After quality differentiation, the trade-off features an additional markup benefit on the buyers' side, namely, the extra margin from high-quality buyers. Thus, the monopolist becomes more efficient in extracting rent from the buyers' side. Given this, the platform has a higher incentive to increase the size of seller demand by lowering the price on the sellers' side.

We also find that quality differentiation leads to greater profit for the platform, as in Corollary 1.4.4. The detailed proof is in the Appendix.

**Corollary 1.4.4.** *The platform strictly prefers to price discriminate by quality on the buyers' side*

Corollary 1.4.4 implies that a platform that provides only one type of quality is able to obtain more profit if it slightly differentiates product quality. Even a minor quality improvement with a small price increase can increase the platform's profit as long as it continues to provide differentiated products, such as low- and high-quality products. Thus, quality differentiation permits the platform to earn more profit by implementing premium service in addition to basic service, which not only expands the total buyer market size, but also extracts more rents from relatively high-quality type of buyers.

As for the the platform's optimal quality provision, we find that the high quality provision is independent of the sellers' side equilibrium whereas the low quality provision is related to the sellers' side equilibrium. By comparing Equations (1.22) and (1.23), we derive the following proposition.

**Proposition 1.4.5.** *The platform's optimal quality level for the high-quality service does not depend on the sellers' side equilibrium, whereas that for low-quality service increases in the marginal revenue from the sellers' side.*

Proposition 1.4.5 states that the platform has an incentive to increase the quality level of the basic service (low-quality) because of the positive network effect coming from the sellers' side, whereas it does not consider the network effect when determining the optimal quality level of the premium (high-quality) service. The intuition behind this finding relates to total buyer demand. As in Equation (1.15), the participation constraint for high-quality buyers is slack, which means that the effect of increasing the quality level of the premium service does not increase the total buyer demand, although it increases the share of high-quality buyers. The platform will be able to increase the total buyer demand, $D^b$, only by attracting more low-quality buyers, and one way of doing so is to increase the quality level of the basic (low-quality) service. More total buyer demand arising from increasing $q_l$ boosts the marginal revenue from the sellers' side due to the positive network effect on the platform. Thus, the optimal condition for $q_l$ is a function of the marginal revenue on the sellers' side.

Moreover, we find that the additional markup from quality differentiation is increasing in the quality gap. Thus, if the platform sufficiently differentiates its product line with respect to quality, it can earn a higher markup. Proposition 1.4.6 summarizes this finding.

**Proposition 1.4.6.** *The platform's incentive to raise the fee charged to high-quality buyers is increasing in the difference in quality, as shown in* $\frac{\partial\left(p_h^b/\varepsilon_h^b\right)}{\partial(q_h-q_l)} > 0$.

Proposition 1.4.6 implies an interesting result—the platform is more likely to charge a higher fee to high-quality buyers if it either provides much better service (e.g., one-day rather than two-day shipping for Amazon Prime members) or maintains the high-quality service at the same level while reducing the quality of the basic service (e.g., increasing the minimum order qualified for free shipping for Amazon Basic members). Section 1.5.2 discusses this point in depth.

In addition, we examine the strategic relationships between the platform's choice variables to derive further implications. We first examine whether the platform's decision variables are strategic complements or substitutes in the game of the one-quality case.

**Proposition 1.4.7.** *The prices on the buyers' side ($p^b$) and the sellers' side ($p^s$) are strategic substitutes. Whether the price variable (buyer or seller side) and quality variable are strategic substitutes or complements is ambiguous.*

From Proposition 1.4.7, reducing $p^b$ is the profit-maximizing response to increasing $p^s$, and vice versa. This means that the platform does not maximize its profits if it increases both prices. Unlike this case, the strategic relationship between price and quality variables is ambiguous. The price for the buyer and quality are strategic substitutes if $\frac{\partial^2\Pi}{\partial p^b \partial q} = \frac{p^b}{Bq^2}\{D^s f^b(\frac{p^b}{Bq}) - \frac{1}{Bq^2}[p^b + p^s - c(q)]f'(\frac{p^b}{Bq})\} + c'(q)\frac{1}{Bq}f^b(\frac{p^b}{Bq}) < 0$. Similarly, the price for the seller and quality are strategic substitutes if $\frac{\partial^2\Pi}{\partial p^s \partial q} = D^s \frac{p^b}{Bq^2} f^b(\frac{p^b}{Bq}) - c'(q)\frac{1}{S^2} f^s(\frac{p^s}{S}) < 0$. If $F^b$ and $F^s$ both have a uniform distribution, $\frac{\partial^2\Pi}{\partial p^b \partial q}$ is always positive (because $f' = 0$), whereas $\frac{\partial^2\Pi}{\partial p^s \partial q}$ is negative if $c'(q) > S(S - p^s)\frac{p^b}{Bq^2}$. This parametric example suggests that if the platform increases quality, it charges buyers a higher price but sellers a lower price (provided that the marginal cost of increasing quality is above a certain threshold).

Next, we derive results from the two-qualities case. For simplicity, we assume that $F^b$ and $F^s$ both have a uniform distribution, which implies that the profit in the two-qualities case is given as follows.

$$\Pi = \left(1 - \frac{p^s}{S}\right)\left\{\left(p_h^b - q_h^2 + p^s\right)\left[\frac{p_l^b - p_h^b}{B\left(q_h - q_l\right)} + 1\right] + \left(1 - \frac{p_l^b}{Bq_l}\right)\left(p_l^b - q_l^2 + p^s\right)\right\}. \tag{1.25}$$

By the first order conditions derived in 1.3.2, we find the following.

$$\begin{aligned}
\frac{\partial^2\Pi}{\partial p_l^b \partial p_h^b} &= \frac{q_l}{2(q_h - q_l)}; \quad \frac{\partial^2\Pi}{\partial p_l^b \partial q_h} = -\frac{q_l\left[p_h^b + q_h\left(q_h - 2q_l\right) + p^s\right]}{2\left(q_h - q_l\right)^2}.\\
\frac{\partial^2\Pi}{\partial p_l^b \partial q_l} &= \frac{1}{2}\left[B + \frac{q_h\left(p_h^b - q_h^2 + p^s\right)}{\left(q_h - q_l\right)^2} + 2q_l\right].\\
\frac{\partial^2\Pi}{\partial p_l^b \partial p^s} &= \frac{q_h}{2(q_h - q_l)} - 1; \quad \frac{\partial^2\Pi}{\partial p_h^b \partial q_h} = \frac{B}{2} + q_h.\\
\frac{\partial^2\Pi}{\partial p_h^b \partial q_l} &= -\frac{B}{2}; \quad \frac{\partial^2\Pi}{\partial p_h^b \partial p^s} = -\frac{1}{2}.
\end{aligned} \tag{1.26}$$

From Equation (1.26), we identify whether two choice variables are strategic substitutes or complements for the platform. The following proposition summarizes the findings.

**Proposition 1.4.8.** *The price for high-quality access for buyers ($p_h^b$) is a strategic complement for the price for low-quality access for buyers ($p_l^b$) and high-quality service ($q_h$), whereas* *it is a strategic substitute for the price for sellers ($p^s$) and low-quality service ($q_l$). Whether the price for low-quality access for buyers is a strategic substitute or complement for quality measures and the price for the seller is ambiguous.*

Note that if the high-quality service is much better than the low-quality service,[9] $p_l^b$ is a strategic substitute for $q_h$ but a strategic complement for $p^s$. This implies that if the platform improves the high-quality service ($q_h$ increases), the profit-maximizing response is to increase the price for high-quality access for buyers while decreasing that for low-quality access. In addition, the best response to an increase in the price for high-quality access for buyers is to decrease the price for sellers. However, if the platform increases the price for low-quality access for buyers, the optimal response is to increase the price for sellers provided that the quality difference is large enough.

Finally, we also conduct comparative statistics to determine how changes in the exogenous parameters affect the equilibrium outcomes. In particular, we are interested in how $B$, which represents the basic benefit from the quality dimension for every buyer, affects consumer demand in the two qualities case. Proposition 1.4.9 summarizes the result.

**Proposition 1.4.9.** *Consumer demand for the high-quality service always increases in the basic benefit from better quality ($B$). Whether consumer demand for low-quality service increases in $B$ is ambiguous.*

In other words, $\frac{\partial D_h^b}{\partial B}$ is always positive (where $D_h^b$ is given by Equation (1.14)), whereas $\frac{\partial D_l^b}{\partial B}$ is positive only if a certain condition is met (where $D_l^b$ is given by Equation (1.13)). Specifically, $\frac{\partial D_l^b}{\partial B}$ is positive if $\frac{p_l^b}{q_l} f^b \left(\frac{p_l^b}{Bq_l}\right) > \frac{p_h^b - p_l^b}{q_h - q_l} f^b \left(\frac{p_h^b - p_l^b}{B(q_h - q_l)}\right)$ and negative otherwise. That is, when the basic benefit from the quality dimension increases, it can reduce the buyers' demand for low-quality access if the price-quality ratio for the quality difference (between high- and low-quality) is greater than that for low-quality service. As in Equation

[9] We can guarantee that $\frac{\partial^2 \Pi}{\partial p_l^b \partial q_h} < 0$ and $\frac{\partial^2 \Pi}{\partial p_l^b \partial p^s} > 0$ if $q_h > 2q_l$.

(1.13), buyers in the middle range of willingness to pay (i.e., who are willing to pay for low-quality service but not for higher priced high-quality service) demand low-quality service access. As $B$ increases, we observe two simultaneous outcomes: (1) more buyers who were not in the market join the low-quality service (measured by $\frac{p_l^b}{q_l} f^b \left( \frac{p_l^b}{Bq_l} \right)$), and (2) more buyers who used to use the low-quality service switch to high-quality service (measured by $\frac{p_h^b - p_l^b}{q_h - q_l} f^b \left( \frac{p_h^b - p_l^b}{B(q_h - q_l)} \right)$). If the latter effect is larger than the first, a greater $B$ leads to fewer buyers for the low-quality service. Per Propositions 1.4.7, 1.4.8, and 1.4.9, we can see that considering quality differentiation for the platform affects its strategic choices on the price and demand structure. If the platform overlooks such dynamic relationship between quality choice and optimal pricing, it could lose a potential opportunity for better business strategies using three-way interaction among quality, buyers', and sellers' side.

## 1.5 Discussion

Quality differentiation is one example of product differentiation, which makes buyers' demand less elastic. In other words, the platform can strategically use quality differentiation to maximize its profit. In Section 1.5, we discuss several business implications based on our theoretical predictions.

### 1.5.1 Quality differentiation and optimal pricing strategy

As in the model, if the platform provides different quality choices to buyers, it faces more inelastic demand from them, which allows the platform to charge a lower price to users on the other side, namely sellers. Specifically, the platform can extract a higher margin from buyers by providing multiple different qualities, which makes their demand inelastic. Given more inelastic demand from buyers, the platform finds it optimal to increase the number of sellers, as this increases the utility of buyers from interactions through the platform. The platform can increase the number of sellers by decreasing the price on the sellers' side, which is one way of subsidizing sellers.

There are several instances in which the platform might find it profitable to subsidize sellers by exploiting buyers' inelastic demand (arising from quality differentiation). One such example is a ride sharing service that connects riders and drivers such as Uber. The model findings predict that if Uber provides two or multiple different quality of services, for example a premium service with a higher usage fee and a basic service with a lower usage fee, which makes consumers' demand inelastic (due to quality differentiation), the platform (Uber) is able to maximize its profit by charging a lower driver's pay rate (a price charged to drivers). In particular, if there is competitive pressure in attracting drivers, the platform can strategically exploit the quality differentiation on the riders' side to subsidize the drivers' side, which helps it overcome intense competition on the drivers' side. Suppose that the competition for drivers is so intense, after Lyft and other competitors start operating, that it is difficult for the company to have sufficient number of drivers during the peak time. The platform wants to attract more drivers because widespread availability of drivers is crucial for expanding its business. Here, if the Uber charges much lower fees for drivers, more drivers will be willing to work for the platform. Our theory predicts that the platform will offer lower fees in equilibrium by differentiating on product quality on the riders' side, which leads to a higher driver participation rate.

### 1.5.2 Quality differentiation and extra markup

The model predicts that the platform will earn more extra markup if it widens the quality gap. This suggests that the platform can adopt either of two strategies: it can improve the high-quality product while maintaining the low-quality product at the same level, or it can reduce the quality of the basic product while maintaining the high-quality product at the same level. Both strategies lead to more quality differentiation, which results in more extra markup being charged to the high-quality buyers per Proposition 1.4.6.

Most incumbent platforms can use their pre-existing resources to develop a much higher quality product, thereby generating more quality differentiation (higher $q_h$). However, because entrants or small platforms lack substantial resources, they might not be able to make large investments to produce better-quality products. What those platforms can do is to provide a very basic quality product at a lower price (or zero price) and a slightly higher quality product at a higher price (lower $q_l$). In other words, by lowering the basic product quality at an almost zero price (or free of charge), the platforms can enjoy more extra markup even without having very high product quality. For example, there are two ways for Amazon to widen the quality gap. It could make its high-quality service more attractive by providing one-day shipping for Prime members. Alternatively, it could offer more differentiated services by maintaining the two-day shipping policy for Prime members while reducing some of the benefits for basic members, for example, increasing the order minimum to qualify for free shipping from $25 to $30.

## 1.6 Conclusion

This paper analyzes a generalized version of quality differentiation by a monopolist in a multi-sided market. The main focus of this paper's analysis is the effect of buyers' side quality differentiation on the optimal pricing strategy for the platform. We first showed that quality differentiation on the buyers' side will decrease the price charged to sellers. To understand the intuition behind this result, recall that the pricing structure for multi-sided firms depends on the relative elasticity of the two sides. Quality differentiation on the buyers' side can increase the surplus extracted from buyers, which means that the buyers' side becomes relatively inelastic. The platform can exploit this inelastic demand among buyers to extract more profit from that side while discounting the sellers' side by decreasing the price charged to them. This finding suggests how a platform can subsidize the more elastic side by introducing quality differentiation on the other side.

We also found that quality differentiation, which leads to a lower price on the sellers' side, ultimately increases the platform's profit. Thus, the platform can strategically use quality differentiation to raise its profits. Another strategic variable that the platform can use is the extent to which the two qualities differ—the platform can earn extra markup from relatively high-valuation buyers by widening the quality gap. Given that the driving force of this greater markup is the quality gap, rather than how high the quality of the better-quality product should be, we derive relevant business implications, especially for small platforms or entrants without substantial resources: if they are unable to make a higher quality product due to their limited resources, they can lower their basic product quality, which would lead to similar consequences in terms of the quality gap.

Overall, our findings suggest one plausible business strategy for the platform: how quality differentiation implemented by the platform can be used as an optimal business strategy. It would be interesting for future research to investigate how competition in the platform market alters our results. In this regard, a model with an asymmetric setup in the competitive market structure faced by the platform could determine the extent to which such asymmetry affects the platform's optimal business strategy.

## CHAPTER 2

## DYNAMIC GAME WITH MULTIDIMENSIONAL TYPE: THE CASE OF CARBON-CREDIT MARKET[†]

### 2.1 Introduction

One of the most popular market-based solutions for limiting greenhouse gas emissions (GHG) is to use Cap-and-Trade schemes. Under the scheme, firms use emission credits to pay for GHG emissions. The scheme's governing body begins by setting a cap on allowable emissions. It then distributes or auctions off emissions allowances that total the cap. The firms that do not have enough allowances to cover their emissions must either make reductions or buy another firm's spare credits. Members with extra allowances can sell them or bank them for future use.

A successful cap-and-trade scheme relies on a strict but feasible cap that decreases emissions over time. If the cap is set too high, an excess of emissions will enter the atmosphere and the scheme will have no effect on the environment. A high cap can also drive down the value of allowances, causing losses in firms that have reduced their emissions and banked credits. If the cap is set too low, allowances are scarce and overpriced. The previous literature has looked at the optimal rate of decrease in supply for the carbon credit market by accommodating the effect on carbon credits being optimally valued.

In this paper, we extend this analysis and look at the optimal per period supply when there is an uncertainty of future demand in the market. We find that in the dynamic setting introduction of demand uncertainty changes the optimal supply rate. In particular, we find that the expectation about future supply and demand changes the equilibrium price and contributes to price fluctuation in the market.

Several papers have shown evidence of price uncertainty in the carbon permits market.

---

[†]This work is joint with Thomas D.Jeitschko and is unpublished.

The demand of carbon permits is uncertain and fluctuates due to multiple reasons such as weather conditions, high correlation with electric power generation, financial trading in carbon permits. For example, in the EU emissions trading system (EU ETS) [1] a dry summer in July 2005 in Southwestern Europe led to under-utilization of hydroelectric plants and a lack of cooling water for nuclear power plants, which therefore increased the demand for carbon permits. Thus, it is integral to accommodate the uncertainty in the demand in the model so as to make it more realistic.

The results show that a decrease in future supply increases the current price. This is because increasing the future supply makes it easier for the firm to buy the carbon credit in the future which makes them less willing to pay a high price in the current period. On the other hand, we find that an increase in expected future demand leads to a decrease in the current price. The intuition behind a negative effect of an increase in expected future demand on the current price is that shift in future demand curve makes the firm demand more in the current period as the carbon credits can be stored and used in future periods. The outward shift of the demand curve consequently increases the equilibrium price. Using these results the paper suggests that the future supply should be a function of expected future demand. In particular, the government should decrease the supply every time the expected future demand decreases and visa-versa. We show that such a policy can decrease the price fluctuation in the market.

The cap and trade scheme uses uniform price auction to distribute the carbon credits. The allowance is auctioned over multiple periods. To study this market, we look at a dynamic repeated uniform price auction setting with stochastic supply and demand. The dynamic setting in the single good case can be of two kind. The first keeps the set of buyers fixed and have their types change over a time as a function of allocations selected in earlier periods (for example Athey and Segal (2007), Eso and Szentes (2007) and Akan et al. (2009). The second is where a finite number of goods are sold to buyers arriving over time. This paper looks

[1] as noted by Thomson Rueters Point Carbon (2006)

at the second type of dynamic setting which we refer to as the changing buyer case[2]. The focus here is on the sale of greenhouse gas emission, as used in the European Union, the US (RGGI, NOx, California Cap and Trade), Switzerland and other countries. The auctioned goods are close substitutes (assumed to be homogeneous) with the bidder having different degrees of urgency to get the permits. This is modeled in the paper as bidders having a different number of periods of active demand. To accommodate the different degrees of urgency among the buyers, the model allows the buyer's type to differ in terms of their demand duration. Thus, buyers have multi-dimensional private information, namely about their valuation and demand duration. The buyers have single-unit demand [3]

In this setting, we derive the equilibrium bid and show that the bid is truthful and ex-post incentive-compatible. The optimal bidding strategy indicates that the buyers take their future period payoffs into consideration when deciding the value of winning in the current period. Thus, the outside option value is endogenously determined through the expected payoff of auctions in future periods. Apart from the carbon credit auction, this model can be applied to any market where the bidders differ in terms of the value of the object as well as the urgency for getting good. This includes markets such as the sale of spectrum (used for a wireless network), auction of house or electricity markets with different delivery dates.

This paper is related to the literature on price and supply restriction imposed in the cap and trade auction. The early work on this was done by Robert & Spence (1976) and Weitzman (1978) . They considered price ceiling and floor under demand and supply uncertainty in the static model. In the dynamic setting papers such as Philibert (2008) and Burtraw et al (2010) look at price control with quantity regulation. Webster et al. (2010) compare price ceiling to a policy that allocates allowance on emission targets. This paper

[2]The term changing buyer type is taken from Dynamic Mechanism design literature review by Rakesh V. Vohra

[3]The objective of this paper is to focus on the question of auction periodicity, hence we abstracting away from multi-unit demand. Another way of thinking of this is the simplification of each bidder having the homogenous number of units demands, thus the item is sold is as a bundle of a specific amount.

contributes to this literature by looking at how instead of a fixed supply rate or price ceiling, we can implement a dynamic control for price fluctuation by changing the supply rate with the change in demand uncertainty and thereby affecting the equilibrium price.

This paper is also related to the literature on dynamic mechanism design with multidimensional private information. It extends this literature to look at the optimal auction for an auctioneer with an unknown number of buyers and sellers in each period. The literature on dynamic mechanism design is vast(see the survey by Bergemann and Said for details). The most relevant papers for current work are Pai and Vohra (2013), where they look at the optimal auction for a single seller selling multiple units to stochastically arriving bidders. These authors also allow for private information about the arrival time. The main difference in this paper is that we are looking at auctioneer's problem of efficient mechanism with stochastically arriving sellers. Also Pai and Vohra paper assumes perfectly patient bidders with a focus on optimal auction for maximizing seller revenue. Another related paper is Mierendorff, K., (2014):, where the author looks at a revenue-maximizing mechanism for a seller selling a single good in a dynamic environment with buyers having multidimensional private information. The main difference with current work is that we are looking at an efficient mechanism in a dynamic environment with multiple and stochastically arriving sellers. This paper is also related to the literature on Efficient Sequential Auction with Impatient Buyers. The most relevant paper is Gershkov and Moldovanu (2010). They examine the allocation of a set of durable goods to the dynamic buyer population. In their setting, objects are durable whereas in this paper objects are non-durable and the total supply in every period is stochastic.

In the current model, we are analyzing sequential auctions with stochastic supply and demand. The early work on stochastic auctions includes Mcafee, R. P., and McMillan, J. (1987), Konrad, M.(2013), looks at an efficient dynamic allocation of a single object when bidder arrival is stochastic. On the supply side, Thomas.J (1999) looks at stochastic supply in a sequential auction.

The remainder of the paper is structured as follows : Section(2.2) and section(2.3) describe the generalized model setup and derive the equilibrium bid. Section(2.4) applies the model to emission trade market and section(2.5) concludes.

