Targeted Search and Platform Design

Job Market Paper, Dissertation Essay 1

Zemin (Zachary) Zhong*

This version: August 18, 2016

For the most recent version, please go to tiny.cc/zhongjmp

Abstract

Major online platforms such as Amazon and eBay have invested significantly in search technologies to direct consumer searches to relevant products. These technologies lead to targeted search, implying consumers are visiting more relevant sellers first. For example, consumers may directly enter their desirable attributes into search queries, and the platform will retrieve relevant sellers accordingly. The platform may also let consumers refine the search outcomes by various criteria. This study characterizes the role of targeted search, and examines how targeted search affects market equilibrium and platform design.

I model targeted search in a differentiated market with many firms where consumers search sequentially for the best product match. Within this setup, I endogenize the search design by allowing the platform to choose the precision of targeted search and the revenue model contract. One of the central results of the analysis is how targeted search affects equilibrium prices. I find its impact on price is not monotonic. When targeting is not too precise, targeted search lowers the equilibrium price. It makes sellers more similar and intensifies price competition, despite the fact that all consumers face sellers with better fit. However, once the targeting becomes sufficiently precise, the equilibrium price increases, because highly targeted search discourages active consumer search and gives sellers monopoly power.

*Haas School of Business, UC Berkeley. Email: zachary_zhong@haas.berkeley.edu. I am indebted to my advisors, Ganesh Iyer, John Morgan, Zsolt Katona, and Ben Handel for input at all stages of this project. I also thank Yuichiro Kamada, Steve Tadelis, and Przemyslaw Jeziorski for helpful suggestions. I gratefully acknowledge financial support from the NET Institute summer research grant. All errors are my own.

Furthermore, I consider two major platform revenue models, commission and promoted slots, with consumer search. The platform, by providing targeted search with precision up to the aforementioned limit, can extract more consumer surplus through higher commission rates, because targeted search improves consumer surplus by lowering search cost, increasing fit, and lowering price. With targeted search up to the limit, the platform can also extract more surplus from sellers by offering promoted slots, because sellers can use promoted slots to better target consumers. However, once targeted search becomes too precise, the market will face a price hike, hurting the platform revenue in both models. Therefore, I find that in both revenue models, the platform may want to limit the precision of targeted search even if improving it is costless, with or without consumer entry. Using a unique dataset from Taobao, I find suggestive evidences that are consistent with the model predictions.

Keywords: Consumer Search, Platform, e-Commerce

1 Introduction

Search technology enables consumers to more easily find their desired goods or services. Imagine a consumer searching for a black leather case for her iPad mini. She can directly type in "iPad mini leather black case" on Amazon, where she will find 87,182 related listings, ranked by "Relevance." She can also use the refinement tools to select the desirable attributes, such as product category, material, and color. Virtually all major platforms adopt similar search technologies. For example, eBay also offers a variety of refinement tools. And the same search query on eBay generates 12,976 listings ranked by "Best Match." In Taobao, the default ranking is by "Comprehensive" which is based largely on relevance, and various refinement tools are also readily available. Table 1 summarizes the default ranking major online platforms use as well as the availability of refinement tools. Most platforms use either relevance ranking or comprehensive ranking by default, and search refinement tools are widely available.

Search refinements are filters that can be applied to the search results to filter out less relevant sellers. In the Amazon search example, users can choose the category (e.g., between Tablet Cases, Tablet Stands, and Tablet Keyboard Cases). They can also refine by attributes, such as material (leather, plastic, etc.), and color, as well as delivery day, Prime or not, and brands, among others. On eBay, buyers can further refine by the location of sellers, listing format (auction vs. buy-it-now), and so on. These refinements help match sellers with search queries more precisely.

Consumers can also directly enter the desirable attributes in their search queries, and the search engine will try to use relevance ranking to match the search query with sellers by algorithms. A ranking mechanism is a retrieval function that takes the user input to rank possible outcomes. There is an extensive literature in Computer Science on training the algorithm using various data. [1] Following our previous example, when a consumer types "iPad mini leather black case" in the search box of Amazon, the website's search engine will select a subset among its hundreds of millions of items according to the relevance of listing attributes to the keywords in the search query. Tablet cases with features that match "iPad mini," "leather," and "black" will get more prominent slots. Furthermore, recent developments in natural language processing and AI allow the platform to understand conversational language in addition to the traditional keywords, and e-commerce platforms are investing in such search technologies. [2]

In addition to refinements and search queries, platforms can use personalized data to optimize the search targeting. It includes, but is not limited to, a consumer's device, location, IP address, cookies, and historical activities such as clicks and purchases. Empowered by developments in machine learning techniques and the availability of data, personalized search is becoming increasingly common, especially in mobile devices. For instance, Google has been using personalized ranking for all searches since 2009. In Table 1, the default ranking of Tripadvisor, "Just for You," uses the customers'

[1] For instance, see Agichtein et al. (2006) and Liu (2009) for advances and reviews on information retrieval. See Dumais et al. (2008),Thirumalai (2013), and Nuzzi (2015) for related patents filed by major platforms such as Microsoft, Amazon, and eBay.

[2] They includes Amazon Echo, Google Assistant, and Apple Siri. Alibaba, the parent company of Taobao and TMall, invested in Twiggle, a company that develops e-commerce search engine, that delivers more accurate results than current options, while allowing shoppers to write queries in conversational language.

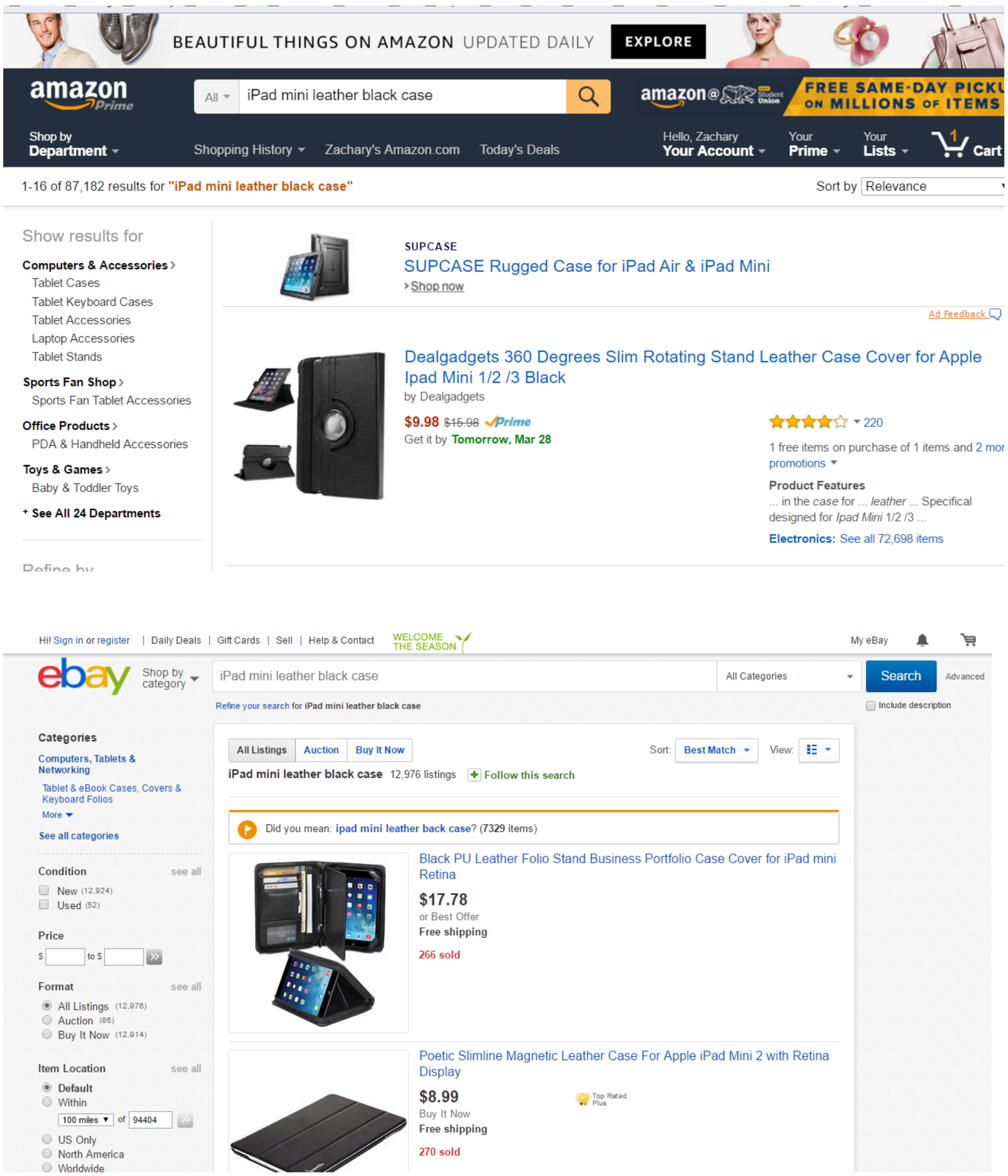

Figure 1: Relevance ranking examples: Amazon (Relevance) and eBay (Best Match)

Table 1: Default rankings by major online platforms and OTAs

Platform	Category	Default Ranking	Refinement Tools
Amazon	general merchandise	Relevance	Yes
eBay	general merchandise	Relevance	Yes
craiglist	general merchandise	Relevance or time of posting	No
etsy	handmade or vintage items	Relevance	Yes
Rakuten	general merchandise	Relevance	Only category
Bestbuy	electronics	Relevance	Yes
Sears	general merchandise	Relevance	Yes
Newegg	electronics	Featured	Yes
Walmart	general merchandise	Relevance	Yes
Taobao	general merchandise	Comprehensive	Yes
Tmall	general merchandise	Comprehensive	Yes
JD	general merchandise	Comprehensive	Yes
Suning	general merchandise	Comprehensive	Yes
Tripadvisor	hotels	personalized ”Just for you”	Yes
Booking	hotels	”recommended”	Yes
Expedia	hotels	”recommended”	Yes
Orbitz	hotels	”recommended”	Yes
Hotels	hotels	”recommended”	Yes
Hotwire	hotels	”popular”	Yes
Priceline	hotels	”popular”	Yes
Kayak	hotels	”recommended”	Yes

past browsing and review data to rank hotels. In a recent Kaggle competition, [3] Airbnb asked participants to design an algorithm to predict users' preferences using demographics, web session records, and attributes such as device and browser for user activities. Jetlore, a start-up that specializes in providing personalized ranking solutions to e-commerce, has provided its service to numerous Internet Retailer 500 firms, including eBay and Paypal. Continuing our example of shopping for iPad cases, eBay adjusts the ranking of listings by a user's past purchase and browsing history as well as her location and device. The personalization of search will lead to different search outcomes and rankings for different customers even if they are using the same search query. And it achieved remarkable precision of matching between consumers and sellers.

These advances in search technology have profound implications for consumer search and platform design. Targeted search changes consumer's optimal search strategy, and hence the demand sellers face. As search becomes more targeted, consumers are facing better-matched sellers, which increases the demand. However, targeted search also makes sellers more similar, leading to fiercer price competition. These two effects have opposite price implications. Therefore, the impact of targeted search on equilibrium prices is not obvious. Furthermore, the price and consumer surplus will not only affect the profitability of sellers, but also the platforms. Because the adoption of search technology is a platform decision, how it interacts with different revenue models is also important. How will targeted search affect the platform's optimal fee structure? Will it help or hurt the platform's monetization such as advertising revenue?

This paper is among the first to study search technology and its implications for e-commerce platform design. I model consumer targeted search in a horizontally differentiated market. In this setting, the platform uses user search query and personalized data to select and rank sellers by the fitness of their match with buyers. And as the platform improves its search technology, the search becomes more targeted, leading to better matches. I show the effect of targeted search is not monotonic on the equilibrium price. When the targeting is not too precise, targeted search makes consumers visit only sellers that have better matches and hence more similar. This intensifies com-

[3] See https://www.kaggle.com/c/airbnb-recruiting-new-user-bookings/data .

petition, leading to more marginal consumers for relevant sellers. And the more elastic demand reduces the equilibrium price. I further prove this effect is equivalent to reducing the search cost by saving on the cost of visiting irrelevant sellers. However, when the targeting becomes so precise that consumers do not search actively, the equilibrium price will be discontinuously higher and increasing in targeting precision, because "no-search" gives sellers monopoly power and the power is increasing when buyers have better matches. This effect is not captured by reducing search cost because negligible search cost implies extremely intense competition and razor-thin margins.

Targeted search can increase consumer surplus in three ways: saving search costs, reducing prices, and offering better matches. Therefore, it encourages consumer entry until it becomes sufficiently precise to discourage consumer search and hence gives sellers monopoly power. Combining with its implication on equilibrium prices, the platform faces a design problem in choosing the optimal search precision, since there are interactions between targeted search and platform revenues. I consider two commonly observed platform revenue models: commission and promoted slots. Platforms often charge a percentage commission for each transaction (for example, 5% for Amazon). I model the platform as a monopoly and chooses a commission rate to maximize its profit. I show that as search outcomes becomes more targeted, the optimal commission rate becomes higher, leading to a higher profit for the platform. When I allow for consumer entry, targeted search increases both the platform profit and consumer surplus up to a threshold. For the optimal search design, I find that the platform may want to limit its precision of search targeting even if improving it is costless, because very precise targeting leads to price hikes in the commission model, with or without consumer entry. Many platforms also sell promoted slots for each search keyword. For instance, Amazon sells "Sponsored Products" slots on page one of search results page that sellers can bid for. In the promoted slot model, targeted search, up to the aforementioned precision limit, increases the value of top promote slots by increasing the relative market share of promoted sellers. The platform gets higher profit from selling the same top promoted slots with targeted search. This effect is exactly the opposite of reducing search cost, because reducing search cost makes promoted slots less valuable as more consumers continue search.

