Advertising on Online Marketplaces: Information Asymmetry and the Relevance of Sponsored Listings*

Vibhanshu Abhishek
vibs@uci.edu
UCI Merage School of Business

Kinshuk Jerath
jerath@columbia.edu
Columbia Business School

Siddhartha Sharma
ssharma3@andrew.cmu.edu
Heinz College, Carnegie Mellon University

December 2019

*The authors thank Patrick Bajari, Vicente Lagos, Dokyun Lee, Abhishek Nagaraj, Harikesh Nair, Martin Peitz, Navdeep Sahni, Andrey Simonov, Param Vir Singh, Shubhranshu Singh, Rahul Telang, Gerard Tellis, Shruthi Venkatesh, Sha Yang and Kaifu Zhang for valuable comments. This paper has also benefited from the comments of seminar participants at Carnegie Mellon University, University of Southern California and Cornerstone Research, and attendees of 2019 Bass FORMS Conference, 2018 Workshop on Information Systems and Economics, 2018 Conference on Information Systems and Technologies, 2018 Summer Institute of Competitive Strategy, 2018 Marketing Science Conference, 2018 and 2019 ZEW Conference on Economics of ICT and 2018 Conference on Digital Economics. The authors also thank Chazen Institute at Columbia Business School and Amazon.com, Inc. for research support. We are also grateful to many Flipkart employees who made the work possible. All remaining errors are our own.

Advertising on Online Marketplaces: Information Asymmetry and the Relevance of Sponsored Listings

Abstract

Advertising on e-commerce marketplaces, wherein sponsored product listings are interleaved with organic product listings, is a large and growing phenomenon. In this paper, we both theoretically and empirically study whether including sponsored listings improves or hurts the overall quality and relevance of the set of products displayed to consumers. Our theoretical analysis reveals that the relevance of the ads to the consumers' search queries depends on the level of information asymmetry between the marketplace owner and the sellers who sell products with different degrees of relevance. Specifically, when information asymmetry is low (high), i.e., the platform can (cannot) easily distinguish between high- and low-relevance sellers, then low-relevance (high-relevance) sellers have a greater incentive to advertise. However, even when low-relevance products are displayed as ads, consumers end up finding well-matching products as long as search and evaluation costs are reasonable; therefore, the overall impact on sales is relatively small while the marketplace benefits from the additional revenue from selling ads. We obtain data from a large-scale field experiment run at Flipkart, a leading online marketplace in India, and find that various empirical patterns implied by our theoretical results hold in the data. Our study provides several practical implications for managers of e-commerce marketplaces.

Keywords: E-commerce Platforms; Sponsored advertising; Asymmetric information; Product heterogeneity

1 Introduction

E-commerce marketplaces, where third-party sellers and consumers transact online, have been growing in importance in the retail industry worldwide. Prominent examples of e-commerce marketplaces include Amazon.com (which accounts for 49% of all online retail sales in USA and 30% in India in 2018), Alibaba (which accounts for 58% of all online retail sales in China in 2018) and Flipkart (which accounts for 28% of all online retail sales in India in 2018).[1] Among these, Alibaba and Flipkart sell exclusively through third-party sellers while Amazon sells both directly and through third-party sellers, and in 2018 third-party sellers accounted for 58% of physical gross merchandise sales on Amazon, amounting to $160 billion (Keyes, 2019). These leading marketplaces have hundreds of thousands of products available on them in most product categories.

Online marketplaces typically charge a commission to the third-party sellers for the transactions taking place on the marketplace (Hagiu and Wright, 2014; Abhishek et al., 2015), and therefore aim to increase purchase rates. Given that a typical user views only a limited set of products on the marketplace due to the search costs involved, a marketplace faces the daunting task of matching the most relevant product listings to a search query (such as "winter jacket" or "digital camera")

[1] See Lunden (2018), Shelley (2019) and Blystone (2019).