## 2.2 Model setup

Consider a sequential auction in infinite, discrete time period model, $t \epsilon \{-\infty.., 0, ..\infty\}$. In each period multiple units of a homogeneous good are auctioned. In the case of a spectrum market, we look at the trading auction where the secondary users bid to buy spectrum units, sold by primary users. Buyers with single unit demand arrive over time. Buyer $i$ has a valuation $v_i$, for one unit of the good, which is an i.i.d random draw from the distribution $F(v)$ on $[\underline{v}, \bar{v}]$. The demand for each buyer lasts for multiple periods. In particular, each buyer has an arrival time $a_i$ and a demand duration $k_i$. This means that the demand lasts for all $t \epsilon \{a_i, ....(a_i + k_i)\}$. Thus apart from heterogeneous valuation, a buyer also differ in terms of their demand lifetime. The type of a buyer is a triplet consisting of his valuation, arrival time and demand duration, $x = (v_i, a_i, k_i)$; and the type space is given as $X = [\underline{v}, \bar{v}] \times [\underline{a}, \bar{a}] \times [1, \bar{k}]$. Buyer's type is an i.i.d random draw from a commonly known distribution $\prod_{i=\{v,a,k\}} F_i = F_v \times F_a \times F_k$ over $X$. We assume that the three components of the type space, i.e. valuation, arrival time and demand duration are independent.

In our model, in period $t$, an agent of type $x_i = (v_i, a_i, k_i)$ and makes a payment $z_t$ derives the following instantaneous utility:

$$U_{i,t}(v_i, k_i) = \begin{cases} (v_i - z_t), & \text{if he wins the auction} \\ 0, & \text{otherwise} \end{cases}$$

The buyers arrival rate is stochastic; In any period $t$, $n_t$ new buyers arrive. The arrival rate, $n_t$, is an i.i.d random draw from the distribution $F^n$ on $\{1, ...., \bar{n}\}$. The distribution of arrival rate of buyers in each period is common knowledge, however the exact number of per period new buyers is unknown. Apart from this uncertainty in the market leads to uncertainty of the current bidders' likelihood of surviving in the future. To capture this, we define $\tau$ as the probability of surviving in the next period for each active buyer. The total number of potential buyers is given by $N \epsilon IN$ such that $N \geq \bar{n}\bar{k}$.

Sellers, with single period supply, arrive over time. Seller $j$ has a per unit value $v^j$ , which is assumed to be the same for all sellers and is normalized to zero, $v^j = 0 \; \forall \; j \; \epsilon \; S$[4]. We allow the sellers to have more than one unit to sell but assume that each sale request is sent to the auctioneer independently[5].

Similar to the buyers, the sellers also arrive stochastically. However unlike the buyers, sellers are active only for one period. The units auctioned by the sellers are identical, and these units remain valid only for one period. Each unit is considered for sale independent of their provider, as sellers offer homogeneous goods. In any period t, $m_t$ new units are available in the market. It is assumed that for each $t^{th}$ period, $m_t$ is an i.i.d random draw from distribution $F^m$ on $[\underline{m}, \bar{m}]$. The total number of potential seller is given by $M \epsilon IN$ such that $M \geq \bar{m}$. The distribution of units available in each period and seller's valuation is common knowledge however the exact per period supply is unknown ex ante. [6].

*Information Structure:* The distribution of buyer's type space, the distribution of available per period units, and seller's valuation for the object are assumed to be common knowledge. The buyer has private knowledge about his own type, composed of his valuation and demand lifetime. He does not have knowledge of exact number of other buyers and sellers, their types and reports , or previous allocation decision by the auction. The arrival rate of buyers and the number of items in each period are stochastic. Thus the exact per period demand and supply is unknown for future periods.

The timeline of the game is as follows. In each period, $t$, the active buyers report a bid for a single unit to the auctioneer. The strategy set of the buyers in period $t$ compromises of

[4]This assumption has been made so that we can concentrate on the buyers side in this paper

[5]Thus this is equivalent to each seller selling single unit item

[6]In this model we consider the case of multiple sellers. The role of a seller is limited upto to reporting the available unit, and thereby accepting to sell at a non negative price as the auctioneer knows seller's valuation.

their bid in period $t$, i.e. $b_t(v_i, r_i)$, where $v_i$ is the value and $r_i$ is the numebr of active periods left at $t^th$ period. The active buyers comprises of new buyers arriving in the current period and the existing buyers which are still active and have unfulfilled demand for the item. The active sellers report the items available for sale. After receiving bids from the active buyers and an estimate of number of items to be sold from the active sellers, the auctioneer holds an auction to decide the number of items to be traded, denoted as $S_t$ and the clearing price, denoted as $z_t$.

We now give a detail description of the auction mechanism. The auctioneer holds a series of static uniform price auctions. Bidders simultaneously submit sealed bids for the item and sellers report the units available for sale. The auctioneer calculates the total number of items to be traded, $S_t$, by equating the demand and supply in the current period. If the number of items traded in the market are equal to the available units i.e. $S_t = m_t$, auctioneer accepts the sale request of all the active sellers. On the other hand if the number of items traded is less than available units i.e. $S_t < m_t$, auctioneer accepts the sale request of randomly selected $S_t$ number of sellers. The buyers are ranked in ascending order of their bids and $S_t$ highest bids are accepted to allocate the item. Each winning bidder gets one unit of the item and pays a price equal to the highest losing bid.

To formally determine the clearing price, we denote the order statistic for bid as $b^{(l)}$, which represents the $l^{th}$ highest order statistic of the bidding values. Note that $S_t$ highest bids are selected in period $t$ to trade. So the bid of a buyer is accepted for trade if $b_t(v_i, r_i) \epsilon \{b^{(1)}, b^{(2)}............b^{(S_t)}\}$. Every winning bidder pays the bid of the highest losing bidder, i.e. $z_t = b^{(S_t+1)}$. The price paid by all buyer in every period is equal to the $(S_t + 1)^{th}$ highest order statistic of that period. All buyers who bid above $z_t$ win the item at price $z_t$ and each accepted sale request yields a payment equal to $z_t$ for sellers.

We solve the repeated period problem using Subgame Perfect Bayesian Nash equilibrium as the solution concept. In each period the buyers reports a bid for single unit of the item to the auctioneer. The equilibrium defines a set of strategy and beliefs, such that given the opponents' strategies, the expected payoff of every buyer is maximized in each period.

**Observation 2.2.1.** *Notice that, although the arrival time $a_i$ is private knowledge, the buyer does not have any incentive to lie about the arrival period. This is due to the independence of buyer's payoff function and the arrival time $a_i$. Thus the relevant private information of buyer $i$ is $(v_i, k_i)$.*

## 2.3 Symmetric equilibrium

A bidding strategy for a bidder $i$ consists of a set of bid functions defined as $b^i = \{b_1(v_i, r_i), b_2(v_i, r_i), ....b_\infty(v_i, r_i)\}$ , where $b_t(v_i, r_i)$ denotes the bid in the $t^{th}$ period given that the bidder's value is $v_i$ and demand will last for $r_i$ periods after period $t$.[7] Note that the bidders arrive and leave at different time periods , thus the subscript on period $t_{r_i}$, i.e $r_i$, denotes the number of active periods for bidder $i$ in period $t_{r_i}$.

As this is a dynamic setting, we start by defining the state variables. Let $\sigma_t = \{m_t, n_t\}$ denotes the auction "state" consisting of the number of bidder $i$ at time period $t$, and it is represented with a k-dimensional matrix $n = [n_{\bar{k}}, ..., n_1]$ and $m$ represents the number of sale units available in that period. Here, $n_x$ denotes the number of bidders with $x$ periods of active demand. Let $N(k, n) = \sum_{x=0}^{\bar{k}} n_x$ denote the set of active buyers in current period.

Let us now consider the maximizing objective function for the buyer which is the total lifetime payoff after realization of demand and supply , denoted by $V(v_i, r_i|\sigma)$

$$
\begin{aligned}
V(v_i, r_i|\sigma_t) \equiv \Bigg\{ & Pr\Bigg(b(v_i, r_i) > b_{j\neq i}^{(m_t)}\Bigg|\sigma_t\Bigg)\mathbb{E}\Bigg[v_i - b^{(m_t)}(v_i, r_i)\Bigg|b(v_i, r_i) > b_{j\neq i}^{(m_t)}\Bigg] \\
& + Pr\Bigg(b(v_i, r_i) > b_{j\neq i}^{(m_t)}\Bigg|\sigma_t\Bigg)\tau\int_{\sigma_{t+1}} V(v_i, r_i - 1|\sigma_{t+1})\Bigg\}
\end{aligned}
\tag{2.1}
$$

The first term represents the expected payoff from auction in the first period and the second term is the future period payoff integrated over possible arrival rate $\sigma_{t+1}$. The future payoff

[7] All this assumes, of course, that he particular bidder has not already won an object so is still active in period $t$

is equal to the total payoff for value $v_i$ and $r_i - 1$ pd left. Now we define the total payoff e-ante, i.e., before the realization of the state variable '$\sigma_t$'

$$W(v_i, r_i) \equiv \int_{\sigma_{tr_i}} \tau \Big\{ Pr\Big( b(v_i, r_i) > b_{j \neq i}^{(m_t)} \Big| \sigma_{tr_i} \Big) \mathbb{E}\Big[ v_i - b^{(m_t)}(v_i, r_i) \Big| b(v_i, r_i) > b_{j \neq i}^{(m_t)} \Big] \tag{2.2}$$
$$+ Pr\Big( b(v_i, r_i) > b_{j \neq i}^{(m_t)} \Big| \sigma_{tr_i + 1} \Big) W(v_i, r_i - 1) \Big\}$$

Using the total payoff ex-ante, we can rewrite the ex-post per period payoff in eqn(2.1) as

$$V(v_i, r_i | \sigma_{tr_i}) \equiv \Big\{ Pr\Big( b(v_i, r_i) > b_{j \neq i}^{(m_t)} \Big| \sigma_{tr_i} \Big)$$
$$\mathbb{E}\Big[ v_i - W(v_i, r_i - 1) - b^{(m_t)}(v_i, r_i) \Big| b(v_i, r_i) > b_{j \neq i}^{(m_t)} \Big] + W(v_i, r_i - 1) \Big\} \tag{2.3}$$

*Pseudo type for each period:*

Observe that we have the probability of winning in each period defined in terms of the bid, in order to define that in term of the bidder's type, we will define a per period pseudo type for each period $t$ and bidder $i$ as $\eta_{i,t}$. Let $G_t(.)$ be the distribution for $\eta_{t,i}$. The pseudo type is defined as follows :

$$\eta_{i,t} = \begin{cases} v_i - \sum_{l=t+1}^{k_i} \int_{\sigma_l} \tau^{l-t} G_l^{(m_l)}(\eta_{i,l}|\sigma_l) \Big( \eta_{i,l} - \mathbb{E}[(\eta_{j,l}^{(m_l)})_{j \neq i} | \eta_{i,l} > \eta_{j,l}^{(m_l)}] \Big), & \text{if } t \epsilon \{a_i, k_i\} \\ 0, & \text{otherwise} \end{cases}$$

The pseudo type helps simplify the payoff function as well as the bid. Following lemma shows that relation between pseudo type and the payoff function

**Lemma 2.3.1.** *The equilibrium bid is increasing in pseudo type and the total payoff in period t can be rewritten as follows :*

$$W(\eta_{i,t}) = \int_{\sigma_t} \tau G_t^{(m_t)}(\eta_{i,t}|\sigma_t) \Big( \eta_{i,t} - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j \neq i} | \eta_{i,t} > \eta_{j,t}^{(m_t)}] \Big) + W(\eta_{i,t+1}) \tag{2.4}$$
$$V(\eta_{i,t}) = G_t^{(m_t)}(\eta_{i,t}|\sigma_t) \Big( \eta_{i,t} - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j \neq i} | \eta_{i,t} > \eta_{j,t}^{(m_t)}] \Big) + W(\eta_{i,t+1}) \tag{2.5}$$

We can use this equation in order to determine equilibrium bid functions, as demonstrated in the following result. We focus on symmetric Bayesian Nash equilibrium.

**Theorem 2.3.2.** *If auction in every period is vickery auction then the equilibrium bidding strategy in the dynamic auction game is given as:*

$$b_i(\eta_{i,t}) = v_i - W(\eta_{i,t+1})$$

or

$$b_i(\eta_{i,t}) = \eta_{i,t}$$

## 2.4 Applications and comparative statics : emission market

Here we look at the case of fixed decreasing supply with an uncertainty of demand in each period. In the emission auction markets, the goal is to reduce the supply of emission credits, thereby reducing the total emission. This is usually done by deciding in advance the decreasing rate of supply of emission credits. The general model, introduced in the previous section, is modified in terms of supply rate to fit the market better. Instead of having stochastic supply, for this section we have the supply decreasing at a constant rate.[8] Let the rate at which supply decreases be denoted by $\lambda \, \epsilon \, [0, 1]$. This changes the state variable $\sigma_t = (\lambda m_t, n_t)$. Additionally, the probability of active demand $\tau$ will affect the future as well as the current value for the object. This is because any change in demand for credit in the future affects current demand as well. Everything else is the same. In this section, we look at comparative statics in terms of how $\lambda$ and $\tau$. This will give insight into how the supply rate and the demand uncertainty affect bidder behavior. The first two propositions look at how the supply rate, i.e. $\lambda$ affects the market variables. This is done to understand how the supply rate can be strategically used to stabilize the price.

**Proposition 2.4.1.** *The bidder's expected payoff is increasing with increase in the future supply rate.*

$$\frac{d(V(\eta_{i,t}|\sigma_t))}{d(\lambda)} \geq 0 \tag{2.6}$$

The above proposition shows that the expected payoff in each period $t$ is positively correlated with the future periods' supply rate. The intuition behind the result is that as the supply in the future increases, the chances of acquiring credits in the future increase, thereby increasing the total expected payoff from current and future periods. However, our purpose is to understand how the supply rate can be used to control the price. In the next

[8]To focus on the effect of supply and demand dynamics we will be abstracting away from the multi-unit demand nature of this market

proposition, we shift our focus to see how the supply rate affects the bid as well as the equilibrium price.

**Proposition 2.4.2.** *The bid and price are decreasing with increase in the future supply rate.*

$$\frac{d(b(\eta_{i,t}))}{d(\lambda)} \leq 0$$
$$\frac{d\mathbb{E}(P_t)}{d(\lambda)} \leq 0$$

The price and the bid in the current period are negatively affected by an increase in the supply rate. Thus, even though the total payoff is increasing for the bidder, they will still decrease the price in the current period. This is because increasing the supply rate increases the payoff in the future periods, which also increases the opportunity cost of winning in the current period ($t^{th}$ period). Therefore, bid and thereby price is negatively correlated to supply rate. We have now established the effect of the supply rate on market variables. However, in order to see how the supply rate can stabilize the effect of demand uncertainty on the price we need to look at the effect of demand uncertainty as well. We now look at how do market variables react to demand uncertainty.

**Proposition 2.4.3.** *The bidder's expected payoff is decreasing in the uncertainty of demand.*

$$\frac{d(V(\eta_{i,t}|\sigma_t))}{d(1-\tau)} \leq 0 \tag{2.7}$$

Recall that $\tau$ represents the probability that bidders will have demand for the credits in the next period.[9] Thus, $1-\tau$ represents the probability of demand reduction in the next period. For this study, we take $1-\tau$ as a proxy for expectations about future demand uncertainty. The proposition(2.4.3) shows that the payoff is negatively affected when the demand uncertainty increases in the market.

[9]this is assuming bidder has active demand next period and is between his entry and deadline period.

**Proposition 2.4.4.** *The current bid and price increase with decrease in expected demand uncertainty.*

$$\frac{d(b(\eta_{i,t}))}{d(1-\tau)} \leq 0$$
$$\frac{d\mathbb{E}(P_t)}{d(1-\tau)} \leq 0$$

This shows that increase(decrease) in expected demand increases(decrease) the price. Now that we have established the individual effects of supply rate and uncertainty on price, we will look at how to utilize it to stabilize the price path. From propositions 2.4.2 and 2.4.4 we see that $\lambda$ and $\tau$ have an opposite effect on the price. Thus, an optimal strategy will negatively correlate the future supply rate and expected demand. Let the proposed supply rate be defined as follows :

$$\lambda_{new} = f(\tau) \tag{2.8}$$

$$f'(\tau) \geq 0 \tag{2.9}$$

The new supply rate changes with the change in demand. The next proposition proves that the change in price due to change in demand uncertainty would be less in case of the new supply rate.

**Proposition 2.4.5.** *If the demand in the future changes by $x$ then, the change in price would be lower in case of the new supply rate.*

$$\frac{d\mathbb{E}(P_t|\lambda_{new})}{d(\tau)} < \frac{d\mathbb{E}(P_t|\lambda)}{d(\tau)}$$

### 2.4.1 Comparison with price control strategies

The literature provides a fixed price ceiling and reserve price as a solution to control the price fluctuation. However, even after the price caps, there can be high variance in price within the bounded prices (as seen in figure(C.1). In this paper, we instead present a more flexible strategy, where the supply rate would be linked to future uncertainties and thereby adjust higher for high-level uncertainty and low for low-level uncertainty. Let's take an example of the RGGI carbon credit market. Figure(C.2) shows the price path with major events during the timeline. We can see that price fluctuates with changes in the demand such as in 2106 we see a big fall in price as the demand uncertainty increased as in 2016 supreme court halted the Environmental Protection Agency's Clean Power Plan. In our model, the supply rate would have temporarily adjusted to decrease the fall.

This work shows that linking the supply rate to changing market factors will stabilize the price. Further, empirical analysis is required to understand the implementation of such a policy.

## 2.5 Conclusion

This paper analyzes a dynamic auction setting with the stochastic arrival of bidders and multidimensional bidder's type. We derive the BNE bid in the repeated game. We apply this setting to the Cap and Trade scheme for the auction of carbon credits. The setup is used to understand the fluctuation of price in this market. Price uncertainty is a major concern in the Cap and Trade market. The reason for that is that for firms to change to more renewable energy they would need a more stable short term supply of carbon credits. The paper identifies two factors affecting price fluctuation, namely the rate of supply and uncertainty in future demand. One of them is not in control of the auctioneer (or the government in this case), that is the future uncertainty in demand. However, the rate of supply is decided by the government. The paper shows that the effect of the rate of supply on price is opposite to the effect of uncertainty in future demand. Thus, one policy suggested in this paper is that the government should correlate the supply rate with the uncertainty in the market. Specifically, they should decrease the future supply rate when the future demand uncertainty in the market increases. The results show that this policy would lead to a more stable price over time. The paper also analyzes the general model setup, which can be applied to other markets as well. One key feature of the paper is that we look at the multi-dimensional type of bidders. Thus, this model can fit markets where the bidders differ in more ways than just the value for the object. Future extensions of this paper can look at extending this setting to multi-unit demand and looking at double auction setting which can further generalize the auction design.

This work shows that linking the supply rate to changing market factors will stabilize the price. Further, empirical analysis is required to understand the implementation of such a policy.

# CHAPTER 3

## REVENUE-MAXIMIZING NUMBER OF ADS PER PAGE IN THE PRESENCE OF MARKET EXTERNALITIES[†]

## 3.1 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 E.1 gives an example of the Yahoo! result page.

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 per page.

[†]This work is unpublished.

[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.

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 E.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

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

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 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

[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) Agarwal & Mukhopadhyay (2016)

[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.

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 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

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

[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.

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 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.

## 3.2 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' will-

[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.

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

ingness 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 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 Vickrey–Clarke–Groves (VCG) auction, for example, in Sayedi, Kinshuk, Baghaie

[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).

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

## 3.3 Market environment and theoretical Model

### 3.3.1 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(E.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(E.2)).

**Stage 3.3.1.** *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 3.3.2.** *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.

**Stage 3.3.3.** *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 E.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[15] 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 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 position won by keyword-ad observation,[16] and on the advertisers' side, it reports the advertiser's bid and

[15] 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)Börgers et al. (2013).

[16] aggregated on a daily basis

winning positions.

### 3.3.2 Theoretical 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. The reason for modeling the consumer side is to get an estimate of relative click rate of different positions and each advertiser's quality score. In the section below, I specify the click behavior using a discrete choice model. In the next section, I present a model of advertiser's equilibrium bid, in which the advertisers choose the bid that maximizes their profit from advertising.

### 3.3.3 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 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:

$$U_{i,j,m} = U(x_j, k_j, x_i; \eta^j, \eta^i) + \epsilon_{i,j,m} \tag{3.1}$$

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[17]. 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[18]). 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 3.3.4.** *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.[19] 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.

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

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

[19]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

Then the click decision can be written as follows:

$$y^*_{i,j} = \begin{cases} 1 & \text{if } U(x_j, k_j, x_i; \eta^j, \eta^i) + \epsilon_{i,j,m} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

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[20] the click probability can be rewritten as:

$$ctr_{j,k} = s_j c_k \qquad \forall\ j \ \epsilon \mathscr{J}\ \&\ \text{k}\ \epsilon\ \{1, 2, ...K\} \tag{3.3}$$

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(3.3) is used in advertiser's maximizing problem as described in the next section.

### 3.3.4 Advertising model

#### 3.3.4.1 Model without externality

In this section, I first specify the auction design. I then model the advertiser's maximization problem and derive the equilibrium bid. In the end, I look at the case of added externality to the per-click ad value.

**Auction Setup**

For each consumer query, a single auction is held to sell multiple ads on the result page. Let there be $K$ ads on sale and $N$ potential advertisers denoted by $j\ \epsilon\ \mathscr{J} := \{1, ...J\}$. Each advertiser submits a single bid $b_j$, which can be interpreted as per click payment.