Using a large and unique dataset from Taobao, I find suggestive evidences that are consistent with the model predictions. In particular, when the measured targeted search precision is not very high, it is negatively correlated with prices, while the opposite is true when the targeted search is sufficiently precise. For the promoted slot model, I find that the measured targeted search precision is positively correlated with the market share of promoted sellers, also consistent with the model prediction.

2 Literature Review

Broadly, this study relates to two sets of literature: consumer search and the economics of platforms. This paper contributes to consumer search literature by modeling targeted search in a e-commerce platform. To the platforms literature, this study contributes by pioneering in investigating how optimal platform revenue models interact with search, and in particular, targeted search.

Studies on consumer search in differentiated market lead to a large literature. Wolinsky (1986) and Anderson and Renault (1999) propose models of sequential search in a market with horizontal differentiation. Kuksov (2004) further endogenizes product design in the search context. He finds that lower search cost increases product differentiation, which may increase prices. Bar-Isaac et al. (2012) explores the possibility of broad versus niche designs in a search context by allowing firms to have both vertical and horizontal differentiation. These studies do not model targeted search. Instead, they interpret the effect of online search technology as reducing search cost, leading to different implications compared to this study.

A growing literature on ordered or directed search has begun that deviates from the standard assumption of random search. Arbatskaya (2007) explores the market equilibrium of homogeneous goods when consumers are forced to visit sellers in a particular order. Armstrong et al. (2009) model one firm as "prominent" as being the first to be sampled, whereas the rest are sampled randomly. Armstrong and Zhou (2011) and Chen and He (2011) further endogenizes the prominence by assuming firms either pay or have lower prices to gain prominence. The main finding is that because the prominent firm faces more elastic "fresh" demand compared to less elastic demand from recall, its price will be lower. By contrast, this study endogenizes prominence by modeling targeted search, that is, making consumers visit more relevant firms first, and more importantly, modeling the design of search targeting as a strategic choice of the platform.

The study is closely related to recent progress in targeted search. Yang (2013) modeled targeted search by endowing consumers with exogeneous discrete types, and as the quality of search increases, the probability of finding matched type increases. Yang (2013) does not capture the idea that as online search technology advances, platforms will be able to better segment consumers *within* each category as is the case in this study. The difference leads to quite different, and sometimes opposite, implications of search technology in equilibrium prices and platform incentives. In De Corniere (2016) model, consumers' types are uniformly distributed in a circle, and improvement of targetability is modeled as a better match between advertisers and consumer types. The major difference is that this model focuses on the organic search results instead of sponsored search. Instead of advertisers choosing their targets, e-commerce platforms decide how sellers match buyers, which leads to different pricing implications. The search engines also have different instruments and incentives than e-commmerce platforms; therefore, the optimal contract and design of the two kinds of platforms are completely different.

Recent empirical studies have investigated some aspects of online targeted search. Ghose et al. (2014) explore how rankings affect consumer search and platform revenues. One of their key findings, that relevance ranking increases platform revenue, is consistent with our model prediction. Chen and Yao (2016) find that refinement tools increase consumer utility of purchased products, also consistent with our model prediction. Ursu (2016) shows that rankings affect consumer search and increase consumer welfare by reducing search cost as well as platform revenue.

This study also contributes to the economics of platforms literature by combining platform incentives with underlying consumer search. Sriram et al. (2015) point out that "advertising and consumer search are critical for web-based commerce....The optimal quantity of search functionality and/or regulated advertising is therefore a crucial platform management question." Eliaz and Spiegler (2011) consider how the search engine might want to include more firms with less relevance to generate more clicks. Hagiu and Jullien (2011, 2014) considered the platforms' incentive to diver searches. More broadly, this study is also related to the two-sided market literature (Rochet and Tirole, 2003, 2006; Armstrong, 2006; Rysman, 2009) by micro-modeling the network externalities in the search process. The ad-slot model is related to the sponsored-search literature (Athey and Ellison, 2011; De Corniere, 2016; De Corniere and De Nijs, 2016). Dukes and Liu (2015) also model search design by letting the platform choose a search cost in a setting where consumers have to search in two dimensions, both between and within products. However, unlike our study, the authors do not explicit model targeted search, nor do they model the revenue models of the platform.

3 Modeling Targeted Search

We start from the classic model of sequential search for horizontal differentiated products. Consider a product sold by n distinct sellers denoted by $i = 1, 2, ...n$. Each seller produces a single product with constant marginal cost c. There is a set of consumers (per seller mass normalized to 1), each with unit demand. The utility of buyer j consuming x_0 numeraire good and a unit of brand i is given by $u_j(x_0, i) = x_0 + v_{ij}$, where v_{ij} captures how good seller i's product matches the need of consumer j. We assume that v_{ij} are realizations of an i.i.d. distribution with CDF G that has finite support $[\underline{v}, \bar{v}]$. We assume its density $g(v)$ is strictly positive, smooth, and log-concave with no further restrictions, unless otherwise specified. The value of the outside option is normalized to zero, an assumption we will relax in one of the extensions.

By assuming v_{ij} is an i.i.d. draw across both i and j, we are focusing on horizontal differentiation. For example, think of sellers as different brands of iPad cases with different colors, styles, and materials. Different types of consumers prefer different combinations of these attributes. The i.i.d. assumption requires that no correlation exists between preferences of different types of consumers. Meanwhile, all sellers are symmetric in the sense that matched values are identically distributed across sellers. These properties are essential to getting a unique pure pricing equilibrium. The set-up is consistent with Wolinsky (1986).

The timing of the game is as follows. All sellers choose prices simultaneously. Consumers know only the number of sellers yet not prices or values v_{ij}. Based on their belief about the equilibrium price distribution, they decide whether to participate, and if they do, they search sequentially. They pay search cost s per visit to learn both the price and value of the particular seller, and decided whether to purchase (stop), continue, or exit. They search without replacement and with perfect recall.

In the baseline, consumers visit sellers randomly. We then introduce targeted search as follows. We assume the platform only has imperfect knowledge about the matched value because of limited consumer inputs by keywords in queries and/or refinements. Following the previous example, a consumer might enter the keywords "iPad mini leather black case." Among all cases for iPad mini, Amazon will firstly show all listings with two features that fit, namely those cases whose color is black and material is leather. Then it will show those cases that are black but not leather, or vice versa, followed by cases that are neither black nor leather. Within cases that are both black and leather, however, Amazon does not have consumer-specific information to further differentiate the sellers. This buyer may be looking for a case with stripe patterns but without providing such a refinement option or recognizing such a keyword, Amazon will not be able to show such cases to her first.

This observation motivates the following modeling of targeted-search technology. Consider a finite set of thresholds $x_i \in X$ such that $\underline{v} < x_i < x_j < \bar{v}$,$\forall x_i, x_j \in X$,$i < j$. Specifically, we assume the platform only observed $\{1_{\{v>x_i\}},\ x_i \in X\}$, that is, whether the matched value exceed thresholds x_i. I interpret the set X as the targeted search technology of the platform: if the platform recognizes more attributes in the search query, offers more refinement tools, or uses more consumer personalized data, there will be more thresholds x_i in the set X, and the highest threshold will get closer to $\bar{v}$, the upper bound of matched value. It will become clear that the key parameter of X is the highest threshold, denoted by $x = \max\{X\}$. x is a measure of the precision of targeted search, because targeted search helps consumers find better matches, that is, sellers with value greater than x. I call the set of sellers with values greater than x relevant sellers (those sellers whose features perfectly fit consumers' input and refinement, as all the sellers with black leather cases in our iPad case example), and call the rest irrelevant sellers.

Upon observing $1_{\{v>x\}}$, the platform ranks every seller with $v > x$ randomly to buyers first, followed by a random sequence of the next best set of sellers, and so forth. [4] It is easy to see that within relevant sellers, where $v > x$, the matched values are i.i.d. with cumulative distribution given by $G_x(v) = G(v|v > x)$. I also assume consumers know the set X and, before deciding whether to incur search cost to visit one seller, know whether the seller is relevant or not. Therefore, consumers behave as if they know the number of firms in each "bucket". This assumption makes the transition between relevant and irrelevant sellers not contingent on the depth of search. It is innocuous when the number of sellers is large, because such a transition probability will be negligible throughout the search. We focus on such cases in this study, and this assumption does not affect the result in any substantive way.

4 Equilibrium Analysis

4.1 Baseline: Random Search

We first consider the case in which the platform does not adopt targeted search, as might be the case in the early days of e-commerce. [5] Consumers search randomly, and the value distribution of any sample visit is $G(v)$. Wolinsky (1986) focuses on symmetric price equilibrium, showing consumers search with reservation value defined by

$$\int_{w^*}^{\bar{v}} (v - w^*) dG(v) = s \tag{1}$$

[4]The platform is assumed to be "faithful" in the sense that it always ranks all relevant sellers above irrelevant ones. Later this strategy can be shown to be readily rationalizable in both commission and ad-slot models: it is profit maximizing for the platform to be faithful in ranking.

[5]Depending on the search query, Craigslist sometimes uses the time of posting as the default ranking, which is quasi-random.

And if no such w^* exists, let $w^* = \underline{v}$. The search strategy is characterized by (1) consumers participating if $w^* \geq p^*$, and (2) consumers stopping if and only if they find a value no less than w^*. In the case they have exhausted all n sellers, they either recall the option with highest value if it is higher than the price, or exit. When s is not too large, such that $s < \int_{\underline{v}}^{\bar{v}} (v - \underline{v}) dG(v)$, which guarantees consumers do actively search, we have demand as follows. Let $w(p) = w + p - p^*$. The demand of any seller with price p (whereas all other sellers charge p^*) is given by

$$D(p, p^*, n) = [1 - G(w(p))] \frac{1 - G(w^*)^n}{n[1 - G(w^*)]} + \int_p^{w(p)} G(p^* - p + v)^{n-1} g(v) dv \qquad (2)$$

in which the first component captures the probability that a consumer stops at and purchases from the focal seller with price p immediately, whereas the second component captures demand from consumers who have exhausted all available options and recall this particular seller who yields the highest surplus.

As with most search models, there exists a set of uninteresting equilibria in which prices are so high that no consumer participates. Wolinsky (1986) shows the existence of symmetric pure price equilibrium with $p^* < w$ such that consumers do participate. The equilibrium price is simply given by

$$\begin{aligned} p^* &= c - \frac{D(p^*, p^*, n)}{D_p(p^*, p^*, n)} \\ &= c + \frac{1 - G(p^*)^n}{g(w^*) \dfrac{1 - G(w^*)^n}{[1 - G(w^*)]} - n \int_p^{w(p)} G(v)^{n-1} g'(v) dv} \end{aligned} \qquad (3)$$

Having derived the equilibrium price, we are interested in how it changes with search cost s. We are especially interested in characterizing the relationship when the number of sellers n is large, as in many if not most online markets. We know that by Equation 3, as $n \to \infty$, $p^* \to c + \dfrac{1 - G(w^*)}{g(w^*)}$, leading to the following corollary:

Corollary 1. *When $n \to \infty$, the equilibrium price p^* is increasing in search cost s.*

Proof. Applying the implicit function theorem to Equation 1, let $f = \int_{w^*}^{\bar{v}} (v - w^*) dG(v) - s$.

We have that $\frac{dw^*}{ds} = -\frac{\partial f/\partial s}{\partial f/\partial w^*} = -\frac{1}{1-G(w^*)} < 0$. Also notice $\lim_{n\to\infty} p^* = c + \frac{1-G(w^*)}{g(w^*)}$. Therefore, $\frac{\partial p^*}{\partial w^*} < 0$ due to the log-concavity of $G(v)$, thus $\frac{\partial p^*}{\partial s} > 0$. $\square$

The intuition is that as search cost s decreases, consumers become more particular about finding the best match (their reservation utility w^* increases). Therefore, the probability of stopping at any seller $1 - G(w^*)$ decreases, implying that, on average, they visit more sellers. Sellers in turn faces more "traffic," and hence more marginal consumers. And more marginal consumers make the demand more elastic. In response, the equilibrium price decreases. In the extreme case, when $s \to 0$, price approaches marginal cost when n is arbitrarily large. The following example uses uniform distribution to illustrate the mechanism.

Example 4.1. When $G(v)$ is uniform, and $G(v) = \frac{v-\underline{v}}{\bar{v}-\underline{v}}$, $g(v) = \frac{1}{\bar{v}-\underline{v}}$, and $g'(v) = 0$, we have

$$p^* = c + \frac{(\bar{v}-w^*)[1-(\frac{p^*-\underline{v}}{\bar{v}-\underline{v}})^n]}{1-(\frac{w^*-\underline{v}}{\bar{v}-\underline{v}})^n} \tag{4}$$

in which w^* is given by $w^* = \bar{v} - \sqrt{2s(\bar{v}-\underline{v})}$, clearly decreasing in s. When $n \to \infty$, the explicit form is given by $p^* = c + (\bar{v} - w^*) = c + \sqrt{2s(\bar{v}-\underline{v})}$, increasing in s, and goes to c when $s \to 0$.

4.2 Equilibrium in Targeted Search

To derive equilibrium under targeted search, we use the same method as in the previous analysis. We also focus on symmetric pure price equilibria. Notice that introducing the cutoff x of relevant sellers changes both the underlying value distribution $g(v)$ to a truncated one $g_x(v) = g(v|v > x)$ and the number of relevant sellers to k, which follows a binomial distribution with n sellers and probability $1 - G(x)$. Define $g_x(v) = g(v|v > x) = \frac{g(v)\cdot 1_{v>x}}{1-G(x)}$ as the truncated PDF. Similarly, the CDF is given by $G_x(v) = \frac{\int_x^v g(y)dy}{1-G(x)}$. Because all sellers with $v > x$ have i.i.d. values for each consumer, we still have that consumers search with a stationary reservation value given by

$$\int_{w^*}^{\bar{v}} (v - w^*) g(v) dv = (1 - G(x))s \tag{5}$$

Let us discuss whether such w^* exists. Clearly, when x gets closer to $\bar{v}$, the incentive to continue search beyond getting a match decreases. For instance, when $\bar{v} - x \leq s$, consumers have no incentive to search beyond the first visit, because even the highest possible marginal benefit $\bar{v} - x$ does not offset the search cost s. In fact, we can find that there exists an unique x^*, such that for any $x > x^*$, consumers have no incentive to search, and for any $x < x^*$, they do, with w^* defined by Equation 5. The cutoff x^* is given by $\int_x^{\bar{v}} (v - x) g_x(v) dv = s$. The following lemma states the assumptions to guarantee its existence and uniqueness.