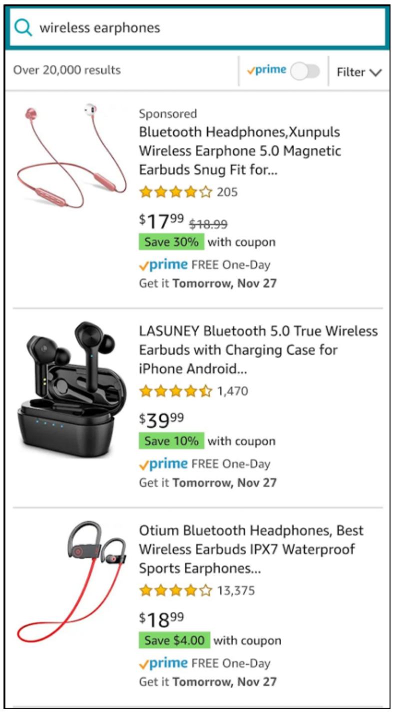
in order to maximize the probability of a positive consumer outcome, such as a click through or a purchase. Most online marketplaces rely on product features, historical data on performance and machine learning methods to perform this matching, returning a list of products sorted by predicted relevance for a consumer query or as a consumer browses the website. These listings are referred to as *organic* listings. However, several new and existing sellers introduce new products every day in the marketplace for which the sellers may have some information that the platform does not have. Furthermore, even for previously existing products, there are short-term trends important for predicting relevance and known to third-party sellers but whose signature is not present yet in the data.[2] Note that we are not assuming that the sellers have *more* information than the platform, rather we are only assuming that the sellers have *some* information about their products that is different from that of the platform, and is relevant to ranking the products.

These reasons create the problem of information asymmetry between the platform and the sellers that reduces the accuracy of the relevance algorithm, and it becomes extremely difficult for the platform to correctly evaluate every listing to match them to a user's search query. Due to information asymmetry, the platform may fail to identify high relevance listings especially when each listing's relevance is subjective and cannot be easily quantified. This problem is more salient for product categories in which the idiosyncratic match component, which is difficult to encode digitally, has high importance. For instance, the asymmetric information problem would be larger for a category such as clothing (under which the query "winter jacket" falls) compared to a category such as electronics (under which the query "digital camera" falls), as it is possible to describe products online more comprehensively in the latter category.

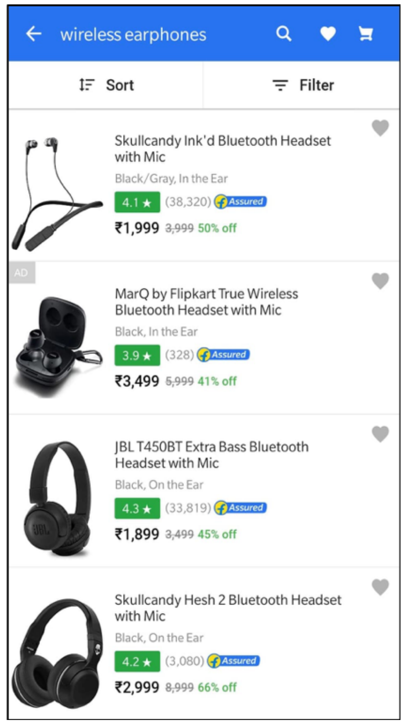
Most e-commerce marketplaces also allow sellers to promote their listings in search results by artificially increasing their ranking in exchange for an additional fee (Borison, 2015; Shields 2017). These promoted listings are referred to as *sponsored* listings. Sponsored product listings appear in the ladder of organic product listings, typically with a "Sponsored" identifier on Amazon (see Figure 1a) or "AD" identifier on Flipkart (see Figure 1b). The proportion of sponsored slots in the search results is about 20%. The motive behind enabling these ads is to give sellers and products more visibility, while creating an additional source of revenue in addition to the sales commission.[3]

[2]E.g., a particular kind of jacket that was recently worn by a celebrity is popular with consumers; sellers of jackets have identified this trend but it is too recent to be identified in data that the platform has available to process.

[3]Source: `https://services.amazon.com/advertising/overview.html`.



(a) Amazon



(b) Flipkart

Figure 1: Sponsored Listings on Mobile Apps

It is estimated that Amazon earned over USD 10 billion from ads in 2018, and this number is expected to grow significantly in the future (Soper, 2019), while Alibaba's marketplace is known to charge a very small sales commission rate and make most of its money through ad sales. Our discussions with managers at several e-commerce marketplaces reveals that advertising is a very lucrative business (Economist, 2018). Although Amazon's advertising revenue is a tiny fraction of its total sales, the profits from advertising could constitute a significant portion of its profits.[4]

While sponsored listings create a new source of revenue for marketplaces, they could be a double-edged sword. Sponsored listings allow third-party sellers to artificially inflate their positions on the product list. This inflation might dilute the quality of the search results, as organic listings in some of the top positions might now be substituted with less relevant or lower quality listings. This, in turn, can adversely affect the user experience and the probability of purchase on these platforms.