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

The $K$ positions are allotted through a weighted Generalized Second Price auction (referred to as $GSP^w$ henceforth). Each advertiser is assigned a quality score, $s_j$, which measures the ad's impact on the consumer's click probability.[21] The advertisers are then ranked in terms of their weighted bids denoted as $b_j^w$, i.e.,

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

The auction assigns position in decreasing order of weighted bids, essentially allotting $k^{th}$ 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:[22]

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

Here $b_w^{[k]}$ 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, which is equivalent to the weighted bid of the advertiser below him divided by his score.

**Advertisers maximization problem**

Each advertiser $j$ has a per click ad value of $v_j$. The advertiser's gain from winning position $k$ is given as

$$\pi_k(v_j) = ctr_{k,j}(v_j - p_k) \tag{3.4}$$

[21]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.

[22]I assume no reserve price for simplicity.

where $v_j$ denotes the per click value to the advertiser and $p_k$ is the per click price for position $j$. The term $ctr_{k,j}$ denotes the probability of a consumer click for ad $j$ in position $k$. The probability of a click can be further broken down as

$$ctr_{k,j} = s_j c_k \tag{3.5}$$

here $s_j$ is the advertiser's affect on click probability and $c_k$ is the position affect on click rate. These variables are estimated on the consumers side.[23]

Note that the advertiser is uncertain as to which position he wins. Thus, the expected profit is given as the profit from each position times the probability of winning that position. Before we write the expected profit, we define the order statistic among bids. let $b^{[k]}_{-j,w}$ denoted the $k^{th}$ highest weighted bid among all other advertisers except $j$.

Advertiser $j$ wins position $k$ if the weighted bid $b^w_j$ is less than $(k-1)^{th}$ weighted bid and more than $k^{th}$ weighted bid, i.e. $b^{[k]}_{-j,w} \leq b^w_j \leq b^{[k-1]}_{-j,w}$. As shown below:

$$\Pi(b_j; v_j, s_j) = \sum_{k=1}^{K} Prob(b^{[k]}_{-i,w} \leq b^w_j \leq b^{[k-1]}_{-i,w}) * ctr_{k,j} \left[ v_j - \mathbb{E}\left( p_k \middle| b^w_j = b^{[k]} \right) \right] \tag{3.6}$$

## 3.4 Equilibrium bid analysis

*Model Assumptions* – This model looks at the incomplete information case. Each advertiser does not know the bid 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 referred as the weighted value, denoted as $\omega_j \equiv s_j \times v_j$), $\omega \sim F_w(.)$. The weighted value is the product of the per click value and the click rate of the advertiser. Apart from this, the number of advertisers and ads per page are common knowledge. For the

[23]Note, that as I assume one ad per advertiser for each auction, I have used subscript $j$ interchangeably for advertiser as well as an ad.

estimate of the externality we also need the additional assumption that the average quality of ads in each position is known to each advertiser[24].

Given the setup, the equilibrium bid maximizes the following objective function:

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

Using $ctr_{j,k} = c_k s_j$ we get,

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

In order for the advertisers to solve for the equilibrium bid they need to know the distribution of other bids, however we assume that they only know the distribution of the value and do not have information about other bids (i.e. incomplete information case). In the standard auction this is solved by inverting bid and using the value distribution. However, in this case the weighted bid $b_w(s_j, v_j))$ is is multi-dimensional as it depends on the value $v_j$ as well as the score $s_j$.

Due to this limitation most of the papers in the literature have concentrated on non-weighted auction in the incomplete information case. This paper over comes the problem of bid by showing equivalence between the weighted GSP and another auction which I refer as the modified GSP auction (GSPM henceforth). In the GSPM auction, instead of a per click bid the advertisers bid $\tilde{b}$ for $s_j$ clicks. The value for $s_j$ click then becomes $w_j$, which recall is the weighted value $w_j = s_j v_j$. Thus, the bid in GSPM serves a similar purpose as the weighted bid in GSP auction. The difference being that in GSPM the bid can be written as a single dimension function given as follows:

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

$b^{GSPM}(\omega_j)$. In the next lemma, I formally prove the equivalence between the weighted bid in weighted GSp to the bid in GSPM.

[24]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.

**Lemma 3.4.1.** *The weighted GSP auction equilibrium bid function* $b^{GSP^w}(v_j, s_j) \rightarrow \mathscr{R}_+$ *is equivalent to the bid in GSPM auction divided by the quality score* $s_j$

$$\frac{b^{GSPM}(\omega_j)}{s_j} = b(v_j, s_j), \qquad \text{where } \omega_j = s_j \times v_j$$
$$\rightarrow b^{GSPM}(\omega_j) = b(v_j, s_j) * s_j = b_j^w(v_j, s_j) \tag{3.7}$$

The lemma shows that the weighted bidding function is equivalent to a function that is dependent only on one-dimensional advertiser's type $\omega_j$. Additionally, it shows that at equilibrium we can rewrite the weighted bid $s_j b_j$ as a function of weighted value , i.e. $b_w(\omega_j)$. This simplification comes in handy for proof of bounds and equilibrium as now inverse of bid function is one dimension.

**Proposition 3.4.2.** *The unique symmetric Bayesian Nash equilibrium of the weighted GSP auction is given by the following bidding strategy:*

$$b^*(\omega) = \omega - \Gamma(\omega) - \sum_{n=1}^{\infty} \int_0^{\omega} M_n(\omega, t)\phi(t)dt \qquad \forall \omega \sim F_w(.) \tag{3.8}$$

*where*

$$\Gamma(\omega) = \frac{\sum\limits_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-2} \int_0^{\omega} F^{N-k}(x)dx}{\sum\limits_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-1} F^{N-k-1}(\omega)}$$

$$M_1(\omega, t) = \frac{\sum\limits_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-2} F^{N-k-1}(t)}{\sum\limits_{k=1}^{K} c_k \frac{N-2}{k-1}(1 - F(\omega))^{k-1} F^{N-k-1}(\omega)}$$

$$M_n(\omega, t) = \int_0^{\omega} M_1(\omega, \epsilon) M_{n-1}(\omega, \epsilon) d\epsilon \qquad \forall n \geq 2$$

The above proposition shows the equilibrium bid[25]. The equilibrium bid in equation (3.8) has some issues regarding empirical identification. First, the structural elements can

[25] Once the lemma(3.4.1) is used to make the objective function depend on weighted value only , the subsequent proof of the equilibrium is similar to BNE derived in Gomes & Sweeney (2014)

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. In order to derive identification of equilibrium bid empirically, I derive bounds on the bid as shown in the next section.

### 3.4.1 Bounds on 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 have a closed and easily tractable equilibrium bid. Specifically, I use Vickrey–Clarke–Groves 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 3.4.3.** *The bid shading in $GSP^w$ auction can be bounded between the bid in analog VCG and generalized first price (GFP) auctions.*[26]

$$(v_j - b_j)_{VCG} \leq \underbrace{(v_j - b_j)_{GSP^w}}_{\textit{price less than bid}} \leq \underbrace{(v_j - b_j)_{GFP^w}}_{\textit{price equal to bid}} \qquad \forall j \epsilon \daleth \tag{3.9}$$

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 3.4.4.** *If the proposition(3.4.3) holds, advertiser's value can be bounded in*

[26]proof of proposition (4) and (5) are given in appendix(G.1)

*terms of observed variables*[27]*:*

$$b_j \leq v_j \leq b_j + \Phi(f_w(.), c_k, b_j) \qquad \forall j \epsilon \daleth \tag{3.10}$$

Here $\Phi(f_w(.), c_k, b_j)$ is the bid shading amount, which shows that advertisers would bid below their value. The term can be expanded in terms as shown below

$$\Phi(f_w(.), c_k, b_j) = \frac{\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)} (1 - F(\omega_j))^{k-1} (F(\omega_j))^{N-k}}{s_j \frac{f(\omega_j)}{b'(\omega_j)} \sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)} \left[ (N-k)(1 - F(\omega_j))^{k-1} F(\omega_j)^{N-k-1} + (k-1)(1 - F(\omega_j))^{k-2} F(\omega_j)^{N-k} \right]}$$

The numerator is the probability of winning any position in the auction and the denominator the differentiation of that probability. Notice that the function $\Phi(f_w(.), c_k, b_j)$ has an unobserved distribution $f_w$ i.e., distribution of weighted value. However, we can substitute that with the bid distribution as the weighted bid is monotonic. I will use the following equality conditions :

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

Thus, the function

$$\Phi(g_b(.), c_k, b_j) = \frac{\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)} (1 - G(b_w))^{k-1} (G(b_w))^{N-k}}{s_j g(b_w) \sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)} \left[ (N-k)(1 - G(b_w))^{k-1} G(b_w)^{N-k-1} + (k-1)(1 - G(b_w))^{k-2} G(b_w)^{N-k} \right]} \tag{3.11}$$

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.

[27] proof in appendix(G.1)

#### 3.4.1.1 Model with externality

Now we look at the Advertiser's model with externality. The aim of this subsection is to show how we can separate the value with and without externality.

The per-click value to the advertiser is compromised of two-part, (i) an advertiser-specific ad value and (ii) a negative value from the number of other ads on the page – externality component. To have a simple and tractable model we assume that the two-component are separable. Altogether the per click value for buyer $j$ is given as:

$$V_j = v_j (EXT_j)^{\beta_1} \tag{3.12}$$

where $v_j$ corresponds to the externality free part as analyzed in the previous section. The second part captures the externality which depends on other ads $EXT_j$ is a function of $K_j$, which denotes the set of ads present next to ad $j$. Note $\beta_1$ captures the intensity of externality's effect on ad value. In the empirical section I well estimate $\beta_1$ ( refer to step(3.7.2) of the estimation section for more details).

**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 eqn(3.12), I can derive the relation between externality and the advertisers' bid; the next proposition formally states the result.

**Proposition 3.4.5.** *The additive separability of externality and individual effect on valuation leads to additive separable bid parameter (after controlling for various factors).*[28]

$$Log(\underbrace{b(v_j, s_j; \alpha_{K,n}, Ext_j)}_{\textit{observed bid}}) = Log(\underbrace{b(v_j, s_j; \alpha_{K,n})}_{\textit{externality free bid}}) + \underbrace{\beta_1 Log(Ext_j)}_{\textit{externality effect on bid}} \tag{3.13}$$

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 $K$ denotes the number of ads and $n$ denotes the number of advertisers. The

[28] proof in the appendix(G.1)

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 $\beta_1$ that captures the intensity of externality's effect on ad value. 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.

## 3.5 Data

### 3.5.1 Data details

The data set is provided by Yahoo! as part of the *Yahoo! Research Alliance Webscope program.*[29] It is a four-month period of data covering search queries from January 2008 to April 2008. The sample covers all search ads in 5 categories, namely *Laptop*, *TV Cable, Cruise, Collectible Coins* and *Car Insurance.*[30] 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.[31] This is mainly done because 90% of the clicks in the data are from first page ads, as can be seen by figure E. This finding is also consistent with the observed pattern in

[29]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

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

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

the industry, which shows that the consumers do not go beyond the first page.[32] For more information on data cleaning refer to appendix(J.0.1).

*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 the $1^{st}$ position, and that translated into five clicks. Table(F.1) gives the list of variables used and table(F.3) gives the summary statistics. Notice that this is similar to feedback given to an advertiser in figure E.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,[33] 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[34] 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.

[32] 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/

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

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

For each ad,[35] 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.[36] Table F.2 lists the variables and table F.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.[37] 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 his corresponding bid.[38] 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-

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

[36] 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

[37] 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.

[38] 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.

groups are created such that only the ones within a market compete with each other.[39] 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(F.5) shows mean values for different variables across categories.[40] The relative popularity of these categories compared to the total daily search can be seen in figure(E.5); laptop has a high popularity for consumer search with value of 55% meaning it is more than half as popular as the peak popular search topic measured as 100%. 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.

### 3.5.2 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

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

[40] 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(J.0.2) for more details

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(E) shows the mean bids for different winning positions in the five categories[41]. 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. Note that as the auction mechanism decides the allocation of the ad position according to the weighted bid, it is not necessary for a higher bid to get a higher position.[42]

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 E.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.

[41]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)

[42]if there are two advertisers such that $b_1 > b_2$, advertiser 2 might get a higher position if he has a higher weighted bid, i.e. $b_s * s_2 > b_1 * s_2$.

3. **Externality:** An important observation from the data, is the preliminary evidence of externality. 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 E.11 shows the linear relation between the average bid and the percentage of high-quality ads in the market.[43] 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.[44]

## 3.6 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.

### 3.6.1 Advertisers 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 (3.10) imply for each advertiser $i$, we have the following

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

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

bounds:

$$\hat{b}^{ext}_{j,m} \leq v_{j,m} \leq \hat{b}^{ext}_{j,m} + \Phi(\hat{g}(.|\hat{b}), \hat{c}_{k,m}, \hat{b}^{ext}_{j,m}, \hat{s}_{j,m}) \qquad \forall \quad j, m \tag{3.14}$$

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(3.14), 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(3.14) gives the following result:

$$H_\phi(.) \leq F_v(.) \leq H_b(.) \tag{3.15}$$

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.

### 3.6.2 Consumers click behavior

I use the utility function of consumers in equation (3.2) 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.

### 3.6.3 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 $j$.[45] The quality estimate, $\hat{s}$, is derived from the consumer side analysis[46]. The quality is denoted on 0 to 1 range with 1 being the highest quality. Fig(E.18) 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.[47] 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.[48]

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

[46]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.

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

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

### 3.6.4 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 ads such that only the ones within a market compete with each other. In this section, I give details of how the markets are created.

Fig(E.14) 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, namely the attributes that capture the similarity between ads and the distance matrix used.

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. Additionally, the score can be improved by putting more weight on a less frequent word being matched, under the assumption that such words are more valuable signal for grouping similar keywords than words that frequently appear across all keywords. In order to do this, I use tf-idf (term frequency inverse document frequency)[49] as a weighting factor. The weight increases proportionally to the number of times a word was used for an ad, and is offset by the overall frequency of the word, which helps to adjust for the fact that some words appear more frequently in general.[50]. Apart from using words as attributes for the ad, I also need to specify the distance metric used between keywords. There are several

[49]tf-idf is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus.

[50]tf–idf is one of the most popular term-weighting schemes today; 83% of text-based recommender systems in digital libraries use tf–id.

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(E.14), the graph plots the distance between keywords for various distance measures[51]. 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 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. 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 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:

$$\lambda_{l,m} = \begin{cases} 1, & \text{if } x_l \text{ is in cluster } m \\ 0, & \text{otherwise} \end{cases} \tag{3.16}$$

Let $\theta_m$ be the chosen centroid of group $m$. The cluster algorithm decides the allocations $(\lambda)$ and centroids$(\theta)$ by maximizing the mean squared distance between points within the

[51]compressed in two dimensional space

group, as shown below:

$$\min_{\theta,\lambda}\left(1-\sum_{m=1}^{m}\lambda_{l,m}\frac{X.\Theta_m}{||X||||\Theta_m||}\right) \tag{3.17}$$

The k-means algortihm used in this paper uses cosine distance to calculate similarity between points[52]. Equation(3.17) 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 $\theta$ is picked given $\lambda$. The algorithm is given as:

$$\text{Step:1} \quad \min_{\lambda}\left(1-\sum_{m=1}^{M}\lambda_{l,m}\frac{X.\Theta_m}{||X||||\Theta_m||}\right) \qquad \forall i$$

$$\text{Step:2} \quad \sum_{l=1}^{N_m}\lambda_{l,m}(x_l-\theta_m)=0 \qquad \forall m$$

The steps are repeated until convergence of $\lambda$ and $\theta$. 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 using 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.

## 3.7 Econometric Method

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

- 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.

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

### 3.7.1 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.[53] 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,[54] 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:

$$U_{i,j,m} = \phi_0 + \Phi_1 X_{\text{position dummy}} + \Phi_2 X_{\text{advertiser dummy}} + \Phi_2 X_{\text{controls}} + \epsilon_{i,j,m} \qquad (3.18)$$

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[55] as well as the keylength which captures how detailed is the search. Lastly the 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

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

[54] similar set up is considered in Athey & Nekipelov (2010)

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

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(3.18).

**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.[56]

- 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 $\hat{Ext}_{j,m} = \sqrt{\sum_{k \neq K_j} \hat{s}^2_{-j,m}}$[57], where $\hat{s}$ is the average quality of an ad in position $k$ for all ads except ad $j$ in market $m$.

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

The goal of this step is to recover the externality impact on the bid. Proposition (3.4.5) 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 Haile et al. (2003) for english auction (this paper shows

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

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

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

$$log(b_{m,j}) = \beta_0 + \beta_1 log(Ext_{m,j}) + Z_m\beta_2 + \beta_3\hat{s}_j + \delta_{N,K} + u_{m,j} \tag{3.19}$$

$$\hat{Ext}_{m,j} = \text{estimated externality in the market}$$

$$Z_m = \text{Market characteristics}$$

$$\delta_{N,K} = \text{Dummy variables for number of advertisers, ads and ad-position}$$

$$\hat{s}_j = \text{estimated ad quality}$$

In above equation, $\beta_1$ 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[58]). 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 $\beta_1$. 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:

$$Ext_{m,j} = \gamma_1 Ext_{m',j} + \Gamma_2 Z_m + \gamma_2\hat{s}_j + \delta_{N,K} + \mu_{j,m} \tag{3.20}$$

**Parameters estimated**

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

- $\beta_1$: This measures the impact of externality on bid. As externality oncreases by 1%, bid will decrease by $\beta_1$ percent.

- Externality free bid - the externality-free bid is residual when the externality effect is removed from the observed bid $b_{j,m}^{ext-free}$.

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

$$\hat{g}(b) = \frac{1}{\delta_g} \sum_j \sum_m 1\{\hat{b}_{j,m}^{[w]ext-free} \leq b\} K\left(\frac{\hat{b}_{j,m}^{[w]ext-free} - b}{\delta_b}\right) \tag{3.21}$$

### 3.7.3 Step 3: 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(3.10) from the theory section,as reproduced below:

$$\hat{b}_{j;m,t}^{ext-free} \leq v_j \leq \hat{b}_{j;m,t}^{ext-free} + \Phi(\hat{g}(.), \hat{c}_k, \hat{b}_j^{ext-free}, \hat{s}_j) \quad \forall \quad j \quad \epsilon \quad \{1, 2, ....n_{m,t}\} \tag{3.22}$$

From the bounds on value in equation(3.14), 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(3.14) gives the following result:

$$H_\phi(.) \leq F_v(.) \leq H_b(.) \tag{3.23}$$

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

$$\hat{H}_b(h) = \frac{1}{\delta_h} \sum_j \sum_m 1\{\hat{b}_{j,m} \leq h\} K\left(\frac{\hat{b}_{j,m} - h}{\delta_h}\right) \tag{3.24}$$

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

$$\hat{H}_\phi \leq F_v \leq \hat{H}_b \tag{3.25}$$

## 3.8 Results

In this section, I discuss the results that are reported in figure(E.15-E.23) and tables (F.6-F.8). 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.

### 3.8.1 Consumers 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.[59]

*Position effect on click probability:* Table(F.8) shows results of average partial effect(APE). Fig(E.15) 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(E.16), 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

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

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 I normalize it to a 0-1 range, with 1 being the highest quality ad.[60] As shown in graph(E.17), 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(E.18) shows the distribution of externality variable across different categories.

### 3.8.2 Externality effect

Table(F.6) shows results of the impact of externality on advertiser's bid using Eqn(3.19). 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(E.19) plots the estimated $\beta_1$ co-efficient that captures the percentage change in bid when externality is increased by 1% . The results show that 1%

[60] The top 0.01% are given a score of 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. In appendix(I.0.1) I look at alternative definition of externality and other robust checks.

### 3.8.3 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(E.20) shows the percentage decrease in the bid when the display goes from an exclusive to two-ad display.[61] The graph shows how the affect of the additional ad depends on its quality.[62] 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

[61] 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

[62] I assume the base externality without any ads on the page is one.

after a certain number, more ads by competitors do not impact the advertiser. The graph(E.21) 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.

### 3.8.4 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(3.22), I get the bounds on the advertiser's maximum willingness to pay for an ad or ad value. The distribution bounds are estimated for each category, as shown in the graph(E.23). The bounds are tight for three out of 5 categories, implying that inequality is sufficient for inference. The graph(E.22) 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.[63]

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

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

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.

## 3.9 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.[64] For each case of the number of ads page, I derive the equilibrium bid and thus, the expected revenue for Yahoo!.[65] 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 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 ($\hat{\beta}_1$).

[64]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.

[65]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(3.4.3). 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.

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).

### 3.9.1 Results

The results from the counterfactual simulation are presented in table(F.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.

## 3.10 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.[66] 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 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

[66] next to a high-quality ad

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.[67]

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

[67] 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/

that show multiple ads to the consumer.[68]

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.

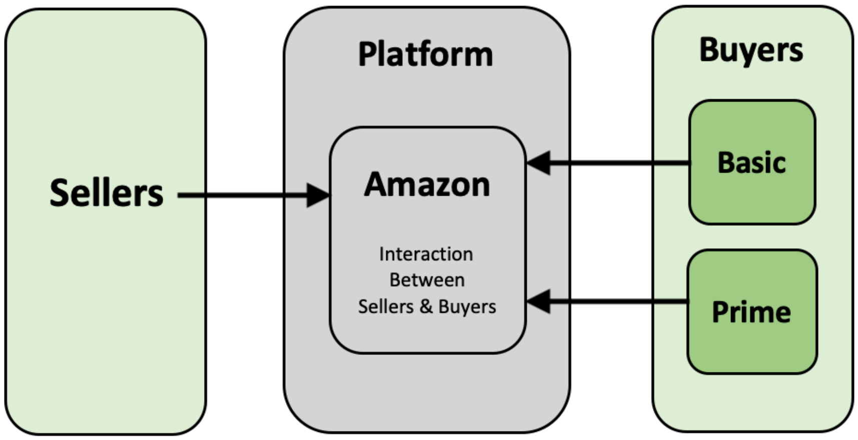
[68] Note that the application is limited to the case of the adjacent ads being from the same ad product category.