Lemma 1. *There exists an unique threshold precision of targeted search that consumers facing more precise targeted search do not have incentive to search further, as long as the search cost is not too large, i.e.* $\exists x^*$ *defined by* $\int_x^{\bar{v}} (v - x) g_x(v) dv = s$ *if* $\int_{\underline{v}}^{\bar{v}} (v - \underline{v}) g(v) dv > s$*. Moreover,* $\int_x^{\bar{v}} (v - x) g_x(v) dv < s$ *if and only if* $x > x^*$*.*

Proof. See Appendix A.1. □

For comparative statics, we have that $\dfrac{\partial w^*}{\partial x} > 0$ and $\dfrac{\partial w^*}{\partial s} < 0$. Both are readily verifiable using the implicit function theorem: Let $f = \int_{w^*}^{\bar{v}} (v - w^*) g(v) dv - (1 - G(x))s$. We have that $\dfrac{\partial w^*}{\partial x} = -\dfrac{\partial f / \partial x}{\partial f / \partial w} = \dfrac{g(x)s}{1 - G(w)} > 0$. This shows that as targeted search becomes more precise (or search cost decreases), consumers are demanding better fits.

Adding some mild conditions (the density being not only log-concave, but also concave), we can also show relevance ranking decreases the search depth of consumer search. See Corollary 2.

Corollary 2. *The probability that consumers stop* $(1 - G_x(w))$ *increases in targeted search precision* x *when* $g(v)$ *is (weakly) concave.*

Proof. See Appendix A.2. □

Example 4.2. Consider the case in which $g(v)$ is uniform at the interval $\underline{v}, \bar{v}$. We then have that $G(x) = \dfrac{x - \underline{v}}{\bar{v} - \underline{v}}$ and $g_x(v) = \dfrac{1}{\bar{v} - x}$. Solving w^*, when $w^* > x$, we have $w^* =$

$\bar{v} - \sqrt{2s(\bar{v} - x)}$. We also have the cut-off x^* given by $x^* = \bar{v} - 2s$, decreasing in s as expected. We can show that when $x > x^*$, the w^* given by Equation 5 will be smaller than x; hence, we need to define $w^* = x \ \forall x > x^*$. When $x < x^*$, the w^* given by Equation 5 will be larger than x as desired.

We also have $1 - G_x(w) = \sqrt{\frac{2s}{v - x}}$ that is increasing in x. Therefore, it implies that targeted search (or higher search cost) increases the possibility of a consumer stopping at any seller, reducing the length of search.

Having shown the search process is stationary with reservation utility given in Equation 5, the following proposition characterizes the symmetric price equilibrium when $x < x^*$.

Proposition 1. *When the targeted search is not too precise ($x < x^*$), there exist a symmetric pure price equilibrium with price given by:*

$$p^* = c - \frac{\sum_{k=1}^{n} \frac{n!(1-G(x))^k G^{n-k}(x)}{k!(n-k)!}[1 - G_x(p^*)^k] + \epsilon}{\sum_{k=1}^{n} \frac{n!(1-G(x))^k G^{n-k}(x)}{k!(n-k)!} \left\{ -g_x(w^*) \frac{1 - G_x(w^*)^k}{1 - G_x(w^*)} + \int_{p^*}^{w(p^*)} G_x(v)^{k-1} g_x'(v) dv + \epsilon_p \right\}} \tag{6}$$

To prove this proposition, we start by introducing the demand function. One main difference between the demand function in targeted search and the one in random search is that the number of sellers in the first block is uncertain. If there were k firms in the first block, the demand one seller faces when $k > 0$ is simply

$$D(p, p^*, x, k) = [1 - G_x(w(p))] \frac{1 - G_x(w^*)^k}{k[1 - G_x(w^*)]} + \int_p^{w(p)} G_x(p^* - p + v)^{k-1} g_x(v) dv$$

When $k = 0$, we denote the demand as $D(p, p^*, x, 0)$ that can be shown to be finite. Notice k follows a binomial distribution of n sellers with probability $1 - G(x)$. Therefore, the expected demand is given by

$$D(p, p^*, x) = \sum_{k=1}^{n} \frac{n!(1-G(x))^k G^{n-k}(x)}{k!(n-k)!} D(p, p^*, x, k) + \epsilon$$

where $\epsilon = G^n(x) D(p, p^*, x, 0)$ is the residual demand when there is no seller in the first block. Clearly it converges to zero very fast when n increases. Therefore, in our following analysis that focuses on large n, its contribution will be negligible. Evaluating it at equilibrium price, we have

$$D(p^*, p^*, x) = \sum_{k=1}^{n} \frac{n!(1-G(x))^k G^{n-k}(x)}{k!(n-k)!} \frac{1}{k}[1 - G_x(p^*)^k] + \epsilon \tag{7}$$

Similarly, we have that the demand's derivative to price at equilibrium is given by

$$D_p(p^*, p^*, x) = \sum_{k=1}^{n} \frac{n!(1-G(x))^k G^{n-k}(x)}{k!(n-k)!} \left\{ -g_x(w^*) \frac{1 - G_x(w^*)^k}{k[1 - G_x(w^*)]} + \int_{p^*}^{w(p^*)} G_x(v)^{k-1} g_x'(v) dv \right\} + \epsilon_p$$

The price is given by $p^*(x) = c - \frac{D(p^*, p^*, x)}{D_p(p^*, p^*, x)}$, which is a function of x, n and s. Expand it, we have the price given by Equation 6. To show the existence of $p^*(x)$, notice that when p^* is large, the right-hand side goes to c (which will be smaller than the LHS), but when p^* is small, the RHS is larger than the LHS. Therefore, by continuity of both sides on p^*, there exist at least one p^* that satisfies $p^*(x) = c - \frac{D(p^*, p^*, x)}{D_p(p^*, p^*, x)}$. Uniqueness, on the other hand, is not guaranteed as in the original Wolinsky (1986) model unless we assume some specific functional form of $g(v)$ such as uniformity. However, when $n \to \infty$, the limit price is unique as the next proposition shows.

When n is relatively large, we can use the normal approximation to binomial distribution. Let $\mu = n(1 - G(x))$, and $\sigma = \sqrt{n(1 - G(x))G(x)}$. We have that the number of sellers in the first block will be approximately normally distributed with $\mathcal{N}(\mu, \sigma)$. The price can be written as

$$p^* = c - \frac{\int_k e^{-\frac{(k-\mu)^2}{2\sigma^2}} \frac{1}{k}[1 - G_x(p^*)^k] dk + \epsilon}{\int_k e^{-\frac{(k-\mu)^2}{2\sigma^2}} \left\{ -g_x(w^*) \frac{1 - G_x(w^*)^k}{k[1 - G_x(w^*)]} + \int_{p^*}^{w(p^*)} G_x(v)^{k-1} g_x'(v) dv \right\} dk + \epsilon_p} \tag{8}$$

Notice that when $x = \underline{v}$, $G(x) = 0$, the distribution of k degenerate to $k = n$ with probability 1. The price given by Equation 6 or Equation 8 is exactly the same as the equilibrium price of random ranking (Equation 3).

The following proposition shows the first main result: the equilibrium price decreases in the precision of targeted search when n is sufficiently large.

Proposition 2. *When there are infinite number of sellers, the equilibrium price is unique, which is given by* $\lim_{n\to\infty} p^*(x) = c + \frac{1 - G_x(w^*)}{g_x(w^*)} = c + \frac{1 - G(w^*)}{g(w^*)}$ *when* $x < x^*$*. Furthermore,*

1. *The equilibrium price is decreasing in targeted search precision x: $\frac{\partial p^*}{\partial x} < 0$.*
2. *The same comparative statics holds for finite number of sellers: There exist a finite N such that $\forall n > N$, $p^*(n)$ is strictly decreasing in x when $x \in [\underline{v}, x^*]$.*

Proof. See Appendix Section A.3. □

Proposition 2 shows that when the targeted search is not too precise ($x < x^*$), the effect of targeted search on equilibrium price is to decrease equilibrium price. At first glance, this might seem counter intuitive. After all, as search becomes targeted, consumers face sellers with better fit. The demand, therefore, should be higher. However, targeted search produces another effect: it makes sellers more similar and hence more competitive. Higher targeted search precision x saves the visits to unfit sellers. It increases the marginal benefit of visiting more sellers. Therefore, consumers visit more relevant sellers with higher reservation utility. This increases the number of marginal consumers a relevant seller faces. Responding to more elastic demand, sellers lower equilibrium prices. In fact, Proposition 2 shows the price effect of targeted search (when it is not too precise, i.e., $x < x^*$) is equivalent to reducing search cost when there are infinitely many sellers: introducing x effectively reduces search cost to $(1 - G(x))s$. Clearly this intuition holds when n is large but finite, as the last claim in Proposition 2 shows. The proof uses the uniform convergence property of equilibrium price with respect to n. Therefore its comparative statics must also converge for a sufficiently large n.

To complete the characterization of equilibrium price, we now focus on the case with sufficiently high targeted search precision: $x > x^*$. When $x > x^*$, which is given by $\int_x^{\bar{v}} (v - x) g_x(v) dv = s$, we have a very different situation. The participation constraint will be binding because the w^* given by $\int_w^{\bar{v}} (v - w) \frac{g(v)}{1 - G(x)} dv = s$ will be smaller than x. Hence, unless the platform compensates for the first search, consumers will not enter the market. Proposition 3 characterizes the equilibrium price when $x > x^*$.

Proposition 3. *When $n \to \infty$ and $x > x^*$, equilibrium price $p^* = x$. And there exist a finite N where the same result hold $\forall n > N$.*

To prove Proposition 3, notice the consumer's optimal search strategy is to stop at the very first shop she visited, unless she gets a matched value that is worth less than the price. The demand has a kink point at $p = x$, being perfectly inelastic when $p \leq x$. The demand is given by

$$D(p,x,n) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}(1-G(x))^k G^{n-k}(x)\frac{n}{k}(1-G_x(p))$$

Obviously, we have $p(x) \geq x$ because lowering price below x will not increase demand (notice that $G_x(v) = 0 \forall v \leq x$). Therefore, for the optimal price, we could either have the interior solution of $p^*(x) > x$ or the corner solution $p^*(x) = x$. To eliminate the interior solution, we only need $x - c > \dfrac{D(p,x,n)}{-D_p(p,x,n)}$, or intuitively, the demand is not too inelastic at x. When $n \to \infty$, we have the condition can be written as $p^*(x) = x$ if $x - c \geq \dfrac{1-G(x)}{g(x)}$. Because $\dfrac{1-G(x)}{g(x)}$ is decreasing in x, we need it to hold at x^* only, or $x^* - c \geq \dfrac{1-G(x^*)}{g(x^*)}$, which will hold by assumption that there exists some $x < x^*$ without market breakdown. The result above can be shown to hold even if n is finite but large enough, using similar arguments on the uniform convergence of $\dfrac{D(p,x,n)}{-D_p(p,x,n)}$ that we used to prove the last claim in Proposition 2.

Notice Proposition 3 is not a Diamond result. When $s \to 0$, we have $x^* \to \bar{v}$, for any $x < \bar{v}$, we still have $\lim\limits_{s\to 0} p^* = c$, that is, the equilibrium price approaches marginal cost when the search cost approaches zero. And $\lim\limits_{s\to 0}\lim\limits_{x\to \bar{v}} p^*$ depends on the convergent rate of the two limits.

At first glance, the finding that consumers do not search with highly targeted search design may look quite striking. Yet as the platforms collect more and more data and develop highly sophisticated search functionality and algorithms, this result is becoming more and more relevant. For instance, Google Now pushes relevant information to users with minimal user input. Amazon has developed its Amazon Echo, which allows consumers to talk to its chatbot (named Alexa) to place orders.[6] AI chatbots such as Siri also directly offer users a single, most relevant search result. In fact, the industry has

[6]See https://techcrunch.com/2016/07/01/alexa-orders.

discussed the future of search being "no-search"[7]. Proposition 3 shows the "no-search" result may have an implication for e-commerce: such a "no-search" future, although greatly simplifying the consumers' search process, could give sellers (and potentially platforms) immense monopoly power to extract consumer surplus. As the technology gets more sophisticated, they can extract still more surplus. In a way, precise targeted search makes consumers "royal" to relevant sellers by making additional searches unnecessary.