The effect of displaying a sponsored listing instead of an organic listing in a search result slot can be seen to be determined by two effects. First, it is determined by the effect that the

[4]Source: `https://www.macrotrends.net/stocks/charts/AMZN/amazon/net-income`.

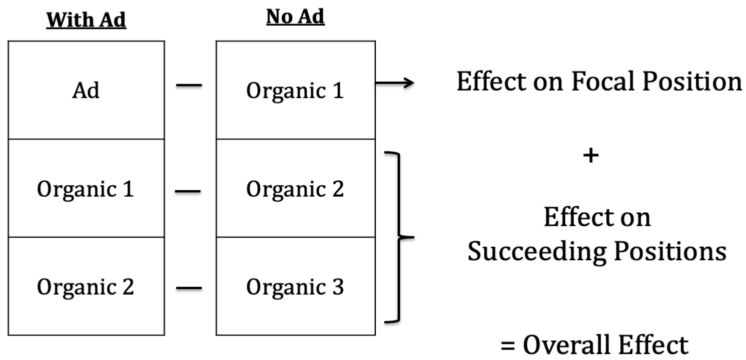


Figure 2: Effects of a Sponsored Listing

sponsored listing has on the performance (in terms of metrics such as click-through rates and purchase conversion) of the slot in which it is placed, i.e., the difference in the performances of the sponsored listing and the organic listing that would have appeared at the same position had the ad been absent. Second, it is also determined by spillover effects, which can be thought of as the difference in the performances of the organic listings that are in the neighborhood of the sponsored listing, due to the presence of the sponsored listing. The aggregation of these two effects determines the extent of the overall effect of displaying a sponsored listing. In Figure 2, we illustrate this with a stylized scenario in which there is an ad in the first slot of the ladder of results (whereas there would be an organic listing in this slot if ads were not shown).

Given this setting of advertising on online marketplaces, we ask the following research questions in this paper. First, how relevant are sponsored listings to a consumer search query compared to the organic listings that they displace in terms of click and conversion performance? Second, what is the effect of sponsored listings on the organic listings at neighboring positions? Third, what is the effect of including sponsored listings on the overall search performance? Finally, what do the answers to these questions imply for the strategies that marketplaces may follow for displaying sponsored ads as part of their search results?

We structure our investigation as follows. To gain insights into the phenomena that we want to study, we first develop a simplified theoretical model of advertising on a marketplace. In response to a consumer search query, multiple heterogeneous sellers offer products that are of high relevance or of low relevance to the query. The consumer is presented with a ranked list, with ads in specific positions, which she searches sequentially. The sellers have private information regarding the relevance of the products to the query, i.e., there is information asymmetry between the marketplace and the sellers, which implies that the marketplace cannot perfectly distinguish high- and low-relevance sellers, while the sellers can bid to be placed in the ad slots. Our main finding from this model is that there is an equilibrium in which low-relevance sellers have a higher incentive to advertise than high-relevance sellers if information asymmetry is low enough (i.e., when the marketplace finds it relatively easy to distinguish between sellers of different relevance), whereas high-relevance sellers have a higher incentive to advertise than low-relevance sellers if information asymmetry is high enough (i.e., when the marketplace finds it relatively difficult to distinguish between sellers of different relevance). Therefore, for high information asymmetry, advertising serves as a screening mechanism for the marketplace with high-relevance sellers promoting their listings, thus improving the quality of the search results. On the other hand, for low information asymmetry, advertising leads to the promotion of low-relevance sellers, thus reducing the quality of the search results.

Our theoretical model produces some key empirical hypotheses. First, for high (low) information asymmetry, a sponsored listing will receive higher (lower) clicks and conversions than the organic listing that it displaces. Second, due to the sequential search process, for high (low) information asymmetry, the listings following a sponsored listing will receive lower (higher) clicks and conversions than the listings following the organic listing that the sponsored listing displaced. Third, due to these two countervailing effects, the overall clicks and conversions after a consumer search may not be very different with and without sponsored ads for both high and low information asymmetry cases. Nevertheless, as the ads are of high-relevance sellers in a high information asymmetry category, the overall quality of results is expected to be better and therefore in a high information asymmetry category one can expect slightly higher search-level clicks and conversions.