# APPENDICES

# APPENDIX A

## Figures for Chapter 1

Figure A.1: Two-sided firm with quality differentiation



## APPENDIX B

### Proofs for Chapter 1

**Details of the equilibrium price and quality equations**: The monopoly problem as in Equation (1.4) can be rewritten as follows:

$$\max_{(p^s, p_l^b, p_h^b, q_l, q_h)} \quad \Pi = [p_l^b + p^s - c(q_l)]D^b D^s + [(p_h^b - c(q_h)) - (p_l^b - c(q_l))]D_h^b D^s. \tag{B.1}$$

Using Equation (B.1), the first order conditions for choice variables can be simplified as follows.

- **Price for low-quality access on the buyer side:** Equation (1.19) can be obtained by the following steps.

$$(D^b - D_h^b)D^s + [p_l^b + p^s - c(q_l)](D^b)'_{p_l^b} D^s + [(p_h^b - c(q_h)) - (p_l^b - c(q_l))](D_h^b)'_{p_l^b} D^s = 0.$$

$$\Leftrightarrow \quad (D_l^b) + [p_l^b + p^s - c(q_l)](D^b)'_{p_l^b} + [(p_h^b - c(q_h)) - (p_l^b - c(q_l))](D_h^b)'_{p_l^b} = 0,$$

where $(D_k^b)'_{p_k^b}$ means that $\dfrac{\partial D_k^b}{\partial p_k^b}$ and $(D_k^b)'_{p_j^b}$ means that $\dfrac{\partial D_k^b}{\partial p_j^b}$, where $j \neq k$.

Using equilibrium condition for $p_h^b$ from Equation (1.20),

$$D_l^b + [p_l^b + p^s - c(q_l)](D^b)'_{p_l^b} + \left[ -\frac{D_h^b}{(D_h^b)'_{p_h^b}} \right] (D_h^b)'_{p_l^b} = 0.$$

Using $(D_h^b)'_{p_l^b} = -(D_h^b)'_{p_h^b}$, $D_l^b + [p_l^b + p^s - c(q_l)](D^b)'_{p_l^b} + \left[ \dfrac{D_h^b}{(-D_h^b)'_{p_h^b}} \right] (-D_h^b)'_{p_h^b} = 0.$

$$\Leftrightarrow \quad D_l^b + [p_l^b + p^s - c(q_l)](D^b)'_{p_l^b} + D_h^b = 0.$$

$$\Leftrightarrow \quad [p_l^b + p^s - c(q_l)] = \frac{D^b}{(D^b)'_{p_l^b}} = \frac{p_l^b}{\varepsilon^b}.$$

- **Price for high-quality access on the buyers' side:** Equation (1.20) can be obtained by the following steps.

$$\{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}(D_h^b)'_{p_h^b} D^s + D_h^b D^s = 0.$$

$$\Leftrightarrow \quad \{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}(D_h^b)'_{p_h^b} + D_h^b = 0.$$

$$\Leftrightarrow \quad \{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\} = \frac{D_h^b}{(-D_h^b)'_{p_h^b}} = \frac{p_h^b}{\varepsilon_h^b}.$$

- **Price for sellers:** Equation (1.21) can be obtained by the following steps.

$$\{[p_l^b + p^s - c(q_l)]D^b + \{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}D_h^b\}(D^s)'_{p^s} + D^b D^s = 0.$$

$$\Leftrightarrow \quad p_l^b + p^s - c(q_l) = \frac{D^s}{(-D^s)'_{p^s}} - \{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}\frac{D_h^b}{D^b}.$$

Using Equation (1.20), $p_l^b + p^s - c(q_l) = \dfrac{D^s}{(-D^s)'_{p^s}} - \left[\dfrac{p_h^b}{\varepsilon_h^b}\right]\dfrac{D_h^b}{D^b}$,

which can be further simplified as $p_l^b + p^s - c(q_l) = \dfrac{p^s}{\varepsilon^s} - \left[\dfrac{p_h^b}{\varepsilon_h^b}\right]\dfrac{D_h^b}{D^b}$.

- **Low-quality service for buyers:** Equation (1.22) can be obtained by the following steps.

$$(D^b - D_h^b)c'(q_l)D^s + [p_l^b + p^s - c(q_l)](D^b)'_{q_l}D^s +$$
$$[(p_h^b - c(q_h)) - (p_l^b - c(q_l))](D_h^b)'_{q_l}D^s = 0.$$

$$\Leftrightarrow (D^b - D_h^b)c'(q_l) + [p_l^b + p^s - c(q_l)](D^b)'_{q_l}$$
$$+ \{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}(D_h^b)'_{q_l} = 0.$$

Using Eqn (1.20), $-D_l^b c'(q_l) + [p_l^b + p^s - c(q_l)](D^b)'_{q_l} + \left[-\dfrac{D_h^b}{(D_h^b)'_{p_h^b}}\right](D_h^b)'_{q_l} = 0.$

Using Eqn (1.23), $D_l^b c'(q_l) + [p_l^b + p^s - c(q_l)](D^b)'_{q_l} + [c'(q_h)\dfrac{D_h^b}{(D_h^b)'_{q_h}}](D_l^b)'_{q_l} = 0.$

Using $(D_h^b)'_{q_l} = -(D_h^b)'_{q_h}$, $-(D_l^b)c'(q_l) + [p_l^b + p^s - c(q_l)](D^b)'_{q_l} - c'(q_h)D_h^b = 0$,

which is simplified as,

$$[p_l^b + p^s - c(q_l)] = \frac{D^b}{(D^b)'_{q_l}}\frac{D_l^b c'(q_l) + D_h^b c'(q_h)}{D^b} = \frac{q_l}{\nu^b}\frac{D_l^b c'(q_l) + D_h^b c'(q_h)}{D^b},$$

where $(D_k^b)'_{q_k}$ means that $\frac{\partial D_k^b}{\partial q_k}$ and $(D_k^b)'_{q_j}$ means that $\frac{\partial D_k^b}{\partial q_j}$, where $j \neq k$.

- **High-quality service for buyers:** Equation (1.23) can be obtained by the following steps.
  $\{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}(D_h^b)'_{q_h} - D_h^b c'(q_h) = 0.$
  Using Eqn (1.20), $c'(q_h)\frac{D_h^b}{(D_h^b)'_{q_h}} = \frac{p_h^b}{\varepsilon_h^b}$, which can be simplified as, $c'(q_h)\frac{q_h}{\nu_h^b} = \frac{p_h^b}{\varepsilon_h^b}$.

**Proof of Proposition 1.4.1.** It shows the platform offers a better price-quality ratio in the case of one quality offering. This basically implies

$$\underbrace{\frac{p^b}{q}}_{\text{Platform (two-sided)}} \leq \underbrace{\frac{p^b}{q}}_{\text{One-sided}}. \tag{B.2}$$

We start by comparing the profit functions for two different cases. For simplicity, we normalize $c^s$ to zero.

$$\begin{aligned} \Pi_{\text{one-sided}} &= [p^b - c(q)]D^b \equiv \Psi(p^b, q). \\ \Pi_{\text{platform}} &= [p^b + p^s - c(q)]D^b D^s \equiv \Psi(p^b, q)D^s + p^s D^b D^s. \end{aligned} \tag{B.3}$$

The one-sided monopolist's profit maximization problem is given in Equation (1.5). Let $\bar{x} = \left(\bar{p^b}, \bar{q}\right)$ be the optimal solution for the one-sided monopoly problem and $\widetilde{x} = \left(\widetilde{p^b}, \widetilde{q}, \widetilde{p^s}\right)$ for the two-sided platform. Given that the monopolist optimal value for $p^s \neq 0$,

$$\Psi(\widetilde{p^b}, \widetilde{q})\widetilde{D^s} + \widetilde{p^s}\widetilde{D^b}\widetilde{D^s} \geq \Psi(\bar{p^b}, \bar{q}). \tag{B.4}$$

As $(\bar{p^b}, \bar{q})$ is the optimal solution $\Pi_{\text{one-sided}}$; this implies that $\Psi(\bar{p^b}, \bar{q}) \geq \Psi(\widetilde{p^b}, \widetilde{q}) \geq \Psi(\widetilde{p^b}, \widetilde{q})\widetilde{D^s}$, so Equation (B.4) holds only if

$$\begin{aligned} &\widetilde{D^s}\widetilde{p^s}(D^b)_{\text{at } \widetilde{x}} > \widetilde{D^s}\widetilde{p^s}(D^b)_{\text{at } \bar{x}}. \\ \Leftrightarrow\ &\underbrace{\frac{p^b}{q}}_{\text{platform}} < \underbrace{\frac{p^b}{q}}_{\text{one-sided}}. \end{aligned} \tag{B.5}$$

Hence proved. ■

**Proof of Corollary 1.4.2.** It is guaranteed by Proposition 1.4.1. ■

**Proof of Proposition 1.4.3.** We use proof by contradiction. Let us suppose that $(p^s)_{1Q} \leq (p^s)_{2Q}$ where the subscripts $1Q$ and $2Q$ denote the model with one quality and with two qualities, respectively. This implies

$$(D^s)_{1Q} \geq (D^s)_{2Q}. \tag{B.6}$$

As $\frac{p^s}{\varepsilon^s}$ is a decreasing function of $p^s$, we derive the following:

$$\left(\frac{p^s}{\varepsilon^s}\right)_{1Q} \geq \left(\frac{p^s}{\varepsilon^s}\right)_{2Q}. \tag{B.7}$$

Using Equation (1.24), we obtain the following:

$$\left(\frac{p^b}{\varepsilon^b}\right)_{1Q} \geq \frac{p_l^b}{\varepsilon^b} + \frac{D_h^b}{D^b}\frac{p_h^b}{\varepsilon_h^b}. \tag{B.8}$$

Using first order condition for $p^b$ in the model with one quality and those for $p_l^b$ and $p_h^b$ in the model with two qualities, we obtain the following:

$$\underbrace{[p^b + p^s - c(q)]}_{\text{average profit for 1Q}} \geq \underbrace{[p_l^b + p^s - c(q_l)] + \frac{D_h^b}{D^b}\{[p_h^b - c(q_h)] - [p_l^b - c(q_l)]\}}_{\text{average profit for 2Q}}. \tag{B.9}$$

Given Equation (B.9), the firm would provide two qualities only if $\underbrace{(D^s D^b)}_{1Q} < \underbrace{(D^s D^b)}_{2Q}$ holds. This condition along with equation(B.6) imply:

$$\underbrace{(D^b)}_{1Q} < \underbrace{(D^b)}_{2Q} \tag{B.10}$$

$$\underbrace{\frac{p_l^b}{q_l}}_{1Q} > \underbrace{\frac{p^b}{q}}_{2Q}. \tag{B.11}$$

By Equation (B.8), we have $\left(\frac{p^b}{\varepsilon^b}\right)_{1Q} \geq \left(\frac{p_l^b}{\varepsilon^b}\right)_{2Q}$, which yields

$$\underbrace{p_l^b}_{1Q} \geq \underbrace{p^b}_{2Q} \tag{B.12}$$

Furthermore from Equation (B.8), we also have

$$\left(\frac{qc'(q)}{\nu^b}\right)_{1Q} \geq \left(\frac{q_l}{\nu^b}\frac{D^b_l c'(q_l) + D^b_h c'(q_h)}{D^b}\right)_{2Q} \tag{B.13}$$

Now, from the equilibrium conditions for 1Q we have

$$\begin{aligned} \frac{qc'(q)}{\nu^b} &= \frac{p^b}{\varepsilon^b} \\ \to c'(q)\frac{1-F(p/q)}{f(p/q)\frac{p}{q^2}} &= \frac{1-F(p/q)}{f(p/q)\frac{1}{q}} \\ \to c'(q)q &= p \end{aligned} \tag{B.14}$$

Similarly using 1.12, for 2Q case we get

$$q_l\frac{D^b_l c'(q_l) + D^b_h c'(q_h)}{D^b} = p^b_l \tag{B.15}$$

substituting eqn(B.14) and eqn(B.15) in eqn(B.13), we get

$$\left(\frac{p}{\nu^b}\right)_{1Q} \geq \left(\frac{p^b_l}{\nu^b}\right)_{2Q}$$

$$\to \left(p\frac{1-F(p/q)}{f(p/q)\frac{p}{q}}\right)_{1Q} \geq \left(p^b_l\frac{1-F(p^b_l/q_l)}{f(p^b_l/q_l)\frac{p^b_l}{q_l}}\right)_{2Q}$$

using eqn(B.12), we get

$$\underbrace{\frac{p^b_l}{q_l}}_{1Q} \leq \underbrace{\frac{p^b}{q}}_{2Q}.$$

This contradicts Equation (B.11). Hence $(p^s)_{1Q} > (p^s)_{2Q}$ ■

**Proof of Corollary 1.4.4.** Let $(\bar{p}^b, \bar{q}, \bar{p}^s)$ be the profit maximization variable for the single-quality case. We prove whether the platform wants to set a nonzero demand for high-quality products at this price and quality level. The demand for high-quality products will be zero if

$$\frac{p^b_h - p^b_l}{B(q_h - q_l)} = 1. \Leftrightarrow p^b_h - p^b_l = B(q_h - q_l). \tag{B.16}$$

We choose $((p_h^b)^*, (q_h)^*)$ such that Equation (B.16) is satisfied and then determine whether the first order condition on $(p_h^b, q_h)$ shows that the platform will attempt to increase demand for the high-quality product above zero. The first order condition with respect to $p_h^b$ at $(\bar{p}^b, \bar{q}, \bar{p}^s, (p_h^b)^*, (q_h)^*)$ is given as follows:

$$\begin{aligned}
\Phi^{p_h^b} &= D^s \left\{ \{[(p_h^b)^* - c(q_h^*)] - [\bar{p}^b - c(\bar{q})]\}(D_h^b)'_{p_h^b} + D_h^b \right\}. \\
&= D^s \left\{ \left\{[(p_h^b)^* - c(q_h^*)] - [\bar{p}^b - c(\bar{q})]\right\} (-f^b(1)) \right\} \quad \text{as } \frac{p_h^b - \bar{p}^b}{B(q_h - \bar{q})} = 1. \\
&= D^s \left\{ \left\{B(q_h^* - \bar{q}) - [c(q_h^*) - c(\bar{q})]\right\} (-f^b(1)) \right\}. \\
&\leq 0 \quad \text{if } f^b(1) \neq 0 \text{ and } c'(\bar{q}) < B.
\end{aligned} \tag{B.17}$$

In a similar manner, we can prove that at $q_h^*$, the first order condition is greater than zero. Thus, the platform will decrease $p_h^b$ and increase $q_h$ such that $D_h^b \neq 0$. Therefore, we see that the profit increases when offering two product qualities as long as $f^b(1) \neq 0$ and $c'(\bar{q}) < B$. $c'(\bar{q}) < B$ is true, as $c'(\bar{q}) < B$ implies that $\frac{\bar{p}}{B\bar{q}} < 1$, which holds for nonzero demand. ■

**Proof of Proposition 1.4.5.** It is trivial to prove from checking Equation (1.22) and (1.23) ■.

**Proof of Proposition 1.4.6.** The proof is in the paper ■

**Proof of Proposition 1.4.7.** The proof is in the paper. ■

**Proof of Proposition 1.4.8.** The proof is in the paper, as shown in Equation (1.26). ■

**Proof of Proposition 1.4.9.** The proof is in the paper. ■

**Proof of equivalence of this model with fixed fee case:**

*Proof.* Suppose that on one of the side the platform also charges a fixed fee. For example in case of Amazon they charge a fixed fee, i.e. Prime membership on the buyers side. The key to the equivalence is that adding fixed fee doesn't change the model as long as there are no fixed benefits for the users. This case fits the Amazon example as even though Amazon charges fixed fee, the buyers only benefit through transactions and do not have a non-transaction benefit from Amazon. **Fixed fee Case** WLOG, let us assume buyers side is also charged a fixed fee Buyers have the following utility:

$$U_i^b = (B\alpha_i^b q - p^b)N^s - P^b, \tag{B.18}$$

Here $P_b$ is the fixed fee charged to the buyers. The seller has the following utility:

$$U_j^s = (S\alpha_j^s - p^s)N^b. \tag{B.19}$$

We first analyze the user side (buyers and sellers) to identify the equilibrium demand. The equilibrium demand functions are derived from the participation constraint:

$$\begin{aligned} D^b &= Prob\left(U_i^b \geq 0\right) \Leftrightarrow D^b = 1 - F^b\left(\frac{p^b + \frac{P^b}{N^s}}{Bq}\right). \\ D^s &= Prob\left(U_j^s \geq 0\right) \Leftrightarrow D^s = 1 - F^s\left(\frac{p^s}{S}\right). \end{aligned} \tag{B.20}$$

The monopoly problem can be written as follows:

$$\max_{(p^s, p^b, q)} \Pi = [p^b + p^s - c(q)]D^b D^s + P^b D^b, \tag{B.21}$$

**Modified fixed fee case**: Now we modify the above case to show the equivalence. Let $p^b_{new} = p^b + \frac{P^b}{N^s}$ be the per transaction fee on the buyers side and let fixed fee be zero. Then the model parameters are as follows:

Buyers have the following utility:

$$\begin{aligned} U_i^b &= (B\alpha_i^b q - p^b)N^s - P^b \\ &= (B\alpha_i^b q - p^b - \frac{P^b}{N^s})N^s \quad = (B\alpha_i^b q - p^b_{new})N^s \end{aligned} \tag{B.22}$$

Thus, this shows the utility is the same in the case of (1) usage fee $p^b$ & fixed fee - $P^b$ (2) usage fee - $p^b_{new}$

The seller has the following utility:

$$U_j^s = (S\alpha_j^s - p^s)N^b. \tag{B.23}$$

We first analyze the user side (buyers and sellers) to identify the equilibrium demand. The equilibrium demand functions are derived from the participation constraint:

$$\begin{aligned} D^b &= Prob\left(U_i^b \geq 0\right) \Leftrightarrow D^b = 1 - F^b\left(\frac{p_{new}^b}{Bq}\right). \\ D^s &= Prob\left(U_j^s \geq 0\right) \Leftrightarrow D^s = 1 - F^s\left(\frac{p^s}{S}\right). \end{aligned} \tag{B.24}$$

The monopoly problem can be written as follows:

$$\max_{(p^s, p^b, q)} \Pi = [p^b + p^s - c(q)]D^b D^s + P^b D^b \tag{B.25}$$

$$= [p^b + p^s - c(q) + \frac{P^b}{D^s}]D^b D^s \tag{B.26}$$

$$= [p_{new}^b + p^s - c(q)]D^b D^s \tag{B.27}$$

Thus, as shown above the profit function for the platform and the users utilities are the same in these two case. ■

# APPENDIX C

## Figures for Chapter 2

Figure C.1: Decrease in fluctuation due to price ceiling and floor

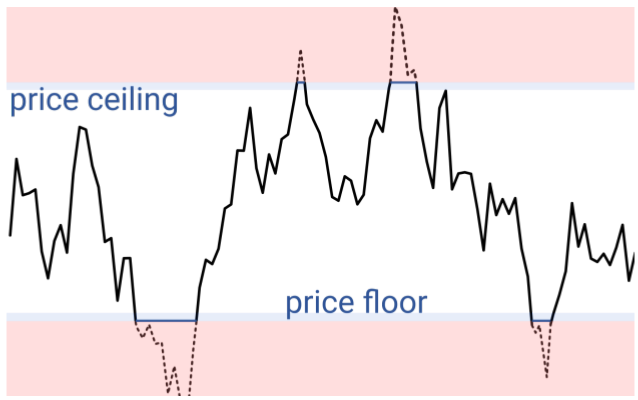
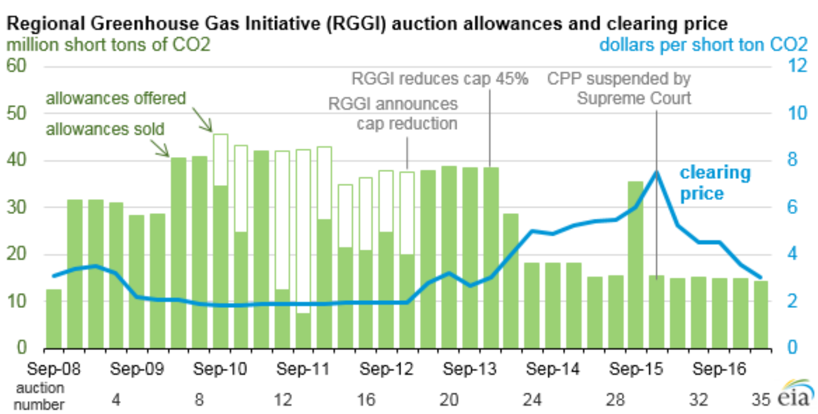



Figure C.2: Price path for carbon credits for RGGI auction



# APPENDIX D

## Proofs for Chapter 2

**Proof of Lemma(2.3.1)** Recall the definition of $\eta$ is as follows:

$$\eta_{i,t} = \begin{cases} v_i - \sum_{l=t+1}^{k_i} \int_{\sigma_l} \tau^{l-t} G_l^{(m_t)}(\eta_{i,l}|\sigma_l)\Big(\eta_{i,l} - \mathbb{E}[(\eta_{j,l}^{(m_t)})_{j\neq i}|\eta_{i,l} > \eta_{j,l}^{(m_t)}]\Big), & \text{if } t\epsilon\{a_i, k_i\} \\ 0, & \text{otherwise} \end{cases}$$