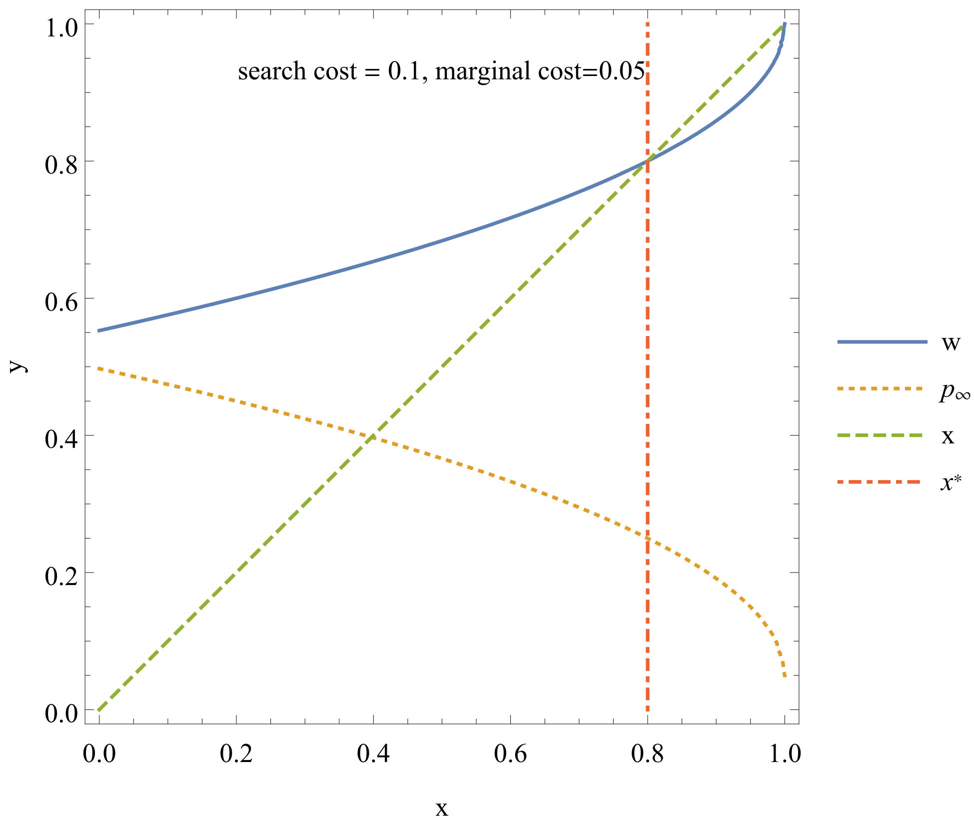

Figure 2: Equilibrium price, w versus x with uniform v

Figure 2 shows one example with uniformly distributed valuation with $\underline{v} = 0$ and $\bar{v} = 1$. As we can see, the equilibrium price decreases with precision parameter x when $x < x^*$. But if $x > x^*$, we have the price equals x. In some cases the market may not exist unless x is large enough. Figure 3 shows such an example with search cost s increased to 0.15. Now we have $w < p^*$ when x is small, which means that only when the platform

[7]For example, see http://tiny.cc/hqrlcy

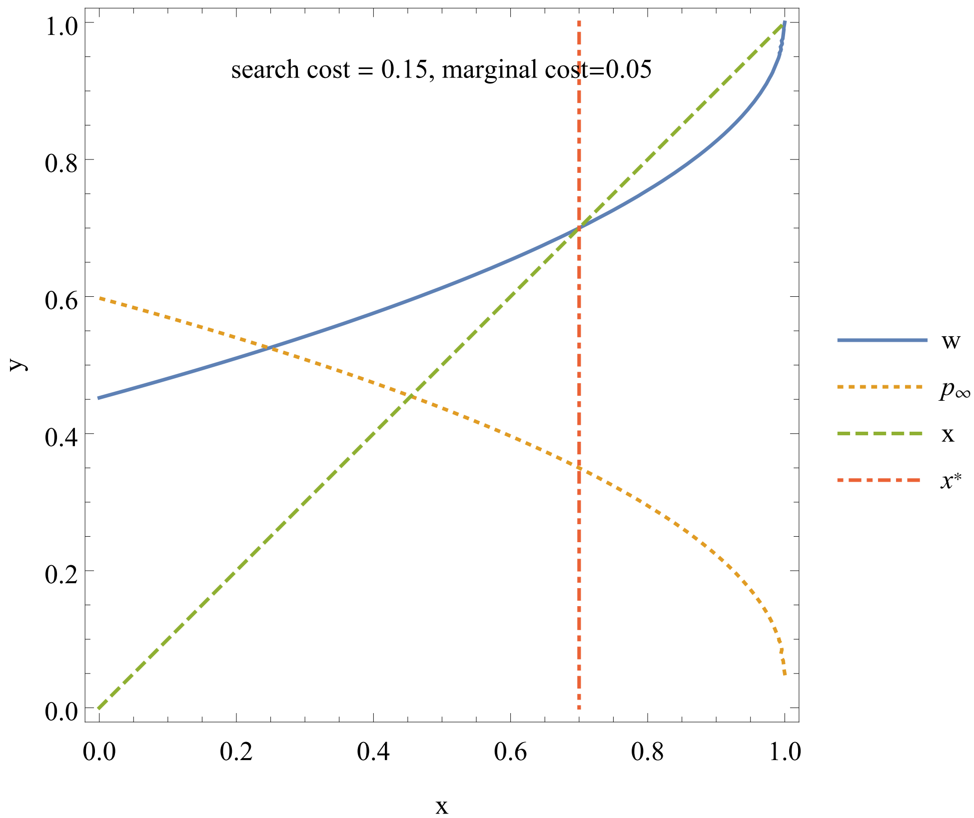

Figure 3: Equilibrium price, w versus x with uniform v

provides sufficiently relevant search results will consumers be willing to actively search.

Because of lower prices, better fits, and fewer visits, higher x will surely increase consumer surplus. In the uniform example, the consumer surplus is given by $w - p^* = 1 - c - 2\sqrt{s(1-x)}$. We can separate it into three components: the expected valuation $1 - \frac{1}{2}\sqrt{s(1-x)}$, disutility of expected search cost $-\frac{1}{2}\sqrt{s(1-x)}$, and disutility from price $-c - \sqrt{s(1-x)}$. We plot the consumer surplus and separate contributions of the three effects in Figure 4. In this particular case, the effect of saving from search cost equals that of increased valuation, and both are exactly half of the price effect.

4.3 Perfectly Targeted Search

This section explores a special case in which the platform has perfect information (or at least the ranking) of the valuations. For each consumer i, the search sequence is ordered by descending values $v_1 > v_2 > ... > v_n$. Since the v are i.i.d. across consumers,

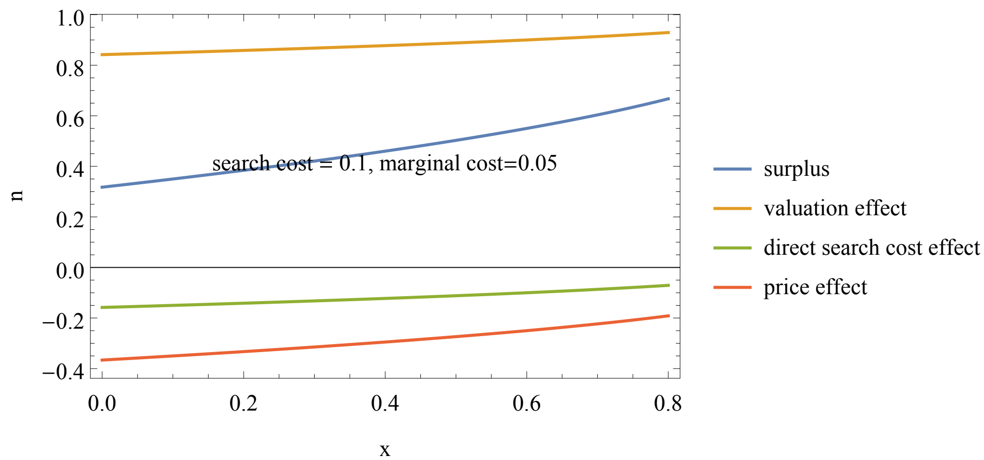

Figure 4: Changes of consumer surplus versus x with uniform v

so the ranking will be i.i.d. as well, retaining the symmetry between sellers. The results of perfectly targeted search are similar to our main model with some slightly different technical properties, and it allows analysis of a finite number of sellers.

Clearly, now the optimal search strategy of consumers is no longer stationary. In fact, in the case of (symmetric) pure pricing strategies, consumers only visit the highest ranked brand and immediately stop if the realization is satisfactory, and otherwise exit.

We introduce the order statistics of v. Let $v_{(i)}$ be the $i-$th order statistic. Therefore, $v_{(n)}$ will be the $n-$th order statistic, or the maximum value out of n brands. We also denote the distribution of $v_{(i)}$ as $G_{(i)}(v)$. Clearly, the order statistics are neither independent nor identically distributed. We consider the optimal searching strategy (under symmetric pure pricing) as follows. First, consider the participation condition, given by

$$\int_{p^*}^{\bar{v}} (v - p^*) dG_{(n)}(v) \geq s \tag{9}$$

The inequality comes from the fact that in equilibrium, any consumer will visit at most one firm: the one with the highest value. If the realized value of $v_{(n)}$ is smaller than equilibrium price, the consumer will exit; otherwise, she stops immediately with payoff $v - p^*$. The following proposition characterizes the key result of equilibrium prices in this setting:

Proposition 4. *In perfectly targeted search, there exist a unique equilibrium where the* *price is given as follows:* $p^* = c + \frac{1 - G(p^*)^n}{n[G(p^*)]^{n-1}g(p^*)}$*, with* $\lim_{n\to\infty} p^* = \bar{v}$*. Furthermore, when the number of sellers is large enough, the equilibrium price under perfectly targeted search is higher than that of random search or targeted search:* $\exists N$ *s.t.* $\forall n > N$, $p^*(n) > p^*_x(n)$ *for any* $x < \bar{v}$*.*

Proof. To begin the analysis, let us consider one firm with price p and all others stay with p^*. This seller will only have demand if it has the highest value, the probability of which is $\frac{1}{n}$. Suppose $p > p^*$. Compared to charging equilibrium price, this seller will lose some consumers with $p^* < v_{(n)} < p$. Notice that for a small deviation, we can ignore the consumers who choose to continue their visit to the second highest rated shop, with a sufficient condition $p - p^* < s$. For $p < p^*$, this seller get some extra consumers with $p^* > v_{(n)} > p$. Therefore, the demand is not inelastic as in Diamond. We derive the optimal prices by showing the demand as follows:

$$D(p, p^*, n) = \frac{L}{n}\int_p^{\bar{v}} g_{(n)}(v)dv \tag{10}$$

At equilibrium, we have $D(p^*, p^*, n) = \frac{L}{n}\int_{p^*}^{\bar{v}} dG_{(n)}(v)$. We also have its derivative with respect to price as follows:

$$D_p(p, p^*, n) = \frac{L}{n}(-g_{(n)}(p)) \tag{11}$$

Therefore, the equilibrium price is given by

$$\begin{aligned} p^* &= c - \frac{D(p^*, p*, n)}{D_p(p^*, p^*, n)} \\ &= c + \frac{\int_{p^*}^{\bar{v}} g_{(n)}(v)dv}{g_{(n)}(p^*)} \end{aligned} \tag{12}$$

Notice that $g_{(i)}(v)$ is given by

$$g_{(i)}(v) = \frac{n!}{(i-1)!(n-i)!}[G(v)]^{i-1}[1 - G(v)]^{n-i}g(v)$$

Therefore,

$$\begin{aligned} g_{(n)}(v) &= \frac{n!}{(n-1)!(1)!}[G(v)]^{n-1}g(v) \\ &= n[G(v)]^{n-1}g(v) \end{aligned}$$

Substituting into Equation 12, we have

$$p^* = c + \frac{\int_{p^*}^{\bar{v}} n[G(v)]^{n-1} g(v) dv}{n[G(p^*)]^{n-1} g(p^*)}$$
$$= c + \frac{1 - G(p^*)^n}{n[G(p^*)]^{n-1} g(p^*)} \tag{13}$$

□

When $n \to \infty$, we have $p^* \to \bar{v}$, a Diamond result in which no consumer ever enters the market.[8] When n is finite, it is possible to have p^* sufficiently small such that the market exists even if the first search is not free.

Note that a particular feature of this finding is that the equilibrium price has nothing to do with search cost s. As long as (9) holds, the equilibrium price will be above the marginal cost, below $\bar{v} - s$, and is not a function of s. It is only a function of n. The next corollary characterizes some comparative statics with respect to the number of sellers and the marginal cost.

Corollary 3. *When G is log-concave and $n > 1$, the equilibrium price under perfectly targeted search is increasing in the number of sellers and marginal cost: $\frac{\partial p^*}{\partial n} > 0$. And $\frac{\partial p^*}{\partial c} > 0$.*

Proof. See Appendix A.4 □

The last piece of Proposition 4 states that the equilibrium price of perfectly targeted search will be higher than (imperfectly) targeted search or random search. The intuition is that perfectly targeted ranking gives monopoly power to sellers with even better fit, allowing them to charge higher prices. The proof is also quite intuitive. When $n \to \infty$, we have $\lim_{n\to\infty} p^*(n) = \bar{v}$ for any distribution of v, and it is strictly higher than the equilibrium price of random search or targeted search. By continuity, the equilibrium price under perfectly targeted search will be higher than random search or targeted search for a sufficiently large n.

[8] Proof: If $p^* < \bar{v}$, according to (13), we have $p^* \to c + \frac{1}{\lim_{n\to\infty} n[G(p^*)]^{n-1} g(p^*)} = c + \infty$.

As a example, consider the case in which v is uniform such that $G(v) = \frac{v - \underline{v}}{\bar{v} - \underline{v}}$, $g(v) = \frac{1}{\bar{v} - \underline{v}}$, and $g'(v) = 0$. We have the price given by

$$p^* = c + \frac{(\bar{v} - \underline{v})^n - (p^* - \underline{v})^n}{n(p^* - \underline{v})^{n-1}}$$

Let $\underline{v} = 0,\ \bar{v} = 1$; then $p^* = c + \frac{1 - (p^*)^n}{n(p^*)^{n-1}}$. Compared to the case with random search under the same parameters, in which we have $p_r^* = c + \frac{\sqrt{2s}(1 - (p_r^*)^n)}{1 - (1 - \sqrt{2s})^n}$, we plot it along with the price under perfectly targeted in the Appendix (see Figure 7 and Figure 8 showing the relationship between p^*, n, and c).