Next, we conduct an empirical investigation to test our hypotheses. We use data from a large-scale randomized field experiment on Flipkart, which is a leading e-commerce marketplace in India. In this experiment, approximately 1.2 million users were randomly divided into different groups that were served varying numbers of ads at different positions, and their click and conversion responses were recorded. The exogenous variation in the positions and the number of the ads across the different groups allows us to identify the effects of sponsored listings. In our analysis, we consider listings from two of the most popular product categories — electronics and clothing. Products in the electronics category are easier to characterize in an online setting compared to products in the clothing category because they have a larger share of digital attributes (Lal and Sarvary, 1999). Therefore, we expect that information asymmetry between the third-party sellers and the platform is lower in the electronics category than in the clothing category. Our conversations with managers at Flipkart confirms this expectation — managers state that the Flipkart system is better able to characterize products in the electronics category than in the clothing category and, in response to a search query, is better able to sort and rank them in the electronics category than in the clothing category.

Our empirical analysis provides strong support for our theoretical predictions. We find that in searches related to the electronics category (which is a low information asymmetry category), compared to the displaced organic listings, both the mean click (click per impression or click-through rate) and the mean conversion (conversion per impression or conversion rate) of sponsored listings are statistically significantly lower (with effect sizes between 5% and 37%, which are large effect sizes). Furthermore, the presence of an ad in a slot leads to statistically significantly higher mean click and conversion for organic listings that succeed the slot (with effect sizes between 8% and 38%, which, again, are large). Both of these patterns are as theoretically expected. In contrast, in the clothing category (which is a high information asymmetry category), compared to the displaced organic listings, both the mean click and the mean conversion of sponsored listings are statistically significantly higher (with magnitudes between 11% and 25%, which, again, are large). For links that succeed a slot, the presence of an ad in the slot leads to no statistically significant difference in outcomes. The first pattern (in clicks) is as theoretically expected; the second finding is inconclusive which may be because the conversion rate in clothing is very small which leads to the analysis on conversions in the clothing category having low power. At the search level, we find either a very small effect or no effect on clicks or conversions in either category. In other words, even though the effects of showing sponsored ads at the focal positions and succeeding positions are quite large in magnitude, countervailing effects essentially cancel out to have minimal or no effect on the search level outcomes, as expected. Overall, while the field experiment does not provide conclusive evidence of our theoretical predictions regarding the impact of information asymmetry (for that, we would have to experimentally manipulate information asymmetry across different scenarios), the results that we obtain for the electronics and clothing categories (low and high information asymmetry categories, respectively) are essentially perfectly in line with our theoretical predictions. The empirical analysis at the sub-category level also yields identical results whenever we have sufficient power.

At a time when advertising on marketplaces is getting increasingly popular, our findings provide a number of implications for practitioners. For managers at marketplaces, we identify that sponsored search can serve as a screening device in categories where they face a high degree of information asymmetry and, more generally, settings in which they find it difficult to distinguish products by relevance to consumers' needs; in fact, using sponsored listings may improve the set of results presented to a consumer and may be an effective way for marketplaces to help solve the classic "cold start"[5] problem for products with non-digital attributes. On the other hand, we find that in low information asymmetry categories like electronics, advertising on marketplaces promotes low-relevance products, which can hurt consumer experience; nevertheless, we find that, at least in the short run, there is limited revenue loss from showing ads in the search results as users tend to purchase at a higher rate from listings at succeeding positions. An implication of this is that the revenue from sponsored products either adds to, or does not hurt, the revenue made from transaction commissions. For sellers, we find the implication that for high information asymmetry categories high-relevance sellers should advertise, while for low information asymmetry categories low-relevance sellers should advertise.

Our paper contributes to the nascent literature on the role of advertising on marketplaces. A recent paper on this topic is Choi and Mela (2019), who study the problem of monetizing online marketplaces and, to that end, develop a joint model of consumer and seller behavior on an online marketplace to find the optimal ad ranking and pricing mechanism. A main theme in their research is that the platform trades off fees from advertising with commissions from product sales. We find, both theoretically and empirically, that in high information asymmetry categories showing sponsored ads may actually increase conversions, i.e., there may not always be a trade-off between advertising and sales revenues. However, we note that how ads are priced and allocated to slots is different in the two settings; also, unlike Choi and Mela (2019) we do not look into the question of mechanism design. Long et al. (2019) conduct a theoretical study to determine how a marketplace that wants to maximize total profit from ad fees and sales can use information learned

[5]Source: `https://en.wikipedia.org/wiki/Cold_start_(computing)`.

from sponsored ad auctions to order the joint list of organic and sponsored results and to determine other parameters such as the commission rate.