Through recursive addition and subtraction, it is easy to see that the above is equivalent to the following :

Replacing probability of bid with probability of pseudo type we can rewrite the payoff functions as

$$W(\eta_{i,t}) = \int_{\sigma_t} \tau G_t^{(m_t)}(\eta_{i,t}|\sigma_t)\Big(v_i - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big)$$
$$+ (1 - G_t^{(m_t)}(\eta_{i,t}|\sigma_t))W(\eta_{i,t+1})$$
$$V(\eta_{i,t}) = G_t^{(m_t)}(\eta_{i,t}|\sigma_t)\Big(v_i - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big) + (1 - G_t^{(m_t)}(\eta_{i,t}|\sigma_t))W(\eta_{i,t+1})$$

This can be rewritten as

$$W(\eta_{i,t}) = \int_{\sigma_t} \tau G_t^{(m_t)}(\eta_{i,t}|\sigma_t)\Big(v_i - W(\eta_{i,t+1}) - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big) + W(\eta_{i,t+1}) \tag{D.1}$$

$$V(\eta_{i,t}) = G_t^{(m_t)}(\eta_{i,t}|\sigma_t)\Big(v_i - W(\eta_{i,t+1}) - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big) + W(\eta_{i,t+1}) \tag{D.2}$$

Now, we rewrite the pseudo type in terms of the payoff function. Using addition and subtraction and using the definition of $\eta$ we can rewrite the definition of $\eta$ as :

$$\eta_{i,t} =$$
$$v_i - \int_{\sigma_{t+1}} \tau G_t^{(m_t)}(\eta_{i,t}|\sigma_l)\Big(v_i - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big) +$$
$$\sum_{l=t+2}^{k_i}\int_{\sigma_l}\Big(\prod_{q=1}^{l-(t+1)}(1-\int_{\sigma_q} G_q^{mq}(\eta_{i,q}))\Big)\tau^{l-t}G_l^{(m_l)}(\eta_{i,l}|\sigma_l)\Big(v_i - \mathbb{E}[(\eta_{j,l}^{(m_l)})_{j\neq i}|\eta_{i,l} > \eta_{j,l}^{(m_l)}]\Big)$$

, if $t \geq a_i$ or $t \leq k_i$

Note that the second term in the equation above is equal to $W(\eta_{i,t+1})$. Thus we have

$$\eta_{i,t} = \begin{cases} v_i - W(\eta_{i,t+1}), & \text{if } t \geq a_i \text{ or } t \leq k_i \\ 0, & \text{otherwise} \end{cases}$$

Thus, equation(D.1) and (D.2) as

$$W(\eta_{i,t}) = \int_{\sigma_t} \tau G_t^{(m_t)}(\eta_{i,t}|\sigma_t)\Big(\eta_{i,t} - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big) + W(\eta_{i,t+1}) \tag{D.3}$$

$$V(\eta_{i,t}) = G_t^{(m_t)}(\eta_{i,t}|\sigma_t)\Big(\eta_{i,t} - \mathbb{E}[(\eta_{j,t}^{(m_t)})_{j\neq i}|\eta_{i,t} > \eta_{j,t}^{(m_t)}]\Big) + W(\eta_{i,t+1}) \tag{D.4}$$

■

**Proof of theorem(3.8)**

The symmetric Bayesian Nash equilibrium bid in period $t$ maximizes the following payoff of bidder $t$:

$$V(\eta_{i,t}) = \Big\{G_l^{(m_t)}(\eta_{i,t}|\sigma_l)\mathbb{E}\Big[n_{i,t} - b^{(m_t)}(n_{i,t})\Big|b(n_{i,t}) > b_{j\neq i}^{(m_t)}\Big] + W(\eta_{i,t+1})\Big\}$$

Notice that $W(\eta_{i,t})$ in the above expression is merely an additive constant. we will use this and backward induction to solve for the equilibrium bidding function.

First from the structure of $V(\eta_{i,t}|\sigma_t)$ , it is clear that after the last active period $k_i$, the buyer's equilibrium bid will be equal to zero, i.e. $b_t^i = 0 \;\forall\; t > k_i$. This is because the buyer is only active till period $k_i$ and would earn a negative profit from winning if he is active after the actual deadline. Thus we can rewrite the equilibrium bidding strategy as a set of finite bids given as $b^i = \{b(\eta_{i,1}), b(\eta_{i,2}), ....b(\eta_{i,k_i})\}$. Now we will show that in the last active round of bidder $i$'s lifetime i.e. $k_i^{th}$ period, bidder bids their pseudo valuation, so $b(\eta_{i,k_i}) = \eta_{i,k_i}$. Note that in the last period pseudo type is equal to value of the bidder, i.e. $\eta_{i,k_i} = v_i$

- If $b' < v_i$.

  In cases where the price for the object is in-between $b(v_i, k_i)$ and $v_i$, i.e. $b(v_i, k_i) < z_{t_{k_i}} < v_i$, the current period discounted utility from winning is positive i.e $(v_i - z_{t_{k_i}}) > 0$ but the buyer does not win. Thus this is not optimal

- If $b' > v_i$.

  In cases where the price for the object is in-between $b(v_i, k_i)$ and $v_i$, i.e. $b(v_i, k_i) > z_{t_{k_i}} > v_i$, the current period discounted utility from winning is negative i.e $(v_i - z_{t_{k_i}}) < 0$. Thus this is not optimal.

From above we get that any other bid than $b(\eta_{i,k_i}) = v_i = \eta_{i,k_i}$ would decreases buyers payoff. Thus $b(v_i, k_i) = \eta_{i,k_i}$ is an optimal bid in the last active period ($k_i^{th}$ period) for bidder $i$.

Next we prove reporting bid equal to $b(\eta_{i,t})$ is optimal in an arbitrary $t$ during the active demand period, i.e., $a_i < t \ k_i$ , assuming it is optimal in all period after $t$. Recall that equilibrium bid maximizes $V(\eta_{i,t})$

$$V(\eta_{i,t}) = \left\{ G_l^{(m_t)}(\eta_{i,t}|\sigma_l)\mathbb{E}\left[ n_{i,t} - b^{(m_t)}(n_{i,t}) \middle| b(n_{i,t}) > b_{j\neq i}^{(m_t)} \right] + W(\eta_{i,t+1}) \right\}$$

Here the first term represents the expected current period discounted utility and the second term represents the expected utility from future periods if he loses the current period auction . Notice that the second term is independent of the bid in period $t_r$. Thus, this is equivalent to the bid maximizing the first term.

Note that $\eta_{i,t} = v_i - W(\eta_{i,t+1})$ represents the adjusted value for bidder $i$ in period $t$.We will now show $b(\eta_{i,t}) = \eta_{i,t}$ maximizes eqn(6). Consider any arbitrary $b' \neq b(\eta_{i,t})$.

- If $b' < b(\eta_{i,t})$

  In cases where the price for the object is in-between $b'$ and $b(\eta_{i,t})$, i.e. $b' < z_t < b(\eta_{i,t})$, the current period discounted utility from winning is positive i.e $v_i - W(\eta_{i,t+1}) - z_{t_r}$ $> 0$ but the buyer does not win. Thus this is not optimal.

- If $b' > b(\eta_{i,t})$

  In cases where the price for the object is in-between $b'$ and $b(\eta_{i,t})$, i.e. $b' > z_t > b(\eta_{i,t})$, the current period discounted utility from winning is negative i.e $v_i - W(\eta_{i,t}) - z_t < 0$. Thus this is not optimal.

Which gives the optimal bidding strategy as, $b(\eta_{i,t}) = v_i - W(\eta_{i,t+1})$. ■

**Proof for proposition(2.4.1)** First we look at how the rate of future supply effects the equilibrium payoff of the bidder . The period $t$ bid maximizes the following

$$b(\hat{\eta})\epsilon \operatorname*{Arg\,max}_{\hat{\eta}} G^{m_t}(\hat{\eta})\mathbb{E}\left[v_i - b^{(m_t)}(\eta_{j,t})_{j\neq i}\middle| b(\hat{\eta}) > b^{(m_t)}(\eta_{j,t})\right] + (1 - G^{m_t}(\hat{\eta}))W(\eta_{i,t+1}|\lambda) \tag{D.5}$$

Note that the bid is made after the supply and demand realization, so the supply rate $\lambda$ only affects payoff from future periods. Also, $W(\eta_{i,t+1})$ is dependent on the supply and demand distributions as well as $\lambda$. Usually we suppress this dependence for easier notation Here we reintroduce its dependence on $\lambda$ as it is critical here. Using integral-form Envelop theorem we get :

$$\frac{\delta(V(\eta_{i,t}|\sigma_t))}{\delta(\lambda)} = \left(1 - G_t^{(m_t)}[\eta_{i,t}|\sigma_t]\right)\frac{\delta(W(\eta_{i,t+1}|\lambda))}{\delta(\lambda)} \tag{D.6}$$

Now we use equation(8) and backward induction to derive the derivative

Let us derive this in the last period of active demand. Recall that the equilibrium bid in last period will be $b_i(\eta(v_i, r_i)) = \eta(v_i, 1) = v_i$ and $W(\eta_i(v_i, 0) = 0$, thus

$$\frac{\delta(W(\eta_i(v_i, 1)))}{\delta(\lambda)} \geq 0$$

Now let us assume for any arbitrary $r'$ period of active demand left, i.e. assume $\frac{\delta(W(\eta_{i,t'})}{\delta(\lambda)} \geq 0$ , we show that this hold in $r'+1$ too. Rewriting equation (2.4) for $t'_r+1$

$$\frac{\delta(W(\eta_i(v_i, t'_r+1)))}{\delta(\lambda)}$$

$$= \frac{\delta \int_{\sigma_{t'_r+1}} \tau G^{\lambda m_{t'_r+1}}(\eta_{i,t'_r+1})\Big(\eta_{i,t'_r+1} - \mathbb{E}[(\eta^m_{j,t'_r+1})_{j\neq i}|\eta_{i,t'_r+1} > \eta^m_{j,t'_r+1}]\Big) + W(\eta_{i,t'_r+1})}{\delta\lambda}$$

Using envelop theorem and $t'_r+1$ maximization equation we get

$$= \int_{\sigma_{t'_r+1}} \tau \frac{\delta G^{\lambda m_{t'_r+1}}(\eta_{i,t'_r+1})}{\delta\lambda}\Big(\eta_{i,t'_r+1} - \mathbb{E}[(\eta^m_{j,t'_r+1})_{j\neq i}|\eta_{i,t'_r+1} > \eta^1_{j,t'_r+1}]\Big)$$

$$- G^{\lambda m_{t'_r+1}}(\eta_{i,t'_r+1})\frac{\delta\mathbb{E}[(\eta^m_{j,t'_r+1})_{j\neq i}|\eta_{i,t'_r+1} > \eta^m_{j,t'_r+1}]\Big)}{\delta\lambda} + \frac{\delta W(\eta_{i,t'_r+1})}{\delta\lambda}$$

using the assumption $\frac{\delta(W(\eta_{i,t'})}{\delta(\lambda)} > 0$, we get

$$\geq 0$$

Thus, $\frac{\delta(W(\eta_{i,t})}{\delta(\lambda)} \geq 0 \ \forall t$ ■

**Proof for proposition(2.4.2)** first we show bid is decreasing in supply rate:

$$\frac{\delta(b(\eta_{i,t})}{\delta(\lambda)} = \frac{\delta(v_i - W(\eta_{i,t+1}|\lambda))}{\delta\lambda} \tag{D.7}$$

$$= -\frac{\delta(W(\eta_{i,t+1}|\lambda))}{\delta\lambda} \tag{D.8}$$

$$\leq \text{as a resutl of proposition(2.4.1)} \tag{D.9}$$

Notice that this also implies $\frac{\delta(\eta_{i,t})}{\delta(\lambda)} \leq 0$ , thus we have

$$\mathbb{E}(P_t) = \mathbb{E}\Big(\eta_t^{(\lambda * m_t)}\Big)$$

$$\rightarrow \frac{\delta\mathbb{E}(P_t)}{\delta(\lambda)} = \frac{\delta(\mathbb{E}[\eta_t^{(\lambda * m_t)}]}{\delta(\lambda * m_t)} m_t + \frac{\delta(\mathbb{E}[\eta_t^{(\lambda * m_t)}]}{\delta(\eta_t)}\frac{\delta(\eta_t)}{\delta(\lambda)} \leq 0$$

■

**Proof for proposition(2.4.3)** First we look at how the rate of future supply effects the equilibrium payoff of the bidder . The period $t$ bid maximizes the following

$$b(\hat{\eta})\epsilon \operatorname*{Arg\,max}_{\hat{\eta}} G^{m_t}(\hat{\eta})\mathbb{E}\left[\tau v_i - b^{(m_t)}(\eta_{j,t})_{j\neq i}\middle| b(\hat{\eta}) > b^{(m_t)}(\eta_{j,t})|\tau\right] + (1-G^{m_t}(\hat{\eta}))W(\eta_{i,t+1}|\tau) \tag{D.10}$$

Note that the bid is made after the supply and demand realization, so the supply rate $\tau$ only affects payoff from future periods. Also, $W(\eta_{i,t+1})$ is dependent on the supply and demand distributions as well as $\tau$. Usually we suppress this dependence for easier notation Here we reintroduce its dependence on $\tau$ as it is critical here. Using integral-form Envelop theorem we get :

$$\frac{\delta(V(\eta_{i,t}|\sigma_t))}{\delta(\tau)} = V(\eta_{i,t}|\sigma_t) > 0 \tag{D.11}$$

■

**proof of proposition(2.4.4)** first we show bid is decreasing in uncertainty:

$$\frac{\delta(b(\eta_{i,t})}{\delta(\tau)} = \frac{\delta(\tau v_i - W(\eta_{i,t+1}|\tau))}{\delta\tau} \tag{D.12}$$

$$v_i - \frac{\delta(W(\eta_{i,t+1}|\tau))}{\delta\tau} \tag{D.13}$$

$$\geq 0 \tag{D.14}$$

Notice that this also implies $\frac{\delta(\eta_{i,t})}{\delta(\tau)} \geq 0$ , thus we have

$$\mathbb{E}(P_t) = \mathbb{E}\left(\eta_t^{(m_t)}\right)$$

$$\rightarrow \frac{\delta\mathbb{E}(P_t)}{\delta(\tau)} = \frac{\delta(\mathbb{E}[\eta_t^{(m_t)}]}{\delta(\eta_t)}\frac{\delta(\eta_t)}{\delta(\tau)} \geq 0$$

■

**Proof of proposition(2.4.5)**

$$\frac{d\mathbb{E}(P_t|\lambda_{new})}{d(\tau)} = \frac{\delta\mathbb{E}(P_t|\lambda_{new})}{\delta(\lambda_{new})}\frac{\delta(\lambda_{new})}{\delta\tau} + \frac{\delta\mathbb{E}(P_t|\lambda_{new})}{\delta(\tau)}$$

as first term is positive we have

$$> \frac{\delta\mathbb{E}(P_t|\lambda_{new})}{\delta(\tau)}$$

$$> \frac{d\mathbb{E}(P_t|\lambda)}{d(\tau)}$$

■

# APPENDIX E

## Figures for Chapter 3

Figure E.1: Sample of Yahoo! search result page

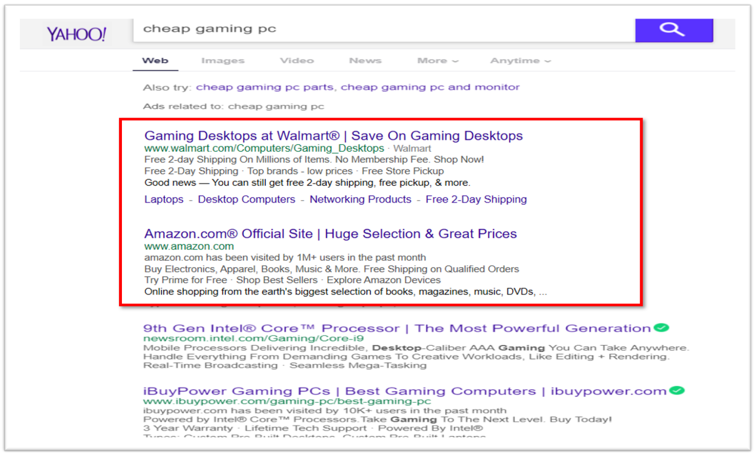
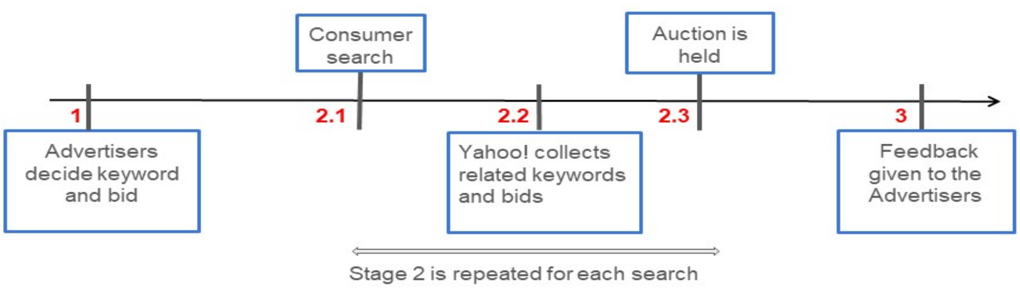
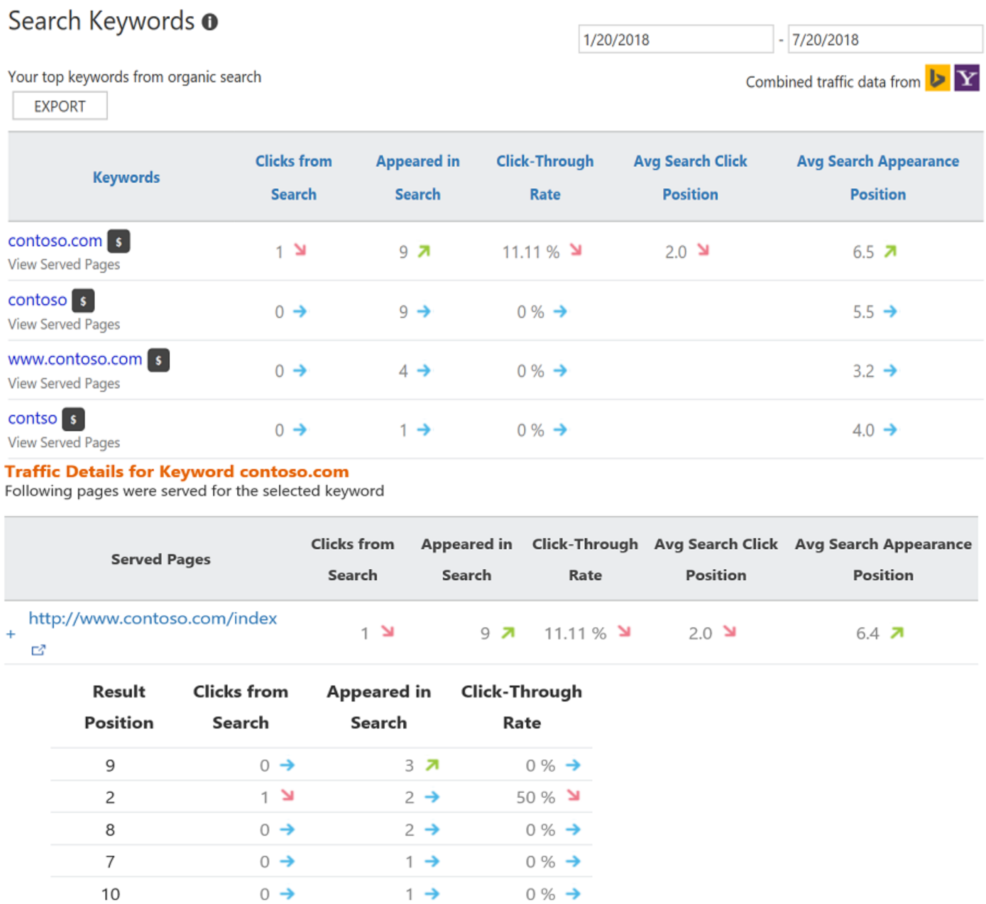



Figure E.2: Timeline




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.

Figure E.3: Snapshot of an advertiser's account on Yahoo!