5 Platform Design

In the previous analysis, we show that targeted search has a non-monotonic effect on equilibrium price. In particular, there exists a limit of search precision x^* that below x^*, targeted search lowers equilibrium price, and above x^*, it increases price. In this section, we consider how the platform designs its targeted search (x) under different revenue models. Broadly, the platform cares about the total trade that takes place on the marketplace (measured by gross merchandise volume, GMV) and the percentage of GMV it can monetize into its own revenue, or monetization rate. Common ways to monetize include commission and advertising. In a commission model, the platform takes a percentage fee out of each transaction. In the advertising model, the platform sells promoted slots in the search outcomes to sellers. As summarized in Table 2, most major platforms use both commission and promoted slots. In the following analysis, the commission model is presented in Section 5.1. Section 5.2 further allows consumer entry. And Section 5.3 introduces the promoted slot model.

5.1 Optimal Commission

Suppose the platform charges a percentage commission z ($z \in [0, 1]$), that is, for items with price p, sellers are receiving $(1 - z)p$ and the platform is getting zp. In terms of equi-

Table 2: Fee structure and promoted slot of major online platforms

Platform	Fee Structure	Promoted slot
Amazon	8-15% commission + listing fee	Yes
eBay	10% commission + $0.30 listing fee	Yes
Sears	$39.99 Monthly fee + 8-17% commission	Yes
Newegg	8-14% commission	Yes
Booking	12-15% basic commission	Yes
Expedia	up to 25% commission	Yes
Orbitz	20-30% commission	Yes
Hotels	20-30% commission	Yes
Hotwire	unknown markup	Yes
Kayak	unknown commission rate	Yes
Tripadvisor	pay-per-click fee	Yes
Taobao	free	Yes
Tmall	2-5% commission + annual fee	Yes
JD	1-10% commission + monthly fee	Yes
Suning	2-5% commission + annual fee	Yes

librium price, charging a percentage commission z is equivalent to a no-commission world where the marginal cost increases from c to $\frac{1}{1-z}c$. Now the optimal price is also a function of z, and $p^*(x,z)$ is increasing in z. The profit for the platform is given by $zp^*(x,z)D(x,p^*)$, where $D(x,p^*)$ is the expected demand.

We focus on the case in which the number of sellers are large, or $n \to \infty$. Because the number of sellers goes to infinity, every consumer who chooses to participate will eventually stop at some seller. There will be neither recall nor exit after consumers exhausting all options yet fail to find any satisfactory match. Therefore, $D(x,p^*) = 1\{w(x) \geq p^*\}$ will be either 1 or 0, where $w(x)$ is given by Equation 5, or $\int_{w^*}^{\bar{v}}(v-w^*)g(v)dv = (1-G(x))s$.

Let us consider the optimal z^* for any $x \leq x^*$. Intuitively, more precise search targeting will improve consumer surplus by lowering price, saving search cost, and improving expected fit. This allows the platform to increase the commission rate to extract the surplus by inducing a higher equilibrium price.

Lemma 2. *When $n \to \infty$, the optimal commission rate $z^*(x)$ is increasing in x when $x < x^*$. The equilibrium price at the optimal commission rate will be $p^*(x,z^*) = w(x)$, which is increasing in x.*

Proof. Given x, changing z^* only changes the equilibrium price and consumers' participation constraint, but not their optimal search strategy. As long as $w(x) > p^*$, the demand is constant. Therefore, the platform will set z^* such that $w(x) = p^*(z^*,x)$. As we've shown, in this case, the price is given by $p^*(x) = \frac{1}{1-z}c + \frac{1-G(w^*)}{g(w^*)}$. Therefore we have that the optimal z^* is given by $w^* = \frac{1}{1-z}c + \frac{1-G(w^*)}{g(w^*)}$. Because w^* is increasing in x, we can easily show that z^* is also increasing in x: the LHS is increasing in w^* whereas the RHS is decreasing in w^*. □

If the platform can choose the precision of targeted search x without transfer to consumers, the following proposition characterize the optimal search design.

Proposition 5. *If the platform cannot transfer to consumers, the optimal strategy for the platform when $n \to \infty$ is to set $x = x^*$ and charge z such that $p^* = x^*$, or z^* is given by $x^* = \frac{1}{1-z}c + \frac{1-G(x^*)}{g(x^*)}$. There also exist a finite N such that $\forall n > N$ the result holds.*

The intuition is clear: higher x (when $x < x^*$) lowers prices, gives consumer better match value and saves on search cost. All three effects increase consumer surplus. The platform should increase the commission rate to capture the surplus. It therefore predicts that targeted search under commission model increases both the monetization rate (by higher z) and the GMV (by higher price). Both effects contribute to the revenue of the platform. However, the situation becomes very different once targeted search is sufficiently precise, that is, $x > x^*$. As our previous result shows, at x^* there is a jump in the equilibrium price. Suddenly, a lot of consumer surplus shifts from consumers to sellers. If the platform cannot compensate for the first search, a market failure with zero demand will occur. Proposition 5 shows that the platform want to limit the precision of targeted search to avoid sellers having monopoly power, even if improving it is costless.

Furthermore, even if the platform can compensate consumers to make their first search free, it may still choose $x = x^*$ and charge the same commission rate in Proposition 5. See the following proposition that shows the platform will want to restrict its precision of targeted search.

Proposition 6. *If the platform can transfer to consumers, there exist a upper bound of $x' > x^*$ such that if the platform can choose $x \in [\underline{v}, x']$, the optimal strategy for the platform when $n \to \infty$ is to set $x = x^*$ and charge z such that $p^* = x^*$, or z^* is given by $x^* = \frac{1}{1-z}c + \frac{1-G(x^*)}{g(x^*)}$. There also exist a finite N such that $\forall n > N$ the result holds.*

To see the proof, suppose that at x', the optimal commission rate is z' and price is p'. Clearly $z' > z^*$ and $p' > p^*$. The intuition of Proposition 6 is that the cost of compensating for first search is s per consumer, yet the potential benefit of increase in commission rate and price is given by $z'p' - z^*p^*$. Since both z and p are increasing continuously in x, there must exist a x' such that $z'p' - z^*p^* = s$, and for which any $\hat{x} \in (x^*, x')$ we have $\hat{z}\hat{p} - z^*p^* < s$, or equivalently the platform profit is strictly smaller compare to that at $x = x^*$.

5.2 Optimal Commission with Consumer Entry

One limitation of the previous analysis is that either all or no consumers enter the market, depending on w^* and p^*. Therefore, we cannot capture the idea that once consumer surplus increases, more consumers are willing to adopt the platform, either due to competition between platforms or market expansion. This section extends the model by generalizing the outside option from zero to a non-degenerated distribution to allow consumer entry.

Suppose the outside option follows some distribution with density $f(v)$ that is smooth and bounded. The corresponding CDF is given by $F(v)$. We assume $f(v)$ is log-concave to ensure that the optimal commission is unique, as is standard in the pricing literature. Notice that introducing heterogeneity in outside option does not change consumers' optimal search strategy if they enter. Consumers still search sequentially with stationary reservation value w as long as they choose to participate. It only changes the number of consumers who will enter the market, which is given by $F(w^* - p^*)$. Therefore, the platform is choosing x and z to maximize its profit when the number of sellers is large so that no consumer exit after choosing to participate:

$$\max_{x,z} \pi(x, z) = F(w^*(x) - p^*)zp^*$$

We first note that allowing consumer entry does not affect the equilibrium price. And we can show that when $x > x^*$ the market equilibrium price could be discontinuously high under optimal commission rate, making it not desirable for platforms, even if improving the precision of targeted search has not cost.

3 shows that the optimal commission rate increases in x. Intuitively, as precision of targeted search x increases, consumer surplus $w^*(x)$ increases and the equilibrium price p^* decreases. Therefore $F(w^*(x) - p^*)$ will be increasing, implying more consumers will enter the market. This effect should increase the incentive for the platform to raise commission rates, as the marginal benefit of raising commission rate increases.

Lemma 3. *The optimal commission rate z^* defined by $\max_z F(w^*(x) - p^*)zp^*$ is unique and increasing in x when $x \leq x^*$ and $n \to \infty$.*

Proof. See Appendix Section A.5. □

Therefore, when targeted search is not too precise ($x \leq x^*$), the platform's commission rate will be increasing in the precision x. Similarly with the case without consumer entry, this result implies a increasing monetization rate as the platform develops advanced search technologies. The next proposition shows our key finding on platform search design with consumer entry. The intuition is similar to Proposition 5: when the precision is below x^*, the platform will always be better off by choosing a higher precision, since it leads to both market expansion and higher monetization. However, once the precision reached x^*, the price could become too high, leading to dead-weight loss due to the monopoly power of sellers that discourages consumer entry.

Proposition 7. *If the platform cannot transfer to consumers, the optimal strategy for the platform when $n \to \infty$ is to set $x = x^*$ and charge z defined by 3, as long as the profit at $x = x^*$ (which is given by $\max_z F(w^*(x^*) - p^*)zp^*$) is higher than that the highest profit for any $x > x^*$ (which is given by $\max_{x>x^*,z} \pi(x, z) = F(w^*(x) - p^*)zp^*$).*

Proof. The proof is quite intuitive. We first show that the platform's profit $\pi^*(x)$ is increasing in x when $x < x^*$. Suppose that we have $x' > x$. We have that $\pi^*(x') = \max_z F(w^*(x') - p^*(x', z))zp^*(x', z)$, and $\pi^*(x) = \max_z F(w^*(x) - p^*(x, z))zp^*(x, z)$. Let $z_1 \in \arg\max F(w^*(x) - p^*(x, z))zp^*(x, z)$. We can find a $z_2 > z_1$ such that $p^*(x', z_2) = p^*(x, z_1)$. And therefore $\pi(x', z_2) = F(w^*(x') - p^*(x', z_2))z_2 p^*(x', z_2) > F(w^*(x) - p^*(x, z_1))z_1 p^*(x, z_1)$ since $z_2 > z_1$ and $x' > x$. Yet by definition $\pi^*(x') = \max_z F(w^*(x') - p^*(x', z))zp^*(x', z) \geq \pi(x', z_2)$, therefore $\pi^*(x') > \pi^*(x)$ as desired.

For the case when $x > x^*$, we show that there are cases in which the profit will be discontinuously lower. This is possible when the price becomes discontinuously higher at x^*. Suppose that the optimal commission rate at x^* is given by z^*, and $p^*(x^*, z^*) < x$. At any $x > x^*$, the price is either an interior solution (which is strictly larger than x) or the corner solution x, both will be strictly and discontinuously higher than $p^*(x^*, z^*)$. Therefore, we have $\lim_{x \to x^{*+}} p^*(x^*, z^*(x)) > p^*(x^*, z^*)$, and more importantly, its revenue will be discontinuously lower, that is, $\lim_{x \to x^{*+}} \max_{x,z} \pi(x, z) < \pi(x^*, z^*)$. If the difference in revenue is high enough, that is, $\max_{x>x^*,z} \pi(x,z) = F(w^*(x) - p^*)zp^* < \pi(x^*, z^*)$, the platform will be better off by choosing a targeted search precision x^*. $\square$

Example 5.1. Suppose that $F(v)$ is uniform on $[-\rho, \rho]$. We have $F(w^*(x)-p^*) = \dfrac{w^*(x) - p^* - \rho}{2\rho}$ if $w^*(x)-p^* \in [-\rho, \rho]$. For any x, optimal commission z^* is given by $\max_z [w^*(x)-p^*+\rho]zp^*$. See Figure 5 for optimal commission rates under different parameters.

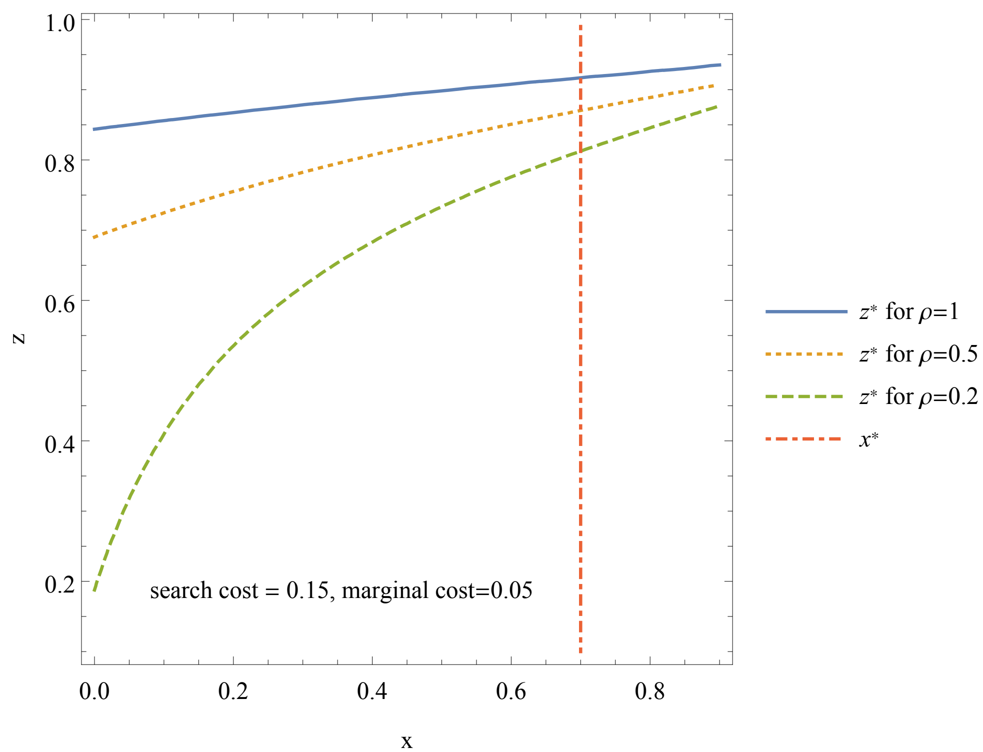

Figure 5: Optimal commission rate with consumer entry

5.3 Promoted Slots Model

Selling promoted slots (also known as sponsored links, ad slots, etc.) to sellers has become an increasingly common way for platforms to monetize the trade. This section models how the promoted slots model interacts with targeted search technology.