Sahni and Nair (2019) run advertising field experiments on a food ordering marketplace and find that advertising serves as a signal of quality to consumers. This is in line with a large theoretical and empirical literature on the signaling effect of advertising (Nelson, 1974; Kihlstrom and Riordan, 1984; Caves and Greene, 1996; Ackerberg, 2003; Feng and Xie, 2012). We find that in settings with high information asymmetry between a platform and sellers, advertising works as a screening mechanism for the platform. Our work also adds to the substantial literature on sponsored search advertising (Edelman et al., 2007; Athey and Ellison, 2011; Jerath et al., 2011; Liu and Viswanathan, 2014; Jeziorski and Segal 2015; Narayanan and Kalyanam 2015; Agarwal and Mukhopadhyay, 2016), and to the growing literature on e-commerce marketplaces (Jiang et al., 2011; Hagiu and Wright, 2014; Abhishek et al., 2015).

The rest of the paper proceeds as follows. In Section 2, we present a parsimonious analytical model of marketplace advertising and obtain theoretical insights. In Section 3, we describe our research setting, the experiment design, and the data used for analysis. In Section 4, we present our empirical findings. In Section 5, we conclude with a discussion.

2 Theoretical Analysis

In this section, we present a stylized theoretical model of advertising on an online marketplace to develop the intuition for the empirical analysis presented later in the paper.

2.1 Model

Our effort is to develop a simplified model that can capture the main forces at play and communicate the main insights. The model has three types of players: the marketplace, heterogeneous sellers and consumers. We assume that there are three sellers, one selling a product of high relevance to consumers (denoted by H), and two selling products of lower relevance to the consumers (denoted by L_1 and L_2); this reflects the reality that typically low-relevance sellers are more in number than high-relevance sellers. Let V_H, V_{L_1} and V_{L_2} denote the expected utilities of H, L_1 and L_2, respectively, where $0 < V_{L_1} = V_{L_2} = V_L < V_H < 1$. This deterministic utility measures the expected relevance of the listing with respect to the search query (e.g., the quality of the product and how well it matches the consumers' needs). A consumer's utility from the product $i \in \{H, L_1, L_2\}$ is given by $u_i = V_i + \epsilon_i$, where ϵ_i is an idiosyncratic i.i.d. shock, $\epsilon_i \sim \text{Uniform}[-1, 1]$. We assume that on viewing a listing the consumer learns the type of the seller as H or L, i.e., learns V_i and also learns the realization of ϵ_i. We normalize the consumer's utility from an outside option to zero.

A consumer searches for a product on the marketplace. In response, the marketplace displays the three sellers in an ordered list to the consumer. Clearly, H is most relevant to the consumer. L_1 and L_2 are equally relevant, but less so than H. When there exists information asymmetry between the platform and the sellers, the platform may misidentify the type of the sellers. We assume that, from the three products, the platform can correctly identify H with probability $\kappa \in [\frac{1}{3}, 1]$. Therefore, κ can be interpreted as a parameter that measures the amount of information asymmetry between the platform and the sellers, with smaller κ implying greater information asymmetry.

If there is no advertising, in response to a consumer query, the platform ranks H above L_1 and L_2 with probability $\kappa \in [\frac{1}{3}, 1]$, or it makes an error and ranks H below at least one of the L-type listings with probability $1 - \kappa$. Specifically, let $\sigma(xyz)$ denote the probability of ranking listing x at the first position, y at the second position and z at the third position. Then, $\sigma(HL_iL_j) = \frac{\kappa}{2}$, $\sigma(L_iHL_j) = \frac{1-\kappa}{4}$ and $\sigma(L_iL_jH) = \frac{1-\kappa}{4}$ for $i, j \in \{1, 2\}$ and $i \neq j$. $\kappa = 1$ implies there is no information asymmetry while $\kappa = \frac{1}{3}$ implies maximum information asymmetry.

If there is advertising, the marketplace auctions the first position as a sponsored slot to the sellers through a second-price pay-per-click mechanism. The losing sellers are ranked in the organic slots (the second and the third positions) in order of their predicted relevance (which, as described, is imperfect in the presence of information asymmetry). Here, we make two assumptions based on the practices of our data partner Flipkart. First, we assume that one product is placed in the list only once, i.e., if a product is placed in the sponsored slot (i.e., the first slot), then it can not also be placed in an organic slot (i.e., the second and third slots). Second, we assume that when the platform auctions the sponsored slot, information from the bids is not used while deciding the organic rankings.[6] To the extent that we want to use our model to obtain hypotheses for empirical

[6] At Flipkart, the customer experience team is separate from the ad products team (this is for various reasons, such as avoiding conflicts of interest). Our conversations with managers at multiple e-commerce marketplaces indicate that this is the case at most marketplaces. Of late, there is an effort at marketplace firms to jointly rank organic and sponsored results. Recent work, such as Long et al. (2019) studies this problem.

testing in our data, it is appropriate to maintain these assumptions. We note that in our model the various forces at play are captured, it is only the optimization of the ranked list that is restricted.