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**

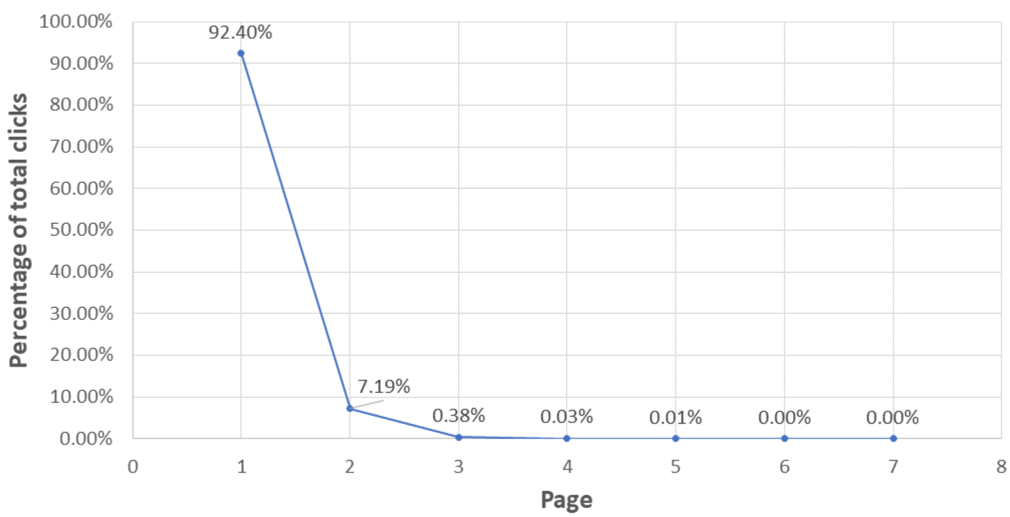
Following pages were served for the selected keyword

| 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 % → |

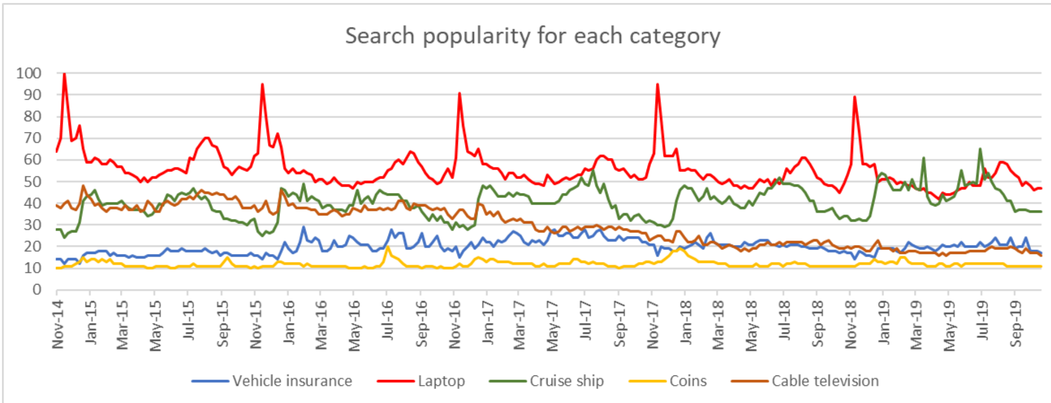
The figure shows a sample snapshot of information available to advertisers. The first one shows aggregated data for each keyword. The second one shows the details for each keyword. Click through rate refers to the ratio of the number of clicks times the number of displays. Refer to this link for more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

Figure E.4: Percentage of clicks across pages



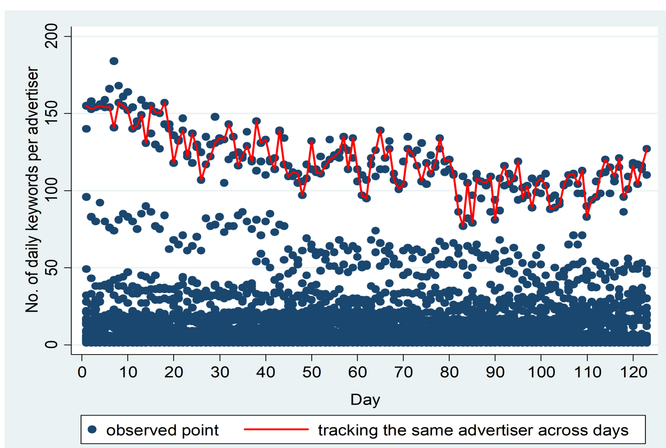

The graph plots the percentage of the clicks received across different pages. This also shows why the paper restricts the analysis to first page. As we need to model consumers click behavior limiting the ads to first position captures most of the click decision. The later pages will show less variation in the consumer's click behavior

Figure E.5: The popularity of the product categories for google in the past five years



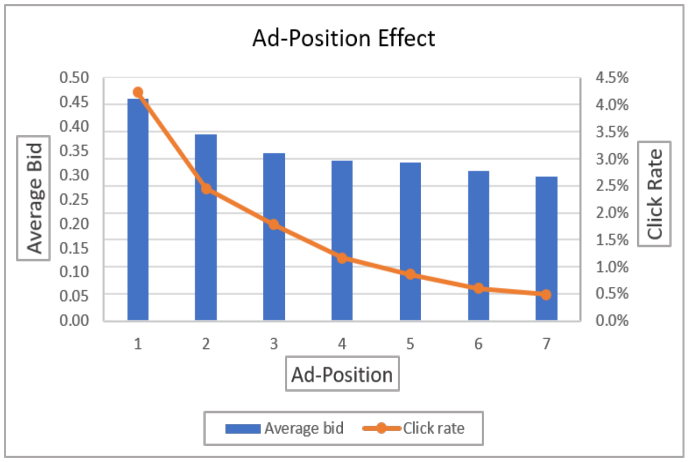

The graph plots the popularity of the categories for google in the past five years. The y-axis numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term. The average popularity score across last five years for laptop, cable, car insurance, cruise, coins are 55, 30, 20, 40, 32 respectively. Source: Google Search Analytics click **here** for details

Figure E.6: No. of keywords specified by advertisers across days



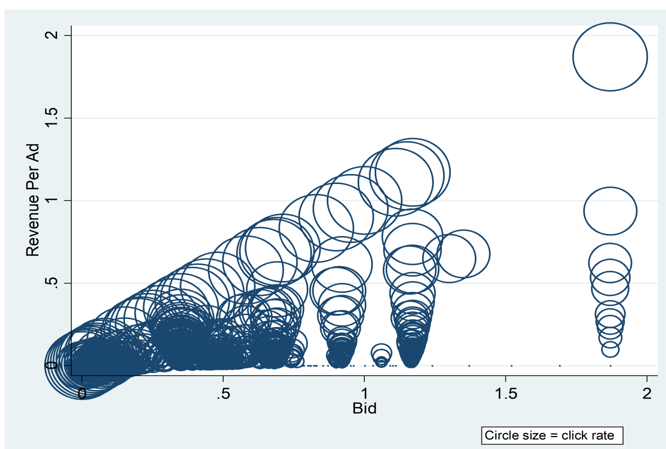

The graph shows a scatter plot of the number of keywords specified per advertiser across days. The red line tracks the change in number of keywords specified by the same advertiser. This variation shows that the advertiser is using keywords to capture difference in ad environment across searches, this variation is used to see how the advertiser bids differently across keywords to accommodate the effect of other ads on the page.

Figure E.7: Average click rate and bid across ad positions



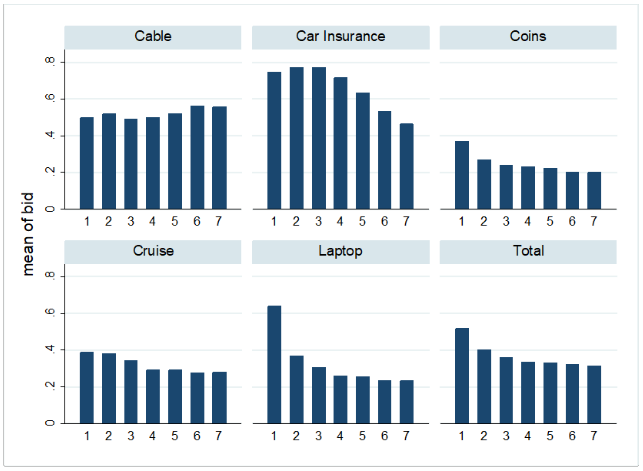

Click rate is the ratio of number of clicks over the number of displays. The car insurance bid is scaled by $1/10^{th}$ of original value for descriptive analysis only.

Figure E.8: Revenue per ad across bid



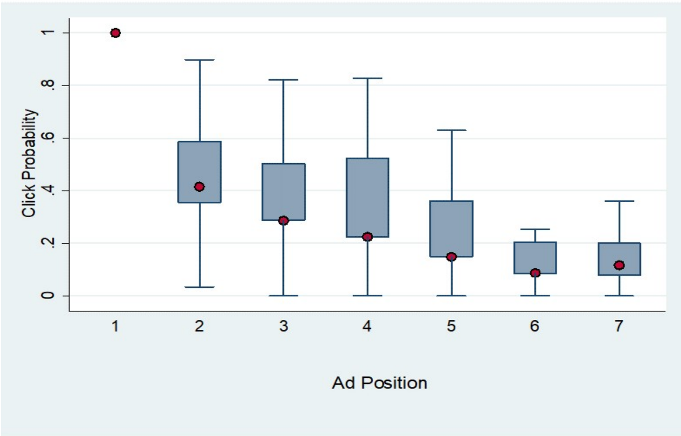

The graph shows the revenue per bid, where the size of the circle shows the click rate of the ad (larger circle signifying higher click rate). This graph motivates the decision of the search engine to weight the bid by the advertiser's effect on click rate. Note: the revenue per ad is equal to price per click multiplied by the quality score. As the price is not given, this graph plots the bid instead of price, which gives the upper bound on the revenue. The results are for cruise category.

Figure E.9: Mean of bids at different winning positions



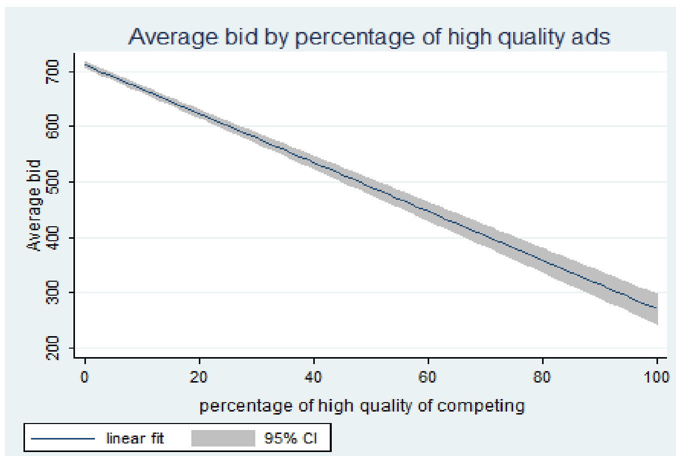

The bid is measured in cents. Note that as the auction mechanism decides the allocation of the ad position according to the weighted bid, it is not necessary for a higher bid to get a higher position. Due to the large difference in bids for car-insurance and other categories, I have scaled the bid in car-insurance to 1/10th of a cent. (This is only done for this graph and not in the data)

Figure E.10: Mean click probability for each ad across position



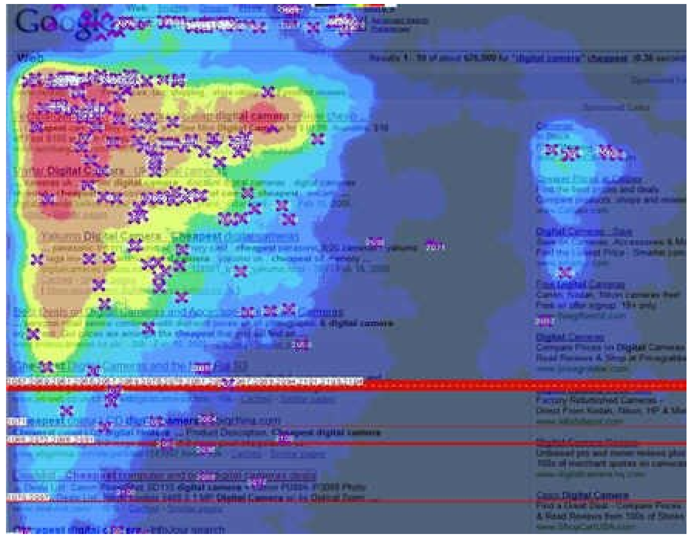

The plot shows the relative click of a single ad in different positions. The click probability is measured relative to the clicks the ad received in the first position.

Figure E.11: Relation between the advertiser's bid and percentage of other high-quality ads on the page



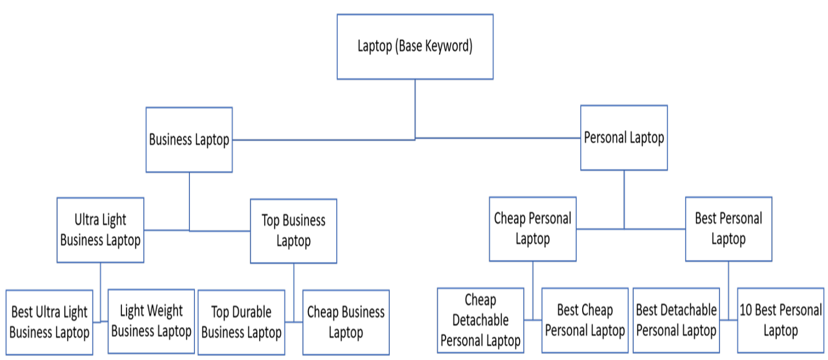

The plot graphs linear regression of bid on the presence of other high-quality ads on the page. The plot shows preliminary evidence that the bid decreases with the increase in other high quality ads on the page. The regression includes controls for auction and consumer heterogeneity. The quality is measured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

Figure E.12: Eye tracking Pattern



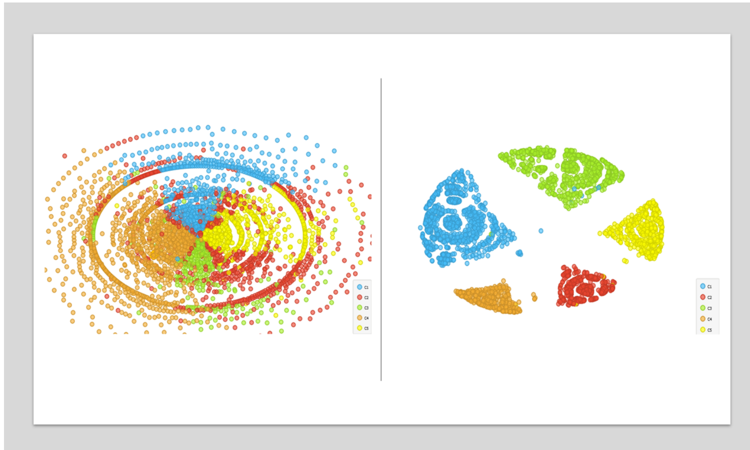
This figure plot the results of a eye tracking research. It shows the eye movement of the consumer on the page. The red region is the on scanned the most followed by yellow and then blue. The picture reconfirms that consumers do a top to bottom scan. Thus, making ads on higher position more valuable.

Figure E.13: Example of sub category within the laptop category



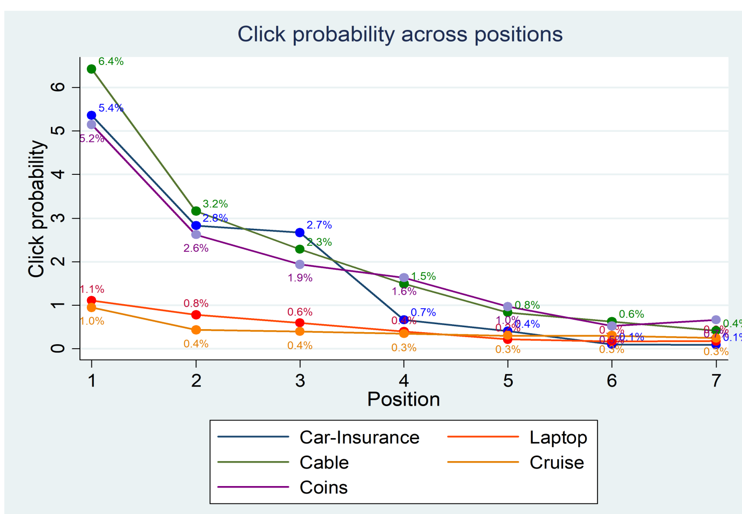

The plot shows an example of what a sub-category would look within a major category such as laptop. I use a machine learning clustering algorithm is to create such sub-groups.

Figure E.14: Keywords for Cosine and Euclidean distance



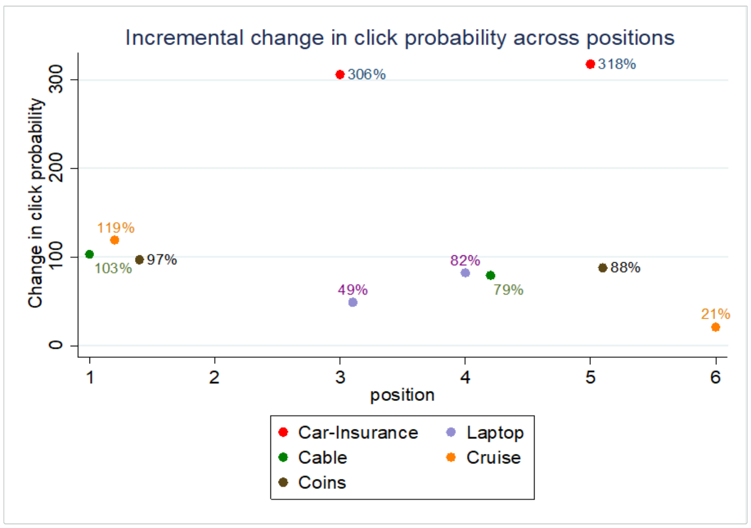
The plot shows keywords in the five main categories. It shows the relative performance of using different distance measure. the one on the left uses Euclidean distance and the one on the right uses Cosine distance. For the analysis cosine distance is used.

Figure E.15: The predicted click probability across positions



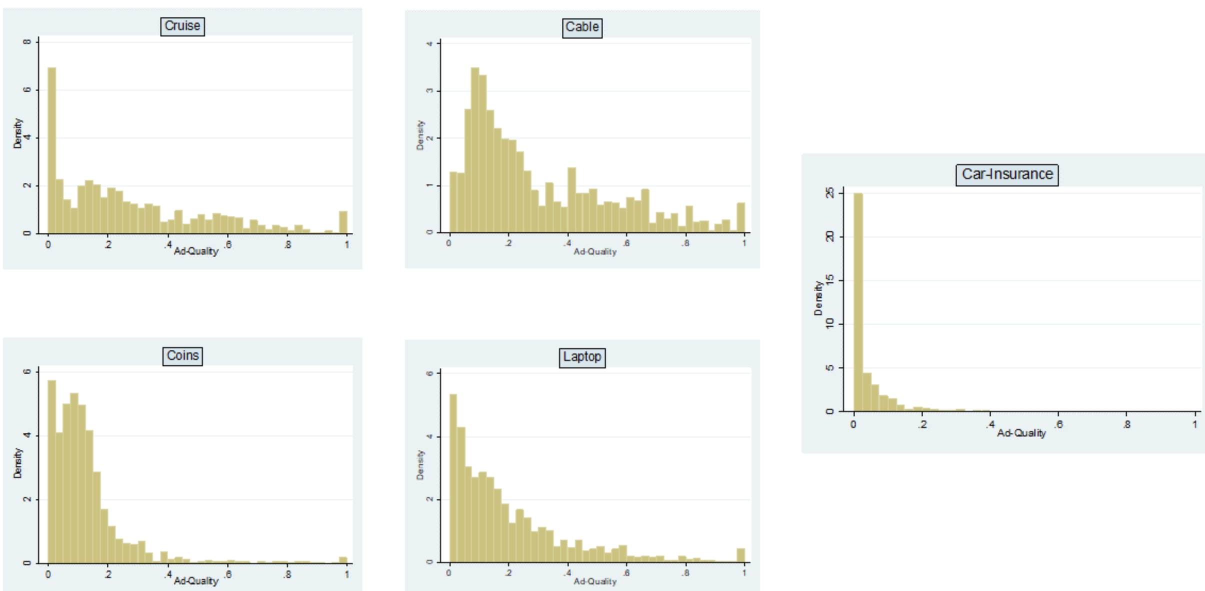

The plot shows the predicted click probability across positions. The click probability is the predicted value from the weighted logit model in step 1 of the estimation method. The values are shown in percentage.

Figure E.16: The affect of choosing higher position on click probability




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 $1^{st}$. The graph plots the two highest jump for each category.

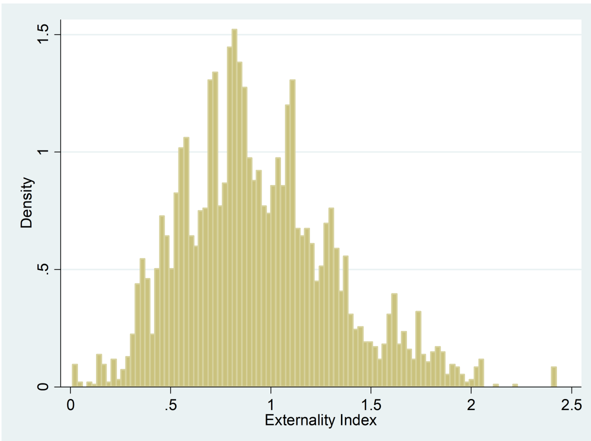
Figure E.17: Ad-Quality



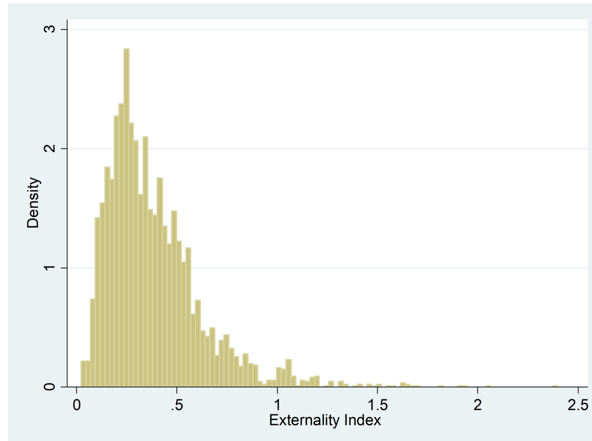

The plot shows distribution of ad quality across the five categories. The quality score is measured as the advertiser's effect on click rate. It is estimated on the consumer side.