To set up the model, consider the platform choosing a set of slots within the search outcomes to be promoted slots for all search queries. We denote the set to be $S = \{i :$

i-th slot being a promoted slot$\}$. For instance, if the platform chooses to make the top two slots as promoted slots, we have $S = \{1, 2\}$.

The timing is as follows. First the platform chooses the set of promoted slots S. The platform then sells the slots through some mechanism such as auctions, such that sellers are indifferent between getting the slot or not. The sellers bid for the slots after observing S. Afterwards, consumers enter and the market clears.

We assume consumers still search sequentially according to the selection of sellers and the ranking that the platform provides. They are fully aware of whether the next seller is an organic search result or a promoted one. And they can choose to skip the promoted seller at no cost. [9]

One important assumption is how much sellers know when they choose to compete for a slot. In this section, we will assume all sellers, when they choose how much to bid for each keyword, have exactly the same information as the platform, whether their valuation is higher than each $x_i \in X$ or not. In particular, they know whether for the keyword and targeted consumers, they are relevant sellers (with value above x) or not. Consistently with the previous analysis, we focus on the case with a large number of sellers.

Lemma 4. *When $n \to \infty$, the equilibrium price for promoted slots are given by $p_i^* = p^* = c + \dfrac{1 - G(w^*)}{g(w^*)}$, $\forall\, i \in S$. The market share is given by $[1 - G_x(w^*(x))]G_x^{i-1}(w^*(x))$, $\forall\, i \in S$.*

Proof. As Armstrong et al. (2009) pointed out, when $n \to \infty$ and $S = 1$, we have that $p_i^* = p^* = c + \dfrac{1 - G(w^*)}{g(w^*)}$, $\forall i \in S$. That is, the promoted sellers have exactly the same price as the sellers in the organic search results. The intuition is as follows: the first (or ith) "prominent" seller faces exactly the same elasticity as the rest, because consumers follow the same search rules with i.i.d. distribution of matched values. The only difference is that the number of consumers is higher for prominent sellers who rank before organic sellers. It leads to a larger market share for the top promoted sellers that given by $1 - G_x(w^*(x))$. As shown in Corollary 2, it is increasing in x when $g(v)$ is weakly concave. $\square$

[9]This assumption makes no difference in this case, because consumers will rationally choose not to skip the promoted sellers, who have the same value distribution and the same or lower prices.

The following proposition characterize the main result for promoted slots model where we focus on a special case in which the platform sells the top k slots as promoted slots of slots: $S = 1, 2, ..., k$.

Proposition 8. *The platform revenue from selling top k promoted slots will be increasing in x for $x \leq x^*$ when $G(v)$ is weakly concave, and the platform will choose $x \leq x^*$ as the optimal targeted search precision to maximize its revenue.*

Notice that the revenue for the top k slots are given by $\pi(x,k) = (p^*(x) - c)\sum_{i=1}^{k}[1 - G_x(w^*(x))]G_x^{i-1}(w^*(x)) = (p^*(x) - c)[1 - G_x^k(w^*(x))]$ when $x \leq x^*$ and zero when $x > x^*$. The profit margin $p^*(x) - c$ is decreasing in x while the quantity $1 - G_x^k(w^*(x))$ is increasing in x. The optimal design of targeted search precision is given by $\arg\max_{x \in [\underline{v}, x^*]}(p^*(x) - c)[1 - G_x^k(w^*(x))]$.

This proposition shows promoted sellers receive higher market shares with targeted search. The intuition is that increasing targeting precision x has two effects on promoted sellers: first it filters out less relevant sellers, making consumers more willing to continue search; second, it improves the *targetability* of sellers. Sellers can use more options to select which consumer segment to target, therefore increase the fit between consumers and sellers. This effect is the opposite of reducing search cost, because reducing search costs will increase w^*, hence making consumers less likely to stop at any seller, including promoted ones. Notice that with the use of quality score and generalized auction, the sellers do not need to explicitly know which keywords or consumer segments to target. The platform, by assigning different quality scores, essentially informs sellers whether they are good fit or not. Again, this results shows that in the promoted slots model, targeted search will improve the monetization rate of the platform.

6 Empirical Study

In this section I use a large and unique Taobao data set to show some suggestive evidence of the theory. I find correlations between a construct of targeted search precision and market price, as well as correlation between precision and market share of promoted sellers in the top slot. Readers should interpret the findings as correlations instead of causality, as in this setting the search precision is not exogenous.

The data consist of search outcomes of over 140,000 keywords, scraped from Taobao during February to April, 2016. The keywords are from Taobao's most searched keywords for each categories. For each keyword, I have the search outcomes (ranked by the default ranking) of the first five pages or the top 224 sellers. In total I have over 31 million observations. And each data point has the price and quantity sold, as well as shipping fee, sellers' attributes including its reputation level, good rating percentage, three dynamic ratings, its major industry, and whether it is a Taobao or TMall seller. For each keyword I also have the number of search outcomes as well as its three level categories. For instance, an iPad case keyword belongs to "3C Digital"-"Tablet Accessory"-"Tablet Case". See Table 3 for the summary statistics at keyword level.

Table 3: Summary Statistics of Keyword Search Outcomes

Variable	Obs	Mean	Std. Dev.	P25	P50	P75
Total # Search-Outcome	141974	434141.1	3236884	11129	42179	155304
Average Quantity Sold	142108	346.116	643.044	66.253	181.9	417.417
Average Price	142133	520.361	13045.03	44.539	91.278	179.345
Average Reputation Level	142118	10.692	1.779	9.717	10.839	11.78
Average Favorable Rate %	142117	99.015	.985	98.853	99.2	99.458
Average Dynamic Rating 1	142117	4.844	.078	4.823	4.852	4.881
Average Dynamic Rating 2	142117	4.848	.075	4.828	4.853	4.881
Average Dynamic Rating 3	142117	4.824	.075	4.803	4.829	4.857
Keyword Length	142133	5.891	2.476	4	6	7

I construct a measure of targeted search precision using the number of search outcomes for each keyword. The intuition is that a specific keyword ("iPad Air 2 Thin Case" compared to "iPad Case") should lead to a small number of search outcomes, and the search should be more targeted. I include the 3rd-Level category fixed effect to control for differences in market sizes across different categories. I also take the log to account for its skewness which greatly increases the model fit.

The first finding is the correlation between the share of promoted seller and search precision. The theory predicts that as search precision increases (in my empirical setting, as the number of search outcomes decreases), the market share of promoted seller increases. I construct the market share by taking the ratio of the quantity sold of promoted seller over the average quantity sold of all sellers for the same search query. As in Table 4, both the share of the first seller (who pays to get promoted) and the share of the first page (first 48 sellers, including one promoted and 47 organic) are negatively correlated with the number of search outcomes. This is a suggestive evidence that is consistent with the model prediction.

Table 4: Correlation between shares of promoted or higher ranked sellers and the number of search outcomes

	First Shop Share		First-Page Share	
Log(# Search Outcome)	-0.738***	-0.814***	-0.0535***	-0.0536***
	(0.0345)	(0.0347)	(0.00115)	(0.00109)
N	141930	141930	141930	141926
3rd Category F.E.	Yes	Yes	Yes	Yes
Average Shop Attributes[1]	Yes	Yes	Yes	Yes
First Shop Attributes[2]	No	Yes	No	No
First-page Shop Attributes[3]	No	No	No	Yes

Standard errors in parentheses. $^{*}\, p < 0.05$, $^{**}\, p < 0.01$, $^{***}\, p < 0.001$.

[1] Including average reputation level, good rating percentage, and three dynamic ratings.

[2] Including the first seller's reputation level, good rating percentage, and three dynamic ratings.

[3] Including the first page sellers' average reputation level, good rating percentage, and three dynamic ratings.

The theory also predicts that increasing targeted search precision (less number of search outcomes) leads to lower prices when the market is not too precise, and the opposite when targeted search is sufficiently precise. Figure 6 shows the relationship between the measure of targeted search precision [10] and the price [11] using local linear fit with 95% confidence intervals. It is very much predicted by the model: when the precision is not too high, increasing targeted search precision is associated with decreasing prices, while the reverse is true when targeted search precision is sufficiently high. Table 5 shows that both correlations are significant in the regression of price over the log number of search outcomes in two subsamples, with high and low search precision as defined before.

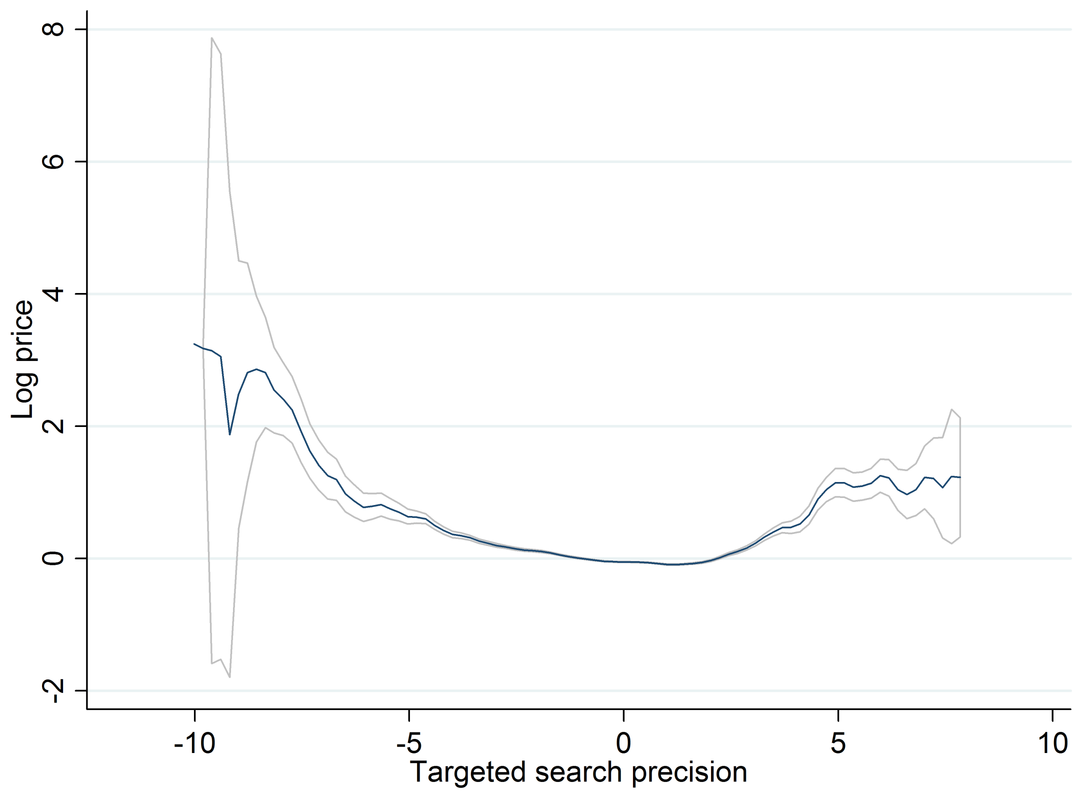

Figure 6: Log prices and targeted search precision

[10] Defined by the negative of the residual of log number of search outcomes regressing on average shop attributes and 3-rd category fixed effects.

[11] Defined by the residual of log prices regressing on average shop attributes and 3-rd category fixed effects.

Table 5: Correlation between prices and the number of search outcomes

	Average Price		
	Full Sample	Low Precision[1]	High Precision[1]
Log(# Search Outcome)	0.0426***	0.121***	-0.118***
	(0.00127)	(0.00238)	(0.00354)
N	141959	70973	70986
3rd Category F.E.	Yes	Yes	Yes
Average Shop Attributes[2]	Yes	Yes	Yes

Standard errors in parentheses. $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

[1] Defined by whether the residual of log number of search outcomes regressing on average shop attributes and 3-rd category fixed effects being smaller than its median (-0.164) or not.

[2] Including average reputation level, good rating percentage, and three dynamic ratings.

7 Discussion and Concluding Remarks

This study focuses on the role of targeted search in online marketplaces. It is a micro model of search engines and its application in platform design questions. There are several new findings. First, targeted search has a non-monotonic effect on the equilibrium price. It decreases the price by intensifying competition when targeting precision is not too high. But when precision is high enough relative to the search cost, targeted search could lead to monopoly (high) prices. The discontinuity in prices is driven by the assumption that all buyers have the same search cost. If we allow some heterogeneity in search cost, we could get a smooth relationship between equilibrium price and search precision. The intuition is when targeting precision is not too high, the competition effect dominates, and targeted search will decrease the equilibrium price. However, as the precision increases, more and more consumers do not want to actively search, because the benefits of search do not compensate their search costs. This effect will be more dominating as search precision becomes sufficiently high, makes the demand more and more inelastic, and leading to higher price. Empirically, we may find some evidence supporting the competition effect across broad categories of goods. The theory also predicts that the monopoly effect of targeted search may be easier to find in platforms with highly targeted search, and for product markets with low differentiation relative to search cost. With the development of search technology, in particular, the artificial intelligence and personalize data, this finding will be more and more relevant.

The second finding is the optimal platform search design and platform contracts. In both commission and promoted slots models, I find that the platform may want to limit its precision even if improving it is costless. This finding is closely related to the non-monotonicity of how targeted search affects the equilibrium price. The monopoly effect of targeted search could lead to either market failure or dead-weight loss, therefore limiting how much surplus the platform can extract using either commission or promoted slots. This finding provides a strategic reason for the platform to protect consumer privacy, as it shows utilizing more consumer data does not necessarily benefit the platform. Also, in both models, I find targeted search having a positive effect on platform revenue when the precision is below a threshold, particularly through increase in monetization rate. We are indeed observing the phenomenon in the real world. For example, the monetization rate of Alibaba has been steadily increasing for its mobile consumers, one reason being that richer data from mobile allows Alibaba to better target its consumers.