Specifically, we assume that if L_i gets placed in the ad slot, then H is placed higher than L_j, i.e., in the second slot (which is an organic slot), with the same probability that H would not be placed below *both* Ls in the case of no ads, which is given by

$$\sigma(HL_1L_2) + \sigma(HL_2L_1) + \sigma(L_iHL_j) = \kappa + \frac{(1-\kappa)}{4}, i, j \in \{1,2\}, i \neq j.$$

If H gets placed in the ad slot, then each L is placed in the second slot with probability 1/2. For the auction, we assume that the reservation price is set equal to zero. We also assume that the profit margins of all sellers are equal to one.

We characterize the consumer's search behavior as follows. The consumer does not incur any cost to view the listing at the first position and thus evaluates it with probability one. However, she incurs an evaluation cost, $s > 0$, to view and evaluate the listing at the second position. As there is one high-relevance listing and two low-relevance listings, we assume that if the first listing is H, the consumer's expected utility from clicking on the next listing is given by V_L. If the first listing is L_1 or L_2, her expected utility from clicking on the next listing is given by $\hat{V}$, where $\hat{V} = \frac{V_H+V_L}{2}$ if there is no sponsored slot, and $\hat{V} = (\kappa + \frac{1-\kappa}{4})V_H + \frac{3(1-\kappa)}{4}V_L$ if there is a sponsored slot. This is because, conditional on an L being ranked first, if there is no sponsored slot, H is ranked at the second and third positions with equal probability,[7] while if there is a sponsored slot, H is ranked at the second position with probability $(\kappa + \frac{1-\kappa}{4})$. The consumer views the second listing if her expected marginal benefit of evaluating the second listing is greater than the marginal search cost. We assume that the consumer's marginal search cost of evaluating the listing at the third position is so high that she never views the third listing.[8]

[7] $\frac{\sigma(L_iHL_j)}{\sigma(L_iHL_j)+\sigma(L_iL_jH)} = \frac{\sigma(L_iL_jH)}{\sigma(L_iHL_j)+\sigma(L_iL_jH)} = \frac{1-\kappa}{4} / \frac{1-\kappa}{2} = \frac{1}{2}$ for $i, j \in \{1,2\}$ and $i \neq j$.

[8] We note that our focus is on the optimal advertising decision of the sellers given the consumer behavior, and we are making certain simplifying assumptions about consumer behavior and beliefs — specifically, that consumers start searching from the top position downwards, and if they see a product of type L in the first position then they assume the second position has a product of type H and L with equal probability. Our assumptions can be interpreted as assuming that consumers are boundedly rational, which has been assumed in various previous studies on position auctions such as in Jerath et al. (2011), and has been assumed and argued more generally such as in Gabaix and Laibson (2006) and Ho et al. (2006), among many others. Furthermore, while we do not do the same here (and do not consider it necessary), we note that Jerath et al. (2011) show in Section 5.2 in their paper that boundedly rational behavior by consumers can be fully rationalized in equilibrium by introducing beliefs of consumers on other constructs, such as firm margins.

Electronic copy available at: https://ssrn.com/abstract=3013468

After evaluating the listings, the consumer makes her click decision. Specifically, if she evaluated only the first listing then she decides whether to click it or not, and if she evaluated the first two listings then she decides which one to click or not to click anything. We assume that γ_i is the probability of conversion conditional on click of a listing of type $i \in \{H, L\}$, and that $0 < \gamma_L < \gamma_H < 1$; for simplicity, we assume that for any listing the other listings do not affect the conversion probability post click.

Before we proceed to the analysis, we note that we have made a number of assumptions in the model either for simplicity or to harmonize our model with the practices at Flipkart, our data partner. However, the results and insights that we obtain could be obtained under a number of alternative assumptions corresponding to different variants of the model.