Figure E.18: Cumulative distribution of Externality Index

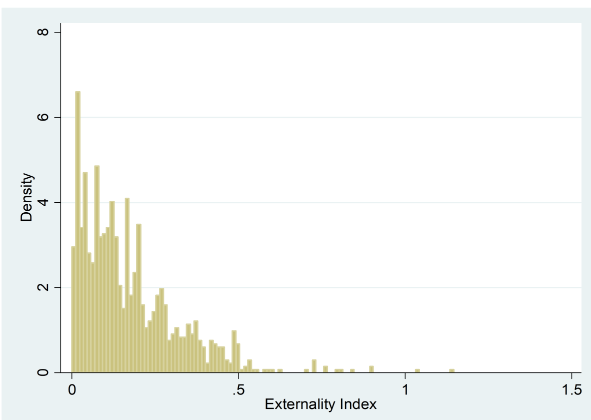
(a) Cable TV



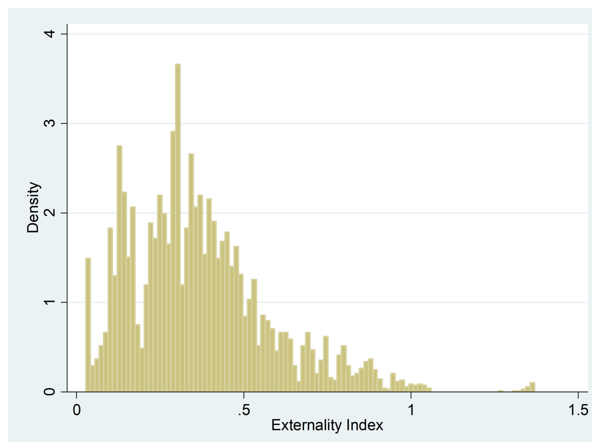

(b) Cruise



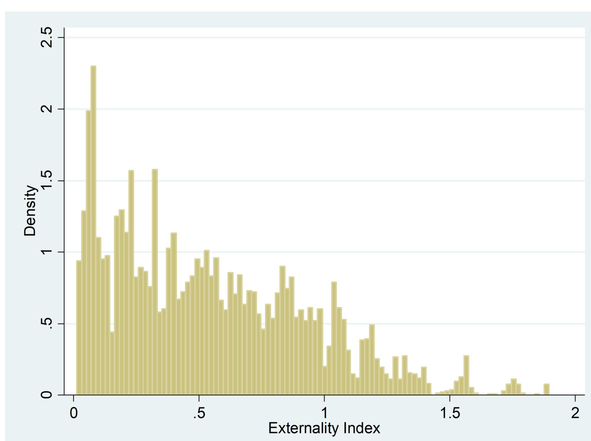

(c) Car Insurance




(d) Coins

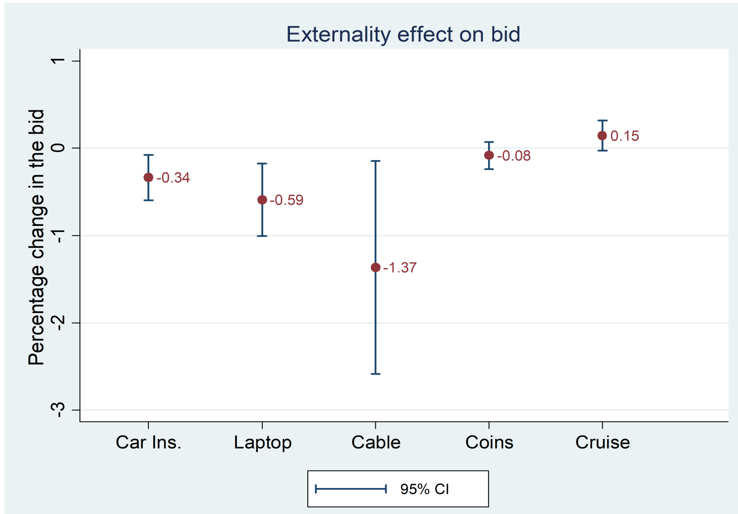



(e) Laptop



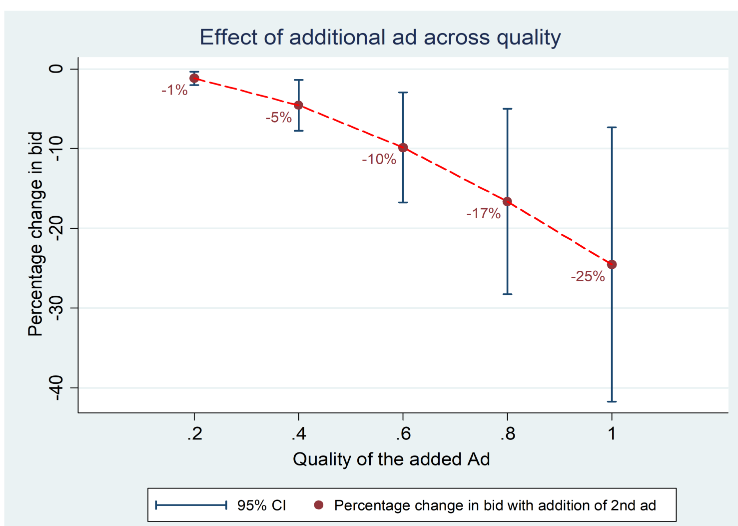

Each graph plots the predicted externality index for the five ad categories. Externality index is measured as a weighted sum of number of other ads where the weight is a function of the quality score of the ad.

Figure E.19: Externality coefficient across categories



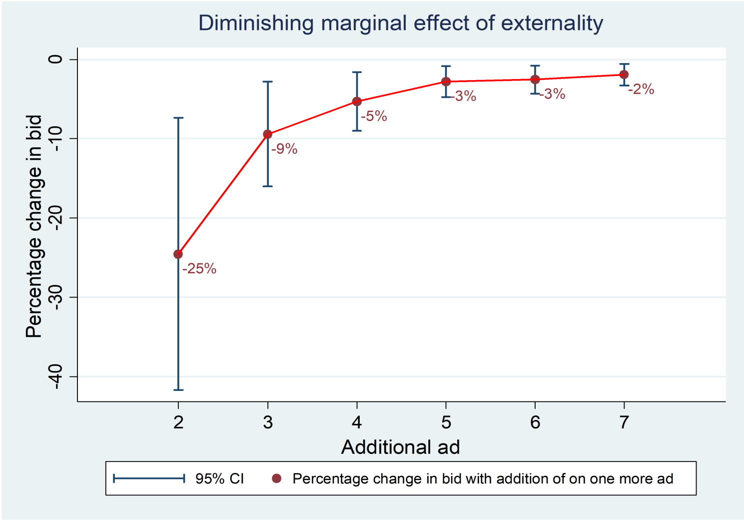

In this graph, I plot the co-efficient from 2SLS log-log regression of log of bid on log of externality. The coefficient is interpreted as the percentage increase in the bid when externality increases by 1%.

Figure E.20: Effect of an additional ad on advertisers bid across different ad quality



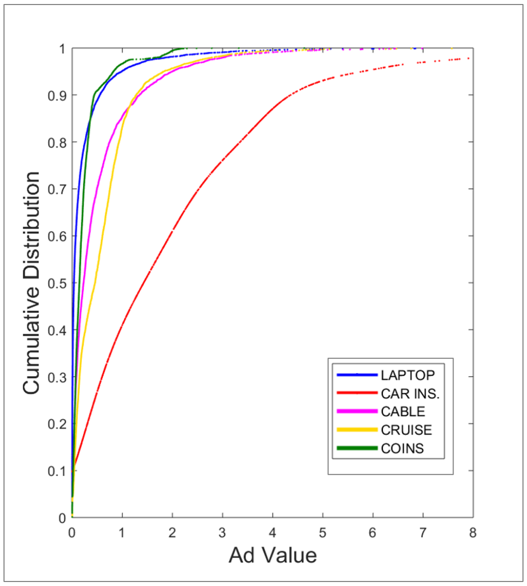

Notes: In this graph, I plot the percentage decrease in a bid for each additional ad. The x-axis shows the quality of the added ad. The graph plots result in the laptop product category. Similar results hold for others as well.

Figure E.21: Diminishing effect of an additional ad on advertisers bid




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

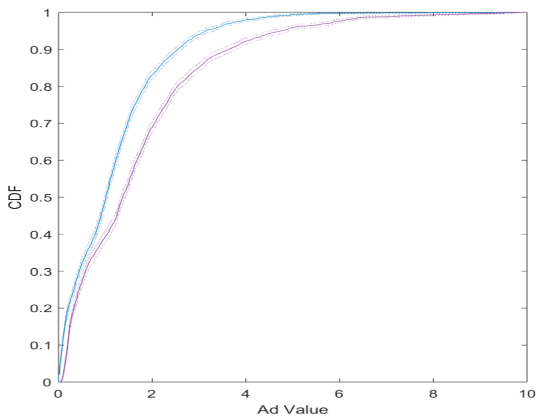
Figure E.22: Upper bound for the cumulative distribution function of the ad value



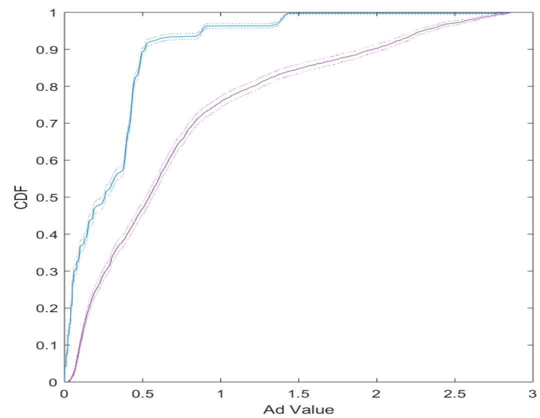

The plot shows the upper bound of the estimated distribution of advertisers' ad value. The distribution of advertisers value for an ad differs across product category. Thus, estiamting the distribution empirically is critical for getting realistic counterfactual results.

Figure E.23: Cumulative distribution of advertiser's ad value

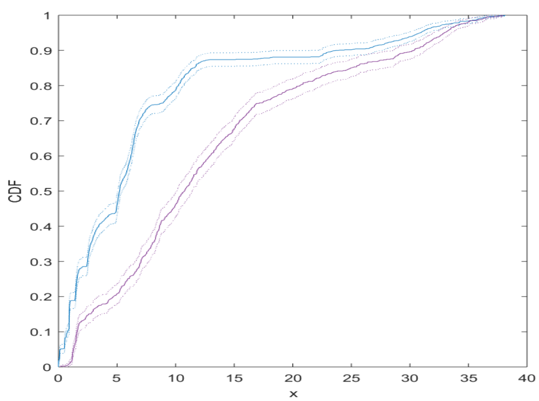
(a) Cable TV



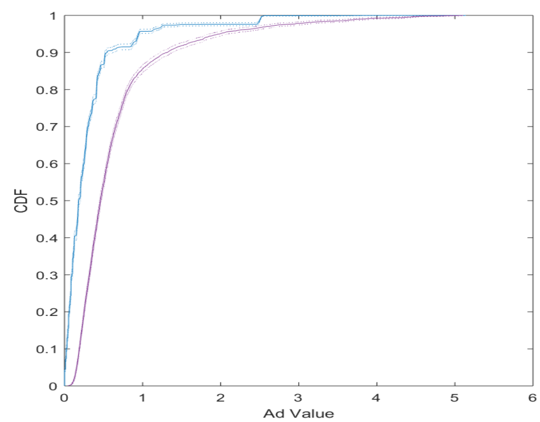

(b) Cruise



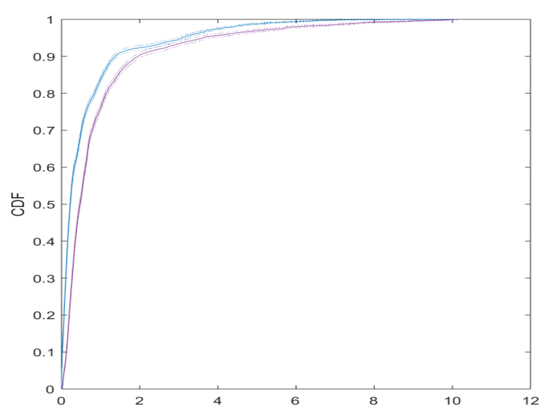

(c) Car Insurance




(d) Coins




(e) Laptop




Notes: This graph plots the upper and lower bound estimated for the ad-value distribution. The x-axis plots the values and y-axis shows the corresponding cumulative distribution at each point. Here ad value captures the advertisers' externality free value that can be thought of as their value for an exclusive ad

Figure E.24: Gain in revenue from using the alternate method in appendix(H)

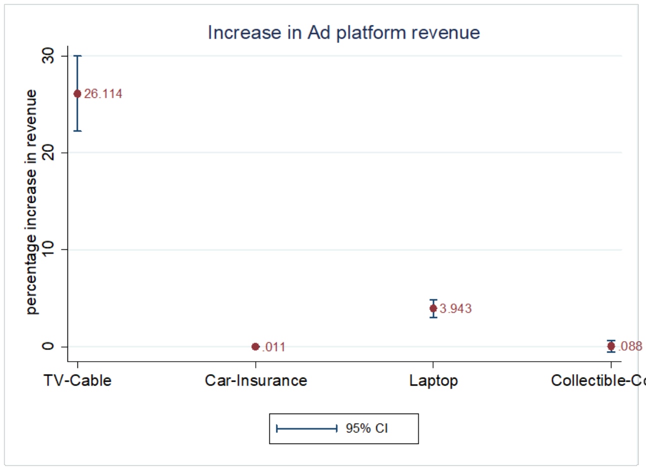
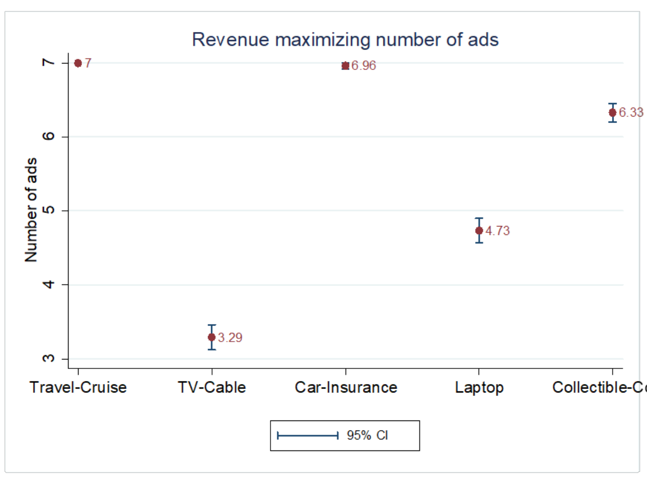



Figure E.25: Revenue-maximizing number of ads from alternate method suggested in appendix(H)




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.

## APPENDIX F

### Tables for Chapter 3

Table F.1: List of variables available for consumers side data

| Variable | Description |
|---|---|
| Day | The day of the month. |
| Advertiser ID | The id for each advertiser. |
| Clicks | The number of clicks received by a advertisement. |
| Ad Displays | The number of times an advertisement was displayed. |
| Keyword | The keyword gives the specified words matched between search and the ad. |
| Keylength | The number of words specified in the keyword. |
| Ad Position | The winning position of the advertisement on the search result page. |

Ad definition: id-keywords combination for the same bid gives the advertisement

Table F.2: List of variables available on advertisers side

| Variable | Description |
|---|---|
| Day | The day of the month. |
| Advertiser ID | The id for each advertiser. |
| Bid | The bid used for an ad. |
| Ad Displays | The number of times an advertisement was displayed. |
| Keyword | The keyword gives the specified words matched between search and the ad. |
| Keylength | The number of words specified in the keyword. |
| Ad Position | The winning position of the advertisement on the search result page. |

Ad definition: id-keywords combination for the same bid gives the advertisement

Table F.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 | - |

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 F.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.
Advertiser side data:Aggregated for each day-ad position observation. Total number of observations: 21,599. Restricting data to ads on first page.

Table F.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 |
| Collectible Coins | 1.7% | .174 | 23 | 103.52 |

Table reports mean value of variables across the five product categories. The bid is measured in terms of cents

Table F.6: 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 advertisers | X | X | X | X | X |
| Sub-category | X | X | X | X | X |
| No. of ads | X | X | X | X | X |
| N | 729 | 6977 | 3219 | 3589 | 5041 |

Table F.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 |

The table summaries statistics of the predicted quality score.

Table F.8: 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) |

The table shows the average partial effect on click probability relative to 1st position as calculated in the step 1 of consumer's click choice model. Fixed effects for market and advertiser included.

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

| | Predicted no. of ads per page | | Predicted Change in profit | | Externality effect | Avg Quality |
|---|---|---|---|---|---|---|
| | lower bound | upper bound | lower bound | upper bound | | |
| Car Insurance | 5 | 3 | 0.6942 % (0.694%, .694%) | 1.698% (1.697%, 1.699%) | Medium | Low |
| Laptop | 5 | 4 | 1.014 % (1.014%, 1.015%) | 2.985 % (2.985%, 2.987%) | Medium | Medium |
| TV Cable | 3 | 2 | 15.575 % (15.574%, 15.577%) | 22.7 % (22.697%, 22.704%) | High | High |
| Coins | 7 | 7 | 0 | 0 | Low | Low |
| Cruise | 7 | 7 | 0 | 0 | Nill | Medium |
| Simulation | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

Please refer to section(3.9) 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.

Table F.10: Identifying different categories

| **Variable** | **Description** | **Clicks** | **Bid** | **No. of adv.** | **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

Table F.11: Robust check: externality co-efficient across categories

| | Car Ins | Laptop | Cable TV | Coins | Cruise |
|---|---|---|---|---|---|
| Externality Index(1) | -.1714833 | -.2950749 | -.6921709 | -.0419864 | .0496461 |
| | ( .072278 ) | (.1050882 ) | ( .319323 ) | (.0399867 ) | ( .0603489) |
| Externality Index(2) | -.3858009 | -.6256772 | -1.391457 | -1.385585 | .1321744 |
| | ( .1372449 ) | ( .2365638 ) | ( .5827886 ) | ( 1.772531 ) | ( .0772414 ) |
| 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 |

The table shows the coefficient of the externality index in step 2 of the estimation of 2SLS regression of log of bid on externality

Table F.12: Robust check: different externality index definitions

| | Car Ins | Laptop | Cable TV | Coins | Cruise |
|---|---|---|---|---|---|
| top 10% | -.3362724 | -1.541917 | -.8776404 | -.1130883 | .1793007 |
| | ( .1330385 ) | (.6485241 ) | (.2510247 ) | ( .0768405 ) | ( .1057224 ) |
| above your quality | -.5183959 | -.2325425 | -.09395 | -1.385585 | -2.85473 |
| | ( .579788 ) | (0.341) | ( .3748843 ) | ( 1.772531 ) | (28.34121 ) |
| 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 |

The table shows the coefficient of the externality index in step 2 of the estimation of 2SLS regression of log of bid on externality

# APPENDIX G

## Proofs for Chapter 3

## G.1 Theoretical Proofs

### G.1.1 Proof for proposition(3.4.5)

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

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

Let $b^*$ be the equilibrium bid of the no externality case, then I show that $b^*(EXT_{K_j})^{b_1}$ is the equilibrium bid of externality case. Now in case of externality the ad value is $V_{j,K_j} = v_j(EXT_{K_j})^{b_1}$ instead of $v_j$. Substituting this in the equation above , I get the profit maximizing objective function in the presence of externality.

$$\Pi(b; v_j, s_j) = \max_b \sum_{k=1}^{K} Prob(b_j^w = b^{[k]}) * c_{k,j} \left[ v_j(EXT_{K_j})^{b_1} - \mathbb{E}\Big( P(k) \Big| b_j^w = b^{[k]} \Big) \right] \tag{G.2}$$

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.[1]. Thus, This can be further solved to:

$$\Pi(b; v_j, s_j) = (EXT_{K_j})^{b_1} \left[ \max_b \sum_{k=1}^{K} Prob(b_j^w = b^{[k]}) * c_{k,j} \left[ v_j - \mathbb{E}\Big( P(k) \Big| b_j^w = b^{[k]} \Big) \right] \right]$$

Thus, the solution to the above objective function is $b^* * (EXT_{K_j})^{b_1}$, 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 $\underbrace{b(v_j, s_j; n, \alpha_{K,n})}_{\text{externality free bid}}{}^*(EXT_{K_j})^{b_1}$. Taking log this implies the following:

$$Log(\underbrace{b(v_j, s_j; \alpha_{K,n}, ext)}_{\text{observed bid}}) = Log(\underbrace{b(v_j, s_j; n, \alpha_{K,n})}_{\text{externality free bid}}) + \underbrace{b_1 Log(Ext_{j,K_j})}_{\text{externality effect on bid}} \tag{G.3}$$

■

[1] assuming monotonic bid

### G.1.2 lemma(3.4.1) proof:

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

$$b^{GSP} = Arg\max_{\hat{b}} \Pi(\hat{b}|v_i, s_i) = \max_{\hat{b}} \sum_{k=1}^{K} \gamma_i c_k \left[ v_i - \frac{\mathbb{E}\left(b_w^{GSP,[k+1]} \middle| b_w^{[k]} = \hat{b} * s_i\right)}{s_i} \right] \times Prob(b_w^{GSP,[k+1]} \leq \hat{b} * s_i \leq b_w^{GSP,[k+1]}) \quad (G.4)$$

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 $b^{GSPM}$ for advertiser $i$ is given as :

$$b^{GSPM} = Arg\max_{\tilde{b}_w} \Pi(b_w|v_i, s_i) = \max_{\tilde{b}_w} \sum_{k=1}^{K} \gamma_i c_k \left[ v_i - \frac{\mathbb{E}\left(b_w^{GSPM,[k+1]} \middle| b_w^{[k]} = \tilde{b}_w\right)}{s_i} \right] \times Prob(b_w^{GSP,[k+1]} \leq \hat{b} \leq b_w^{GSP,[k-1]}) \quad (G.5)$$