This study provides a framework of targeted search that can be applied to a wide range of research topics. First, one can be investigate how targeted search interacts with not only platform, but also product design problems. In particular, does targeted search encourage niche design or main-stream design? Second, this study focuses on horizontal differentiation. How targeted search could affect a vertically differentiated market, or a market with both vertical and horizontal differentiation, remains a open question. Third, if sellers can manipulate the search outcomes by search engine optimization, it could also change the implications of targeted search. Finally, although this study is motivated by the development of search technology in online platforms, its intuition could be applied to other intermediaries such as retailers and recruitment platforms.

References

Agichtein, E., Brill, E., and Dumais, S. (2006). Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '06, pages 19–26, New York, NY, USA. ACM. 1

Anderson, S. P. and Renault, R. (1999). Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model. *The RAND Journal of Economics*, pages 719–735. 2

Arbatskaya, M. (2007). Ordered search. *The RAND Journal of Economics*, 38(1):119–126. 2

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691. 2

Armstrong, M., Vickers, J., and Zhou, J. (2009). Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233. 2, 5.3

Armstrong, M. and Zhou, J. (2011). Paying for prominence*. *The Economic Journal*, 121(556):F368–F395. 2

Athey, S. and Ellison, G. (2011). Position auctions with consumer search. *The Quarterly Journal of Economics*, page qjr028. 2

Bar-Isaac, H., Caruana, G., and Cuñat, V. (2012). Search, design, and market structure. *American Economic Review*, 102(2):1140–1160. 2

Chen, Y. and He, C. (2011). Paid placement: Advertising and search on the internet. *The Economic Journal*, 121(556):F309–F328. 2

Chen, Y. and Yao, S. (2016). Sequential search with refinement: Model and application with click-stream data. *Management Science, forthcoming*. 2

De Corniere, A. (2016). Search advertising. *American Economic Journal: Microeconomics, forthcoming.* 2

De Corniere, A. and De Nijs, R. (2016). Online advertising and privacy. *The RAND Journal of Economics*, 47(1):48–72. 2

Dukes, A. and Liu, L. (2015). Online shopping intermediaries: The strategic design of search environments. *Management Science*. 2

Dumais, S., Peltonen, K., Gupta, A., Brunell, B., Gates, W., Flake, G., Naam, R., Horvitz, E., Huang, X., Platt, J., et al. (2008). Context-based search, retrieval, and awareness. US Patent App. 11/426,981. 1

Eliaz, K. and Spiegler, R. (2011). A simple model of search engine pricing*. *The Economic Journal*, 121(556):F329–F339. 2

Ghose, A., Ipeirotis, P. G., and Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60(7):1632–1654. 2

Hagiu, A. and Jullien, B. (2011). Why do intermediaries divert search? *The RAND Journal of Economics*, 42(2):337–362. 2

Hagiu, A. and Jullien, B. (2014). Search diversion and platform competition. *International Journal of Industrial Organization*, 33:48–60. 2

Kuksov, D. (2004). Buyer search costs and endogenous product design. *Marketing Science*, 23(4):490–499. 2

Liu, T.-Y. (2009). Learning to rank for information retrieval. *Found. Trends Inf. Retr.*, 3(3):225–331. 1

Nuzzi, F. A. (2015). Search system utilizing purchase history. US Patent 9,105,029. 1

Rochet, J.-C. and Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029. 2

Rochet, J.-C. and Tirole, J. (2006). Two-sided markets: a progress report. *The RAND journal of economics*, 37(3):645–667. 2

Rysman, M. (2009). The economics of two-sided markets. *The Journal of Economic Perspectives*, 23(3):125–143. 2

Sriram, S., Manchanda, P., Bravo, M. E., Chu, J., Ma, L., Song, M., Shriver, S., and Subramanian, U. (2015). Platforms: a multiplicity of research opportunities. *Marketing Letters*, 26(2):141–152. 2

Thirumalai, S. (2013). Ranking item attribute refinements. US Patent 8,620,891. 1

Ursu, R. M. (2016). The power of rankings: Quantifying the effects of rankings on online consumer search and choice. *Available at SSRN 2729325*. 2

Wolinsky, A. (1986). True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, pages 493–511. 2, 3, 4.1, 4.1, 4.2

Yang, H. (2013). Targeted search and the long tail effect. *The RAND Journal of Economics*, 44(4):733–756. 2

A Appendix

A.1 Proof to Lemma 1

Proof. $\int_x^{\bar{v}}(v-x)g_x(v)dv = s$ can be written as $\int_x^{\bar{v}}(v-x)g(v)dv = [1-G(x)]s$. Denote the LHS as $f_1(x) = \int_x^{\bar{v}}(v-x)g(v)dv$ and the RHS as $f_2(x) = [1-G(x)]s$. Both will be smooth and decreasing in x since $f_1'(x) = G(x) - 1 \leq 0$ and $f_2'(x) = -g(x)s < 0$. Notice $f_1(\underline{v}) > f_2(\underline{v})$ by assumption (ii) of s being not too large, $f_1(\bar{v}) = f_2(\bar{v})$ (both equal zero), and $f_1'(\bar{v}) = 0 < f_2'(\bar{v})$, by continuity of both functions, there exists at least one x^* such that $f_1(x^*) = f_2(x^*)$. To prove uniqueness, take the smallest of all such x^* that satisfy $f_1(x^*) = f_2(x^*)$. Notice that $\dfrac{f_1'(x)}{f_2'(x)} = \dfrac{1-G(x)}{g(x)s}$. Since $g(v)$ is log concave, the hazard rate $\dfrac{g(x)}{1-G(x)}$ will be monotone increasing, therefore we have at most one $\hat{x}$ such that

$\frac{f_1'(\hat{x})}{f_2'(\hat{x})} = 1$. Therefore, there does not exist any $x^* < x'^* < \bar{v}$ such that $f_1(x'^*) = f_2(x'^*)$, otherwise there will be at least one $\hat{x_1} \in (x^*, x'^*)$ and one $\hat{x_2} \in (x'^*, \bar{v})$ both satisfying $\frac{f_1'(x)}{f_2'(x)} = 1$ by mean value theorem, violating the monotone hazard rate property of log concave $g(v)$. Also, from the analysis above, it is clear that the second condition can be relaxed to that $\exists x_0 \in [\underline{v}, \bar{v}]$ such that $\int_{x_0}^{\bar{v}}(v - x_0)g(v)dv > [1 - G(x_0)]s$.

To show that when $x > x^*$, we have $\int_x^{\bar{v}}(v-x)g_x(v)dv < s$, notice that $f_1'(\bar{v}) = 0 < f_2'(\bar{v})$, therefore when $x \in (x^*, \bar{v})$ we have $f_1(x) < f_2(x) \to \int_x^{\bar{v}}(v - x)g_x(v)dv < s$. Since x^* is unique, given $f_1(\underline{v}) > f_2(\underline{v})$, whenever $x < x^*$ we have $f_1(x) > f_2(x)$. □

A.2 Proof to Corollary 2

Proof. $G_x(w) = \int_x^w g_x(v)dv = \frac{1}{1 - G(x)} \int_x^w g(v)dv$. Furthermore we have $\frac{\partial w}{\partial x} = \frac{g(x)s}{1 - G(w)}$ Therefore, we have

$$\frac{\partial G_x(w)}{\partial x} = \frac{g(x)}{1 - G(x)}\left[\frac{G(w) - 1}{1 - G(x)} + \frac{g(w)s}{1 - G(w)}\right] \tag{14}$$

$$= \frac{g(x)s}{1 - G(x)}\left[\frac{G(w) - 1}{\int_w^{\bar{v}}(v - w)g(v)dv} + \frac{g(w)}{1 - G(w)}\right] \tag{15}$$

$$= \frac{g(x)s}{(1 - G(x))(1 - G(w)) \int_w^{\bar{v}}(v - w)g(v)dv}\left[g(w)\int_w^{\bar{v}}(v - w)g(v)dv - (1 - G(w))^2\right] \tag{16}$$

The second equality is given by substituting $\frac{1}{1 - G(x)} = \frac{s}{\int_w^{\bar{v}}(v - w)g(v)dv}$. Let us denote $F(w) = g(w)\int_w^{\bar{v}}(v - w)g(v)dv - (1 - G(w))^2$. To show that $F(w) \leq 0$, notice that $F(\bar{v}) = 0$, therefore $F'(w) \geq 0 \ \forall \ w \in [w(0), \bar{v}]$ is a sufficient condition. $F'(w) = g'(w)\int_w^{\bar{v}}(v - w)g(v)dv + 2g(w)(1 - G(w))$. This is clearly positive if $g(w)$ is concave ($g'(w) \geq 0$). □

A.3 Proof to Proposition 2

The following proof shows the explicit form of $\lim_{n\to\infty} p^*(x)$ and its comparative static with respect to x.

Proof. Notice that $p^* - c$ is the ratio of two integrals. We let $f_1(n) = \int_k nD(p^*, p^*, k)\Phi_n(k)dk+$

ϵ, and $f_2(n) = \int_k nD_p(p^*, p^*, k)\Phi_n(k)dk + \epsilon_p$, in which $\Phi_n(k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu)^2}{2\sigma^2}}$ where $\mu = n(1 - G(x))$ and $\sigma = \sqrt{n(1 - G(x))G(x)}$. We can denote the price as $p^* - c = \frac{f_1(n)}{f_2(n)}$. To get $\lim_{n\to\infty} p^* - c$, if we can show that both $\lim_{n\to\infty} f_1(n)$ and $\lim_{n\to\infty} f_2(n)$ exist, as well as $\lim_{n\to\infty} f_2(n) > 0$, we have $\lim_{n\to\infty} p^* - c = \frac{\lim_{n\to\infty} f_1(n)}{\lim_{n\to\infty} f_2(n)}$.

Consider the interval $k \in [\mu - n^{1/4}\sigma, \mu + n^{1/4}\sigma]$. Clearly, as $n \to \infty$, the probability that $k \in [\mu - n^{1/4}\sigma, \mu + n^{1/4}\sigma]$ goes to one. Now we consider the value of $nD(p^*, p^*, k)$. It is bounded for all k and therefore,

$$\lim_{n\to\infty} \int_{-\infty}^{\mu - n^{1/4}\sigma} nD(p^*, p^*, k)\Phi_n(k)dk + \lim_{n\to\infty} \int_{\mu + n^{1/4}\sigma}^{\infty} nD(p^*, p^*, k)\Phi_n(k)dk = 0 \tag{17}$$

Also notice that when $k \in [\mu - n^{1/4}\sigma, \mu + n^{1/4}\sigma]$, we have $\frac{k}{n} \in [(1 - G(x)) - n^{-1/4}G(x)(1 - G(x)), (1 - G(x)) + n^{-1/4}G(x)(1 - G(x))]$, and the interval approaches $1 - G(x)$ as n increases. We then have

$$\begin{aligned}\lim_{n\to\infty} \int_{\mu - n^{1/4}\sigma}^{\mu + n^{1/4}\sigma} nD(p^*, p^*, k)\Phi_n(k)dk + \epsilon &= \frac{1}{1 - G(x)} \lim_{k\to\infty} kD(p^*, p^*, k) \\ &= \frac{1}{1 - G(x)} \end{aligned} \tag{18}$$

Combining Equation 17 and Equation 18, we have that $\lim_{n\to\infty} f_1(n) = \frac{1}{1 - G(x)}$. Similarly, we apply the same methods to $f_2(n)$, we will get $\lim_{n\to\infty} f_2(n) = \frac{1}{1 - G(x)} \frac{g_x(w^*)}{1 - G_x(w^*)}$. which is clearly bounded away from zero. Therefore, we have $\lim_{n\to\infty} p^*(x) = c + \frac{\lim_{n\to\infty} f_1(n)}{\lim_{n\to\infty} f_2(n)} = c + \frac{1 - G_x(w^*)}{g_x(w^*)}$ as desired.

The last piece can be derived from the simple fact that $\frac{\partial p^*}{\partial x} = \frac{\partial p^*}{\partial w^*}\frac{\partial w^*}{\partial x}$, and $\frac{\partial w^*}{\partial x} > 0$ whereas $\frac{\partial p^*}{\partial w^*} < 0$ due to log concavity of $g(v)$. □

And the following proof shows the second claims in Proposition 2: there exists a finite number n such that $\frac{\partial p^*(n)}{\partial x} < 0$ when $x < x^*$.

Proof. By the convergence argument (we need uniform convergence of $\partial p^*(n, x)$ to $p^*(\infty, x)$ which is strictly negative), it is easy to see that the result of Proposition 2 can be extended to large but finite n. Since $p^*(\infty, x)$ is decreasing in x, if $p^*(n, x_1) < p^*(n, x_2)$ for some $x_1 < x_2$ for all n, we can let $\epsilon = \frac{1}{2}[p^*(n, x_1) - p^*(n, x_2)]$, and since $p^*(\infty, x_1) > p^*(\infty, x_2)$, as well as that there exist some N such that for all $n > N$ we have $|p^*(n, x) - p^*(\infty, x)| < \epsilon$, there is a contradiction.

To show the uniform convergence of $p^*(n, x)$ to $p^*(\infty, x)$, notice that we only need $p^*(n, x)$ being equicontinuous in x when $x \in [\underline{v}, x^*]$ (it is differentiable with uniformly bounded derivatives with $g(v)$ being smooth).