2.2 Analysis and Results

We now solve the model. The proof for all the lemmas and propositions are provided in Appendix A. Let δ^i denote the probability of evaluating the second listing if the type of the first listing is $i \in \{H, L\}$. Then, $\delta^H = \Pr(V_L - u_H > s)$ and $\delta^L = \Pr(\hat{V} - u_L > s)$, where $\hat{V}$ could be different with and without the sponsored listing. Note that $\delta^L > \delta^H$ and $\frac{d\delta^i}{ds} < 0$ for $i \in \{H, L\}$. Let $\theta_i = \Pr(u_i > 0)$ denote the probability of click on listing $i \in \{H, L\}$ when only i, the first listing, is evaluated, and $\theta_{ij} = \Pr(u_i > \max\{u_j, 0\})$ denote the probability of click on listing i when both listings $i, j \in \{H, L\}$ are evaluated, where $\theta_H > \theta_L$ and $\theta_{HL} > \theta_{LL} > \theta_{LH}$. Let β_k^{jl} and π_k^{jl} denote the probability of click and the probability of conversion, respectively, of the listing at position $k \in \{1, 2\}$ when $j \in \{H, L\}$ occupies the first and $l \in \{H, L\}$ occupies the second position. The expressions for all δ, θ, β and π are derived in Appendix A. We assume $0 < \gamma_L < \gamma_H < 1$ such that $\pi_1^{HL} > \pi_1^{LH}$ and $\pi_2^{HL} < \pi_2^{LH}$.

The value of advertising per click for the two types of listings is as follows:

$$v_H = \gamma_H \left\{ 1 - \frac{(1+3\kappa)}{4} \frac{\beta_2^{LH}}{\beta_1^{HL}} \right\}, \tag{1}$$

$$v_L = \gamma_L \left\{ 1 - \frac{(1+3\kappa)}{4} \frac{\beta_2^{HL}}{\beta_1^{LH}} + \frac{3(1-\kappa)}{4} \frac{(\beta_2^{LL} - \beta_2^{HL})}{\beta_1^{LL}} \right\}. \tag{2}$$

Since bidding one's own valuation is a weakly dominant strategy for each bidder in a second-price auction, the seller of listing type $i \in \{H, L\}$ bids v_i. The following lemma states how κ affects the incentives of the two types of listings.

Lemma 1 $\frac{dv_i}{d\kappa} < 0 \ \forall \ \kappa \in [\frac{1}{3}, 1]$, *where* $i \in \{H, L\}$

Lemma 1 states that both types have a higher valuation for advertising when information asymmetry is higher, i.e., as κ decreases. This happens for the high type because as information asymmetry increases, the probability of it being ranked at the third position increases, where it gets no views. This creates a higher incentive for the high type to advertise when κ is lower. As for the low types, the incentive increases because, with lower κ, the two low types can occupy the first two positions more frequently, which increases the returns to winning the first position (which is the sponsored slot). Since, both v_H and v_L are increasing in κ, the marketplace earns a higher price per click for the sponsored slot when there is more asymmetric information (as cost per click is $\min\{v_H, v_L\}$).

We now compare the valuations of the high- and low-relevance sellers to determine which one wins the sponsored slot. We present the key result in the following proposition. We define the following quantities that are used in the proposition:

$$\underline{r} = \left. \frac{1 - \frac{\beta_2^{HL}}{2\beta_1^{LH}} + \frac{\beta_2^{LL} - \beta_2^{HL}}{2\beta_1^{LL}}}{1 - \frac{\beta_2^{LH}}{2\beta_1^{HL}}} \right|_{\kappa = \frac{1}{3}}, \overline{r} = \left. \frac{1 - \frac{\beta_2^{HL}}{\beta_1^{LH}}}{1 - \frac{\beta_2^{LH}}{\beta_1^{HL}}} \right|_{\kappa = 1}$$

$$\text{and } \overline{\kappa} = 1 - \frac{4\beta_1^{LL}(\gamma_H(\beta_1^{HL} - \beta_2^{LH})\beta_1^{LH} - \gamma_L\beta_1^{HL}(\beta_1^{LH} - \beta_2^{HL}))}{3(\gamma_L\beta_1^{HL}(\beta_1^{LL}\beta_2^{HL} - \beta_1^{LH}(\beta_2^{HL} - \beta_2^{LL})) - \gamma_H\beta_2^{LH}\beta_1^{LH}\beta_1^{LL})}.$$

Note that $\underline{r} \leq \overline{r}$.