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

$$b^{GSPM} = Arg\max_{\tilde{b}} \Pi(b_w|v_i, s_i) = \max_{\tilde{b}} \sum_{k=1}^{K} \gamma_i c_k \left[ v_i - \frac{\mathbb{E}\left(b_w^{GSPM,[k+1]} \middle| b_w^{[k]} = \tilde{b}\right)}{s_i} \right] \times Prob(b_w^{GSP,[k+1]} \leq \hat{b} \leq b_w^{GSP,[k-1]}) \quad (G.6)$$

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 :

$$b^{GSPM} = Arg\max_{\check{b}} \Pi(\check{b}|v_i, s_i) = \max_{\check{b}} \sum_{k=1}^{K} \gamma_i c_k \left[ v_i - \frac{\mathbb{E}\left(b_w^{GSPM,[k+1]} \middle| b_w^{[k]} = \check{b} * s_i\right)}{s_i} \right] \times Prob(b_w^{GSPM,[k+1]} \leq \check{b} * s_i \leq b_w^{GSPM,[k-1]}) \quad (G.7)$$

The optimization problem in equation(G.5) and (G.7) are equivalent, Thus, if all other advertisers $j \neq i$ have $b_j^{GSP} = \frac{b_j^{GSPM}}{s_j}$, then advertiser will also have $b_i^{GSP} = \frac{b^c GSPM_i}{s_i}$. 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. ■

### G.1.3 Proof of proposition (3.4.3)

Suppose bidder $i$ bids $\tilde{b}$ and everyone else is playing according to the equilibrium increasing bidding strategy $b(v_j, s_j)$. This equivalently means that bidder $i$ has weighted bid $\tilde{b}_w$ and everyone else bids the equilibrium weighted bid $b_w(w_j)$ (refer to lemma(3.4.1) for more details). Recall that $F_w(.)$ is the distribution of the weighted value $\omega$ 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 :

$$z_k(\tilde{b}_w) = \sum_{k=1}^{K} \frac{\binom{N-1}{}}{\binom{K-1}{}} (1 - G(\tilde{b}_w))^{k-1} (G(\tilde{b}_w))^{N-k} \tag{G.8}$$

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

$$z_k(b_w^{-1}(\tilde{b}_w)) = \sum_{k=1}^{K} \frac{\binom{N-1}{}}{\binom{K-1}{}} (1 - F(b_w^{-1}(\tilde{b}_w)))^{k-1} (F(b_w^{-1}(\tilde{b}_w)))^{N-k} \tag{G.9}$$

$$\tag{G.10}$$

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

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

Substituting value of $z_k(b_w^{-1}(\tilde{b}_w))$ from equation(G.9) and then differentiating we get the above equation we get :

$$\left(v_j - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right) \tag{G.11}$$

$$\times \Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(b_w^{-1}(\tilde{b}_w))^{k-1})F(b_w^{-1}(\tilde{b}_w))^{N-k-1} f(b_w^{-1}(\tilde{b}_w))b^{'}(b_w^{-1}(\tilde{b}_w))$$

$$+ c_k \frac{(N-1)}{(K-1)}(k-1)(1-F(b_w^{-1}(\tilde{b}_w)))^{k-2} F(b_w^{-1}(\tilde{b}_w))^{N-k} f(b_w^{-1}(\tilde{b}_w))b^{'}(b_w^{-1}(\tilde{b}_w))\Big]$$

$$-\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(b_w^{-1}(\tilde{b}_w)))b^{'}(b_w^{-1}(\tilde{b}_w))^{k-1}(F(b_w^{-1}(\tilde{b}_w)))^{N-k} \frac{d\left(\mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)}{d(b)} = 0$$

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

(G.12)

$$\left(v_j - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_j)^{k-1})F(\omega_j)^{N-k-1}\frac{f(\omega_j)}{b^{'}(\omega_j)}$$

$$+ c_k \frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_j)^{k-2})^{N-k} f(\omega_j)b^{'}(\omega_j)\Big]$$

$$-\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_j))b^{'}(\omega_j)^{k-1}(F(\omega_j))^{N-k} \frac{d\left(\mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)}{d(b)} = 0 \tag{G.13}$$

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

$$\mathbb{E}\bigg(\frac{b_w^{[k+1]}}{s_i}\bigg|b_w^{[k]} = \tilde{b}_w\bigg) \tag{G.14}$$
$$= \mathbb{E}(\frac{b_w^{k+1:N}}{s_i}|b^{k+1:N} \leq \tilde{b}_w \leq b^{k-1:N})$$
$$= \mathbb{E}(\frac{b_w^{1:N-k}}{s_i}|b^{1:N-k} \leq \tilde{b}_w)$$
$$= \int_0^{\tilde{b}_w} \frac{x}{s_j} \underbrace{\frac{(N-k)F^{N-k-1}(b^{-1}(x))g(b^{-1}(x))}{F^{N-k}(b^{-1}(\tilde{b}_w))} dx}_{\text{conditional distribution}}$$

Using integration by parts, we get :

$$= \frac{\tilde{b}_w}{s_j} \int_0^{\tilde{b}_w} \frac{(N-k)F^{N-k-1}(b^{-1}(x))f(b^{-1}(x))}{F^{N-k}(b^{-1}(x))} dx - \int_0^{\tilde{b}_w} \frac{F^{N-k}(b^{-1}(x))}{s_j F^{N-k}(b^{-1}(\tilde{b}_w))} dx$$
$$\Rightarrow \mathbb{E}\bigg(\frac{b_w^{[k+1]}}{s_i}\bigg|b_w^{[k]} = \tilde{b}_w\bigg) = \frac{\tilde{b}_w}{s_j} - \Gamma(\tilde{b}_w, f(.)) \tag{G.15}$$

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 $\Gamma(\hat{b}, s_j, g(.))$ which is equal to $\int_0^{\hat{b}s_j} \frac{G^{N-k}(x)}{s_j G^{N-k}(b)} dx$.substituting the expected price in equation(G.13) we get:

$$\left(v_j - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)\frac{f(\omega_j)}{b'(\omega_j)}\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_j)^{k-1})F(\omega_j)^{N-k-1}$$

$$+ c_k \frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_j))^{k-2}F(\omega_j)^{N-k}\Big]$$

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

$$\Rightarrow \left(v_j - \mathbb{E}(\frac{b_w^{[k+1]}}{s_i})\right)\frac{f(\omega_j)}{b'(\omega_j)}\Big[\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(N-k)(1-F(\omega_j)^{k-1})F(\omega_j)^{N-k-1}$$

$$+ c_k \frac{(N-1)}{(K-1)}(k-1)(1-F(\omega_j)^{k-2})F(\omega_j)^{N-k}\Big]$$

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

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

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

using $1 - \frac{d(\Gamma(\tilde{b}_w, f(.)))}{d(b)} < 1$ I get

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

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

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

$$\Rightarrow v_j \leq \frac{b_w}{s_j} \tag{G.16}$$

$$+ \frac{\sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}(1-F(\omega_j))^{k-1}(F(\omega_j))^{N-k}}{s_j \frac{f(\omega_j)}{b'(\omega_j)} \sum_{k=1}^{K} c_k \frac{(N-1)}{(K-1)}\Big[(N-k)(1-F(\omega_j)^{k-1}F(\omega_j)^{N-k-1} + (k-1)(1-F(\omega_j))^{k-2}F(\omega_j)^{N-k}\Big]} \tag{G.17}$$

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 :

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

Thus, equation(G.17) can be written as :

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

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

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

the equilibrium bid will then be :

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

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

# APPENDIX H

## 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. [1]
    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. [2]
- 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.

[1] 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 alternative specification would be the advertisers only report the bid, and the auctioneer applies the pre-specified percentage decrease.

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

### H.0.1 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 ($\beta_1$).

3. Pick the revenue-maximizing number of ads:

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

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

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

### H.0.2 Results

- *The revenue-maximizing number of ads:* The revenue-maximizing number of ads changes with each category. The graph(E.25) 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.

- *Revenue gain across categories:* The new auction design leads to higher revenue for the search engine. Graph(E.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.

# APPENDIX I

## Robustness Check

### I.0.1 Robust checks

- **Definition of Externality index**:

  The externality index captures the effect of the quality as well as the number of other ads on the page. In this paper the relation between the quality and the number of ads is captured by using a weighted sum of number of ads where the structure is similar to a euclidean distance. For convenience the index is reproduced below:

  $$Ext_{j,m} = \sqrt{\sum_{k \neq K} \hat{s}^2_{-j,m}}$$

  The above definition is used as it captures the differences over a single ad having high quality compared to 4-5 ads having lower quality that sum upto the high quality of single ad. Note that a linear sum over the quality other ads would not have been able to capture such difference. The square root is added to normalize the effect of squaring the average quality ($\hat{s}_{-j,m}$). To check the sensitivity of the results to externality index the step was rerun for alternative definition of the externality index as shown in Table(F.11). The alternative definitions are as follows:

  Externality index(1) = $Ext_{j,m} = \sum_{k \neq K} \hat{s}^2_{-j,m}$

  Externality index(2) = $Ext_{j,m} = \sum_{k \neq K} \hat{s}_{-j,m}$

  Overall the table shows that the sign and significance is not affected by the definition.

- **Type of externality:**

  Another assumption made in the main analysis is that the externality is generated from the presence of all the other ads. However the externality can be also be generated from only the high quality ads, or instead only from the ads above you. Table(F.12) presents results for externality when the effect is only from wither the top ten percent of the ads or from ads with quality above the advertiser's quality. The results show that externality is still negative and significant. However, it might be more or less for certain type of ads depending on the ad category. Another interesting observation is that in cruise category the advertisers bid is negatively effected if only the ads with quality above the advertisers ad is considered. This is insightful as for the general externality that captures the effect of all the ads on the page, showed positive and insignificant effect. This can be an indication that for cruise the advertisers are specifically focusing only on the competitors that have a higher or similar quality to theirs. Further analysis on this topic can be done if additional information about the ads is provided such as the similarity of product sold and price charged for the product. As the advertisers names are masked such analysis can not be done with the current data

# APPENDIX J

## Data cleaning for Chapter 3

### J.0.1 Data cleaning

The data set is provided by Yahoo! as part of the Yahoo! Research Alliance Webscope program. It gives details about five different categories, namely laptop, cable, coins, cruise, and car insurance, over 123 days from January 2008 to April 2008. The data has information about keywords, bid, clicks, ad position, and display frequency. The keywords in the data set include one of the base category word 'coin,' 'laptop,' 'cable,' and 'car insurance.' Apart from the base category word, the keywords also include one or more additional words. For example, 'business laptop' and 'student laptop' are two keywords within the base category laptop. The additional words provide a more targeted ad. For instance, an ad 'business laptop' targets consumers that are specifically looking for business usage; however, keyword 'laptop' captures a broader search for any laptop need of a consumer. The maximum number of words in a keyword is 10.

Another key characteristic of this data is the keywords, and the advertiser's id is masked. This means I can track the same advertiser across ads and time; however, the actual identity of the advertiser is masked. More details on identifying the base categories are given in the appendix(J.0.2). To do the analysis, I restrict the dataset to ads on the first page and consider only the first seven ads. The total observations in the raw data set are 207982. After restricting the data to the first page, the data has 131524. The restriction to ads on the first page is important as the consumers click probability drops drastically as the ads go beyond the first page. As seen in fig(E) 92% of the clicks are on first page ads. Additionally, the data is divided into two parts according to the aggregation needed for consumers or advertisers analysis. On the consumer's side, the analysis focuses on the click decision on the consumer. The data provider, Yahoo! Research Lab, aggregated the data for each day-keyword-advertiser-position combination. Let us take an example of observation, for January 1st, 2008, the data reports that the keyword 'business laptop' specified by Walmart that was displayed in the first position got 100 displays and five clicks. For the same keyword and advertiser, i.e. Walmart ad with keyword 'business laptop,' I will have a different observation for the ad displayed in the second position, which got 50 displays and two clicks. This means that the data does not aggregate over different positions and reports results for each position and keyword separately. This is an advantage for the consumer side, as the different keywords can help in capturing the difference in consumer search. Thus, the total observation on the consumer side analysis is 131524.

On the other hand, for analysis on the advertiser's side, I need information on how the advertiser maximized profit for each ad. This means here the relevant variation is the bid over ads and not the keywords. Therefore, on the advertisers side, I aggregate the data on day-ad-advertiser-position combination.[1] This means that when deciding the bid, the advertiser accommodates the expected clicks and price for each ad across different winning positions over a day. An example of an observation is; on January 1st, 2008, the data reported that an ad by Walmart was displayed in the first position, got 200 displays, and ten clicks. The results of the estimation without the aggregation on the consumer side are similar to the one without the aggregation. However, as the model assumes that the advertiser maximizes profit over an ad, the results from aggregation are more aligned with the advertiser's profit-maximizing strategy. Note that the advertiser can always specify a different bid for each keyword, and in this case, each keyword would be treated as a separate ad in this data set. So the aggregation is only on the keywords for which the advertiser has specified the same bid, meaning the bid was maximized over all the keyword that has the same bid per day. The aggregated data gives 21,599 observations.

The estimation uses subgroups in different categories on the basis of similar keywords. For more detail refer to section(3.6). On the consumer side, the analysis needs an estimation

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

of click probability across a position for each market. However, there are some markets that do not observe any clicks and therefore have to be dropped from the analysis. This is less than 1% of the data. Apart from this, there are few advertisers that do not get any click in the data. For these advertisers, I am not able to get the quality score as the quality score is captured through the advertiser's effect on click probability. Therefore, I assume the minimum observed quality score in the data as the imputed quality score for these advertisers. The analysis is robust to assuming less than the minimum quality score, as well. Note that dropping these observations would have introduced bias in the estimation as we need to accommodate all the advertisers in the market. Additionally, there was some market that did not have a consistent observation, meaning they had an observation for lower ad position but not for higher ad position. These were treated as inconsistent with the market rules as a lower ad position is only shown when there are ads at higher ad positions. Therefore these were also dropped from the data.

### J.0.2 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.[2] 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(F.10) gives a summary of how variables differ among categories. Additionally the table(F.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

[2]the categories are identified through specificity of keyword

pay per click.[3] 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 at least 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.[4]

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.

[3] 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

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

# REFERENCES

# REFERENCES


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

Aggarwal, G., Feldman, J., & Muthukrishnan, S. (2006). Bidding to the top: Vcg and equilibria of position-based auctions. In *International workshop on approximation and online algorithms* (pp. 15–28).

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

An industry survey conducted by PwC. (2017). *Iab internet advertising revenue report.*

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, *37*(3), 668–691.

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

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

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

Athey, S., & Segal, I. (2013). An efficient dynamic mechanism. *Econometrica*, *81*(6), 2463–2485.

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

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

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

Bergemann, D., & Said, M. (2010). Dynamic auctions. *Wiley Encyclopedia of Operations Research and Management Science*.

Bolt, W., & Tieman, A. F. (2008). Heavily skewed pricing in two-sided markets. *International Journal of Industrial Organization*, *26*(5), 1250–1255.

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

Brynjolfsson, E., Hui, X., & Liu, M. (2019). Does machine translation affect international trade? evidence from a large digital platform. *Management Science*, *65*(12), 5449–5460.

Burtraw, D., Palmer, K., & Kahn, D. (2010). A symmetric safety valve. *Energy Policy*, *38*(9), 4921–4932.

Caillaud, B., & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *RAND journal of Economics*, 309–328.

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

Crampes, C., Haritchabalet, C., & Jullien, B. (2009). Advertising, competition and entry in media industries. *The Journal of Industrial Economics*, *57*(1), 7–31.

Deb, R., & Pai, M. M. (2013). Ironing in dynamic revenue management: posted prices & biased auctions. In *Proceedings of the twenty-fourth annual acm-siam symposium on discrete algorithms* (pp. 620–631).

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

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

De Smet, D., & Van Cayseele, P. J. (2010). Product differentiation on a platform: the informative and persuasive role of advertising.

Dou, Y., & Wu, D. (2019). Platform competition under network effects: piggybacking and optimal subsidization. *Georgia Tech Scheller College of Business Research Paper*(18-3).

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

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

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

Gatti, N., & Rocco, M. (2013). Which mechanism for sponsored search auctions with externalities? In *Proceedings of the 2013 international conference on autonomous agents and multi-agent systems* (pp. 635–642).

Gershkov, A., & Moldovanu, B. (2009). Dynamic revenue maximization with heterogeneous objects: A mechanism design approach. *American economic Journal: microeconomics*, *1*(2), 168–98.

Gershkov, A., & Moldovanu, B. (2010). Efficient sequential assignment with incomplete information. *Games and Economic Behavior*, *68*(1), 144–154.

Ghosh, A., & Mahdian, M. (2008). Externalities in online advertising. In *Proceedings of the 17th international conference on world wide web* (pp. 161–168).

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

Giotis, I., & Karlin, A. R. (2008). On the equilibria and efficiency of the gsp mechanism in keyword auctions with externalities. In *International workshop on internet and network economics* (pp. 629–638).

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

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

Gomes, R., Immorlica, N., & Markakis, E. (2009). Externalities in keyword auctions: An empirical and theoretical assessment. In *International workshop on internet and network economics* (pp. 172–183).

Gomes, R., & Pavan, A. (2016). Many-to-many matching and price discrimination. *Theoretical Economics*, *11*(3), 1005–1052.

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

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

Hagiu, A. (2011). *Quantity vs. quality: Exclusion by platforms with networks effects.* Citeseer.

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

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

Heish, Y., Shum, M., & Yang, S. (2014). *To score or not to score? structural estimation of sponsored search auctions* (Tech. Rep.). mimeo.

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

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

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

Itoh, M. (1983). Monopoly, product differentiation and economic welfare. *Journal of Economic Theory*, *31*(1), 88–104.

Iyengar, G., & Kumar, A. (2006). Characterizing optimal keyword auctions. In *Second workshop on sponsored search auctions.*

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

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

Jeitschko, T. D. (1999). Equilibrium price paths in sequential auctions with stochastic supply. *Economics Letters*, *64*(1), 67–72.

Jeon, D.-S., Kim, B.-C., & Menicucci, D. (2016). Second-degree price discrimination by a two-sided monopoly platform.

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

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

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

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

Lahaie, S. (2006). An analysis of alternative slot auction designs for sponsored search. In *Proceedings of the 7th acm conference on electronic commerce* (pp. 218–227).

Lu, S., Zhu, Y., & Dukes, A. (n.d.). Position auctions with budget-constrained advertisers.

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

Maskin, E., & Riley, J. (1984). Monopoly with incomplete information. *The RAND Journal of Economics*, *15*(2), 171–196.

McAfee, R. P., & McMillan, J. (1987). Auctions with a stochastic number of bidders. *Journal of economic theory*, *43*(1), 1–19.

Mierendorff, K. (2013). The dynamic vickrey auction. *Games and Economic Behavior*, *82*, 192–204.

Mierendorff, K. (2016). Optimal dynamic mechanism design with deadlines. *Journal of Economic Theory*, *161*, 190–222.

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

Mussa, M., & Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, *18*(2), 301–317.

Muthukrishnan, S. (2009). Bidding on configurations in internet ad auctions. In *International computing and combinatorics conference* (pp. 1–6).

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

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

Pai, M. M., & Vohra, R. (2013). Optimal dynamic auctions and simple index rules. *Mathematics of Operations Research*, *38*(4), 682–697.

Parkes, D. C., & Singh, S. P. (2004). An mdp-based approach to online mechanism design. In *Advances in neural information processing systems* (pp. 791–798).

Pavan, A., Segal, I., & Toikka, J. (2008). Dynamic mechanism design: Revenue equivalence, profit maximization and information disclosure preliminary and incomplete.

Peitz, M., & Valletti, T. M. (2008). Content and advertising in the media: Pay-tv versus free-to-air. *international Journal of industrial organization*, *26*(4), 949–965.

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

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

Roberts, M. J. (n.d.). Spence, michael (1976).". *Effluent Charges and Licenses Under Uncertainty." Journal of Public Economics*, *5*, 193–208.

Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the european economic association*, *1*(4), 990–1029.

Rochet, J.-C., & Tirole, J. (2006). Two-sided markets: a progress report. *The RAND journal of economics*, *37*(3), 645–667.

Said, M. (2011). Sequential auctions with randomly arriving buyers. *Games and Economic Behavior*, *73*(1), 236–243.

Said, M. (2012). Auctions with dynamic populations: Efficiency and revenue maximization. *Journal of Economic Theory*, *147*(6), 2419–2438.

Satterthwaite, M., & Shneyerov, A. (2007). Dynamic matching, two-sided incomplete information, and participation costs: Existence and convergence to perfect competition. *Econometrica*, *75*(1), 155–200.

Schmalensee, R. (2011). Why is platform pricing generally highly skewed? *Review of Network Economics*, *10*(4).

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

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

Spence, A. M. (1975). Monopoly, quality, and regulation. *The Bell Journal of Economics*, 417–429.

Spence, A. M. (1980). Multi-product quantity-dependent prices and profitability constraints. *The Review of Economic Studies*, *47*(5), 821–841.

Suarez, F. F., & Cusumano, M. A. (2009). The role of services in platform markets. *Platforms, markets and innovation*, 77–98.

Szalay, D. (2009). Regulating a multi-attribute/multi-type monopolist. *University of Bonn, unpublished working paper*.

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

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

Viecens, M. F. (2006). Two-sided platforms with endogenous quality differentiation.

Watts, A. (2016). Two ways to auction off an uncertain good. *Journal of Economics*, *119*(1), 1–15.

Webster, M., Wing, I. S., & Jakobovits, L. (2010). Second-best instruments for near-term climate policy: Intensity targets vs. the safety valve. *Journal of Environmental Economics and Management*, *59*(3), 250–259.

Weitzman, M. L. (1978). Optimal rewards for economic regulation. *The American Economic Review*, *68*(4), 683–691.

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

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