We only need to show that $\frac{\partial p^*(n)}{\partial x}$ is uniformly bounded. Notice that $g(v)$ is sufficiently smooth (with bounded first and second derivatives). The details are as follows:

To get $\frac{\partial p^*}{\partial x}$, we use implicit function theorem. Let

$$f(p^*, x) = p^* - c + \frac{\sum_{k=0}^n \frac{n!}{k!(n-k)!}(1 - G(x))^k G^{n-k}(x)[1 - G_x(p^*)^k]}{\sum_{k=0}^n \frac{n!}{k!(n-k)!}(1 - G(x))^k G^{n-k}(x) \left\{ -g_x(w^*) \frac{1 - G_x(w^*)^k}{1 - G_x(w^*)} + \int_{p^*}^{w(p^*)} G_x(v)^{k-1} g_x'(v) dv \right\}}.$$

We have $\frac{\partial p^*}{\partial x} = -\frac{\partial f / \partial x}{\partial f / \partial p^*}$. To show that it is uniformly bounded, we only need to show that $\partial f / \partial x$ is uniformly bounded and $\partial f / \partial p^*$ is uniformly bounded away from zero.

We let

$$\Delta_1 = \sum_{k=0}^n \frac{n!}{k!(n-k)!}(1 - G(x))^k G^{n-k}(x)[1 - G_x(p^*)^k]$$
$$\Delta_2 = \sum_{k=0}^n \frac{n!}{k!(n-k)!}(1 - G(x))^k G^{n-k}(x) \left\{ -g_x(w^*) \frac{1 - G_x(w^*)^k}{1 - G_x(w^*)} + \int_{p^*}^{w(p^*)} G_x(v)^{k-1} g_x'(v) dv \right\}$$

Then we have $f(p^*, x) = p^* - c + \frac{\Delta_1}{\Delta_2}$.

Firstly, notice that both Δ_1 and Δ_2 are uniformly bounded. The numerator Δ_1 is simply the mean of $[1 - G_x(p^*)^k]$ (with k binomial distributed) that bounded in $(0, 1)$. The denominator Δ_2 is similarly the mean of $-g_x(w^*) \frac{1 - G_x(w^*)^k}{1 - G_x(w^*)} + \int_{p^*}^{w(p^*)} G_x(v)^{k-1} g_x'(v) dv$ which lives in $(\frac{-g_x(w^*)}{1 - G_x(w^*)} + \int_{p^*}^{w(p^*)} g_x'(v) dv, -g_x(w^*))$ if $g(w^*) < g(p^*)$ or $(\frac{-g_x(w^*)}{1 - G_x(w^*)}, -g_x(w^*)) + \int_{p^*}^{w(p^*)} g_x'(v) dv$ if $g(w^*) > g(p^*)$. Either way it is bounded and bounded away from zero.

Secondly, $\partial f / \partial x = \frac{\partial(\Delta_1 / \Delta_2)}{\partial x}$ is uniformly bounded. To prove it we only need to show that both $\frac{\partial \Delta_1}{\partial x}$ and $\frac{\partial \Delta_2}{\partial x}$ are uniformly bounded. For the former, notice that it is given by:

$$\frac{\partial \Delta_1}{\partial x} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{\partial[(1-G(x))^k G^{n-k}(x)]}{\partial x}[1 - G_x(p^*)^k] + \sum_{k=0}^{n}(1-G(x))^k G^{n-k}(x) \frac{\partial[1 - G_x(p^*)^k]}{\partial x}$$

The first component is given by

$$\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}(g(x)(n-k)(1-G(x))^k G^{n-k-1}(x) + g(x)k(1-G(x))^{k-1}G^{n-k}(x))[1 - G_x(p^*)^k]$$

which is uniformly bounded given both series converges. The second component is given by

$$\sum_{k=0}^{n}(1-G(x))^k G^{n-k}(x)[-g(x)\frac{1-G(p^*)}{(1-G(x))^2}]$$

which is also uniformly bounded for all $x < \bar{v}$. Similarly, we can also expand $\frac{\partial \Delta_2}{\partial x}$ and show it is uniformly bounded.

Finally we consider $\partial f / \partial p^*$. Clearly, it is given by $1 + \frac{\partial(\Delta_1/\Delta_2)}{\partial p^*}$. Yet both $\frac{\partial \Delta_1}{\partial p^*}$ and $\frac{\partial \Delta_2}{\partial p^*}$ are uniformly bounded. $\frac{\partial \Delta_1}{\partial p^*}$ is bounded by the mean of $-g(p^*)kG_x(p^*)^{k-1}$, which is also bounded given that its series is convergent (for all $p^* < \bar{v}$ which always holds). $\frac{\partial \Delta_2}{\partial p^*}$ is bounded by the mean of $G_x(w^*)^{k-1}g'_x(w^*) - G_x(p^*)^{k-1}g'_x(p^*)$ which also has a convergent series. Moreover, we have $\Delta_2 < 0$ and $\frac{\partial \Delta_1}{\partial p^*} < 0$, as well as $\Delta_1 > 0$ and $\frac{\partial \Delta_2}{\partial p^*} > 0$. Therefore, $\partial f / \partial p^*$ is uniformly bounded and bounded away from zero.

Putting the parts together, we've shown that $\frac{\partial p^*}{\partial x}$ is uniformly bounded. Therefore $\partial p^*(n, x)$ converges to $p^*(\infty, x)$ uniformly. Therefore there exists N such that $\forall n > N$, $\partial p^*(n, x)$ is decreasing in x.

□

A.4 Proof to Corollary 3

Proof. Use implicit function theorem. Let $f(p^*, n, c) = -p^* + c + \frac{1 - G(p^*)^n}{n[G(p^*)]^{n-1}g(p^*)}$. We have that p^* is an implicit function of n and c defined by $f(p^*, n, c) = 0$. Therefore $\frac{\partial p^*}{\partial n} =$

$-\dfrac{\partial f \,/\, \partial n}{\partial f \,/\, \partial p^*}$. We have

$$\begin{aligned}
\frac{\partial f}{\partial p^*} &= -1 - \frac{n[G(p^*)]^{n-1}g(p^*)}{n[G(p^*)]^{n-1}g(p^*)} - \frac{(1-G(p^*)^n)(n[G(p^*)]^{n-1}g'(p^*) + n(n-1)[G(p^*)]^{n-2}g^2(p^*))}{(n[G(p^*)]^{n-1}g(p^*))^2} \\
&= -\frac{n[G(p^*)]^{n-2}(1-G(p^*)^n)}{(n[G(p^*)]^{n-1}g(p^*))^2}\left[G(p^*)g'(p^*) + (n-1)g^2(p^*)\right] \\
&< -\frac{n[G(p^*)]^{n-2}(1-G(p^*)^n)}{(n[G(p^*)]^{n-1}g(p^*))^2}\left[G(p^*)g'(p^*) + g^2(p^*)\right] < 0
\end{aligned}$$

The last inequality is from the log-concavity of G. Now we show that

$$\begin{aligned}
\frac{\partial f}{\partial n} &= \frac{-n[G(p^*)]^{n-1}g(p^*)G(p^*)^n \log G(p^*) - (1-G(p^*)^n)([G(p^*)]^{n-1}g(p^*) + n[G(p^*)]^{n-1}g(p^*)\log G(p^*))}{(n[G(p^*)]^{n-1}g(p^*))^2} \\
&= \frac{-nG(p^*)^n \log G(p^*) - (1-G(p^*)^n)(1+n\log G(p^*))}{n^2[G(p^*)]^{n-1}g(p^*)} \\
&= \frac{-1+G(p^*)^n - n\log G(p^*)}{n^2[G(p^*)]^{n-1}g(p^*)} > 0
\end{aligned}$$

To see how the last inequality works, let $z = \dfrac{1}{G(p^*)}$, clearly $z \in (1, +\infty)$. The numerator can be written as $y = -1+z^{-n}+n\log z$. It equals 0 when $z = 1,\ \forall n$, and $\dfrac{\mathrm{d}y}{\mathrm{d}z} = \dfrac{n}{z}(1-z^{-n}) > 0$. Therefore the numerator is positive, so is the denominator and hence the ratio.

Having proved that $\dfrac{\partial f}{\partial p^*} < 0$ and $\dfrac{\partial f}{\partial n} > 0$, we have $\dfrac{\partial p^*}{\partial n} = -\dfrac{\partial f \,/\, \partial n}{\partial f \,/\, \partial p^*} > 0$. Also note that $\dfrac{\partial f}{\partial c} = 1$, hence we have $\dfrac{\partial p^*}{\partial c} > 0$ which is quite expected. □

A.5 Proof to Lemma 3

Proof. Firstly we prove the uniqueness of z^*. Take the F.O.C., we have that

$$-f(w-p)pz + F(w-p)z + F(w-p)p = 0$$

Rearrange the terms, we have $\dfrac{pz}{p+z} = \dfrac{F(w-p)}{f(w-p)}$. We want to show that its LHS (denoted by $\Phi_1 = \dfrac{pz}{p+z}$) is increasing in z whereas the RHS (denoted by $\Phi_2 = \dfrac{F(w-p)}{f(w-p)}$) is decreasing in z. Notice we have $\dfrac{\partial \Phi_1}{\partial z} = \dfrac{1}{(p+z)^2}\left(p^2 + \dfrac{\partial p}{\partial z}(p+z)z\right) > 0$ given $\dfrac{\partial p}{\partial z} > 0$. And we have Φ_2 decreasing in p by the log concavity of $f()$, therefore it is also decreasing in z given $\dfrac{\partial p}{\partial z} > 0$. Therefore the F.O.C. has a unique solution. We can also rule out corner solution by observing $\pi(x,0) = \pi(x,+\infty) = 0$, whereas any z s.t. $0 < z$ and $p^*(z) < w$ will lead to strictly positive profit.

Next we show that z^* is increasing in x when $x \leq x^*$. Notice that $w - p^*$ is increasing in x when $x \leq x^*$: w is increasing in x whereas p^* is decreasing in x, as we've shown in Proposition 2. Therefore, Φ_2 is increasing in x and decreasing in z. Also notice that Φ_1 is decreasing in x ($\frac{\Phi_1}{\partial x} = \frac{\partial \Phi_1}{\partial p}\frac{\partial p}{\partial x} < 0$) and increasing in z as we've just shown. Combining two pieces together, we have z^* to be increasing in x by implicit function theorem. $\square$

A.6 Figures

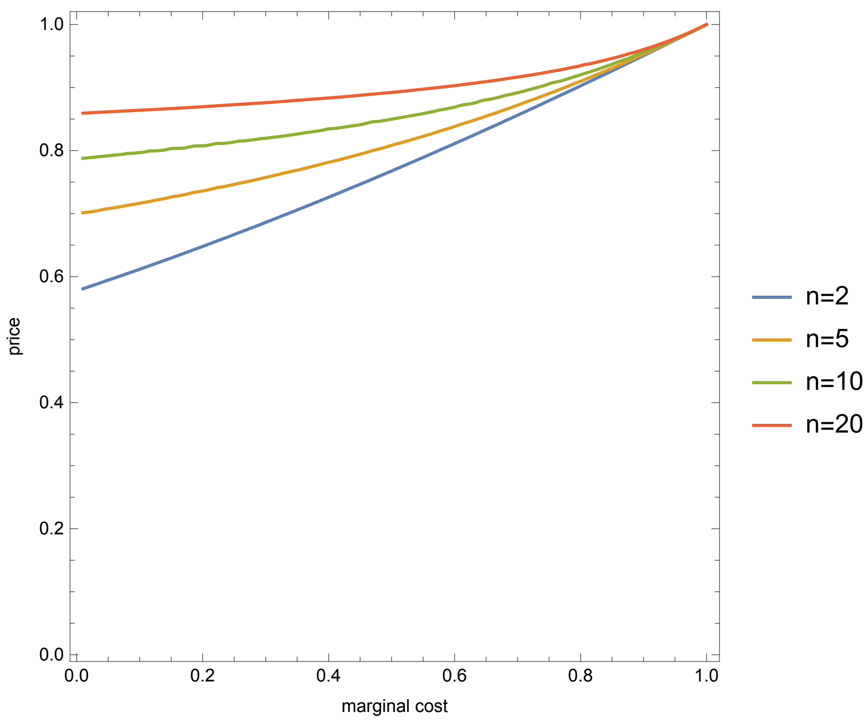

Figure 7: Relevance ranking prices versus n and c with uniform v

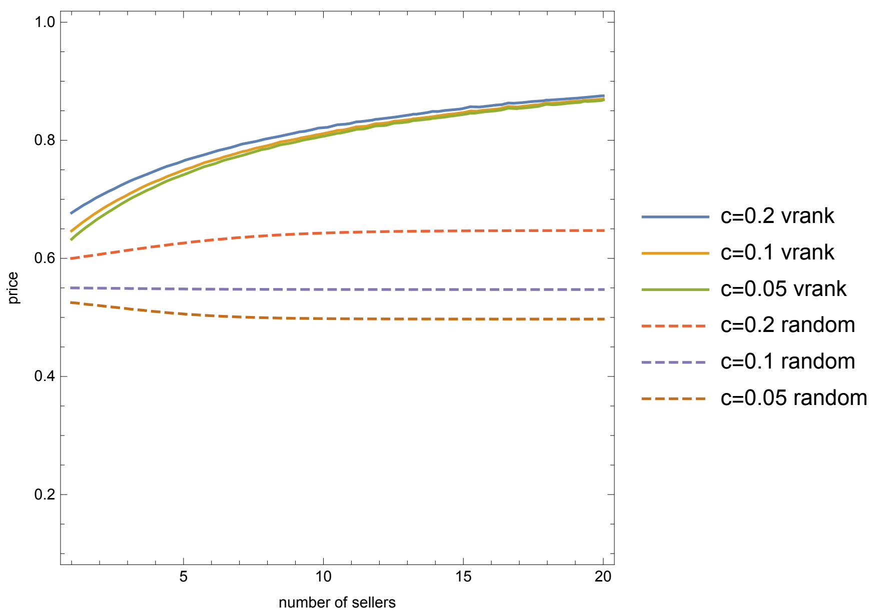

Figure 8: Prices versus c and n with uniform v, for both relevance ranking and random ranking $(s = 0.1)$