Proposition 1 *Assume that* $\underline{r} < \frac{\gamma_H}{\gamma_L} < \overline{r}$. *Then for* $\kappa \in \left[\frac{1}{3}, \overline{\kappa}\right]$ *(high information asymmetry), the high-relevance listing is displayed in the sponsored slot (i.e.,* $v_H > v_L$*), whereas for* $\kappa \in (\overline{\kappa}, 1]$ *(low information asymmetry) a low-relevance listing is displayed in the sponsored slot (i.e.,* $v_L > v_H$*).*

For completeness, we also state the following: if $\frac{\gamma_H}{\gamma_L} \leq \underline{r} < \overline{r}$, a low-relevance listing is always displayed in the sponsored slot (i.e., $v_L > v_H \ \forall \ \kappa$), and if $\underline{r} < \overline{r} \leq \frac{\gamma_H}{\gamma_L}$, the high-relevance listing is always displayed in the sponsored slot ($v_H > v_L \ \forall \ \kappa$).

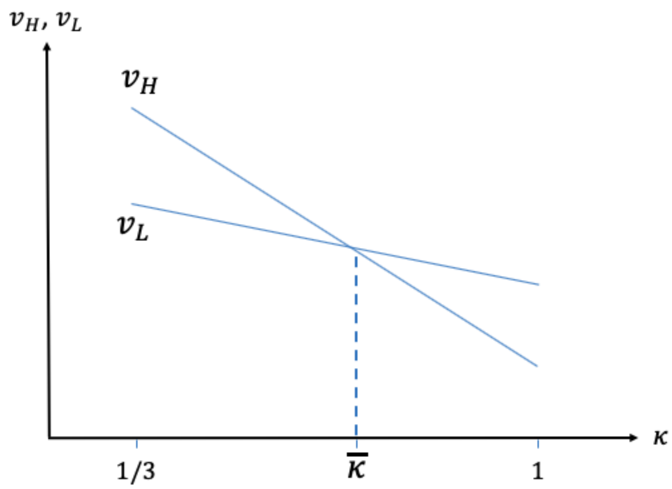


Figure 3: Incremental value of being placed in the sponsored slot for high- and low-relevance sellers with respect to κ

Now consider the statement of the proposition. Suppose $\frac{\gamma_H}{\gamma_L} < \overline{r}$, i.e., when γ_H is not too large as compared to γ_L. Then, for large enough κ, i.e., for low enough information asymmetry, the Ls value the sponsored slot higher than H (i.e., $v_L > v_H$). This is because if a low-relevance seller loses the ad auction, it will be relegated to the third position with a high probability where it will get no conversions, whereas the high-relevance seller knows that even if it loses the auction it will be placed in the second slot with high probability where it will get sufficient conversions; in other words, under low information asymmetry, a low-relevance seller is hurt more on not being placed in the sponsored slot and therefore values it more than the high-relevance seller.

From Lemma 1 we know that, as κ decreases (i.e., information asymmetry increases), sellers of both types of listings value the sponsored slot more. When $\frac{\gamma_H}{\gamma_L} > \underline{r}$, i.e., when γ_H is relatively large compared to γ_L, then v_H increases at a faster rate than v_L as κ decreases. Under this condition, the high type can have a higher incentive to advertise than the low type when information asymmetry is large enough. Indeed, when information asymmetry is large, if the high-relevance seller wins the auction it will obtain a large number of clicks and conversions while if it loses the auction it may be placed in the third slot with a reasonable probability, where it will get no clicks or conversions. On the other hand, if a low-relevance seller wins the auction it will be placed at the top but still not get many clicks and conversions (due to its low relevance) while even if it loses the auction it has a reasonable probability of being placed in the second slot. Therefore, when information asymmetry is large, the high-relevance seller has more to lose from losing the auction, and therefore bids higher than the low-relevance seller. Figure 3 depicts such a scenario.

Based on the above proposition, we derive the following corollaries that compare outcomes when the marketplace allows sponsored listings versus when it does not. The aim of conducting this comparison is to obtain results that enable us to determine the empirical patterns that we expect to see in our data in the empirical analysis. We will assume for these corollaries that the conditions of Proposition 1 hold, such that the high-relevance (low-relevance) seller is displayed in the sponsored slot when information asymmetry is high (low) enough.

Corollary 1 ***Clicks and conversions on the first slot.*** *Under high (low) information asymmetry, if the first slot displays a sponsored listing, it gets more (fewer) clicks and conversions than if it displayed an organic listing.*

Corollary 2 ***Clicks and conversions on the second slot.*** *Under high (low) information asymmetry, if the first slot displays a sponsored listing, the second slot gets fewer (more) clicks and conversions than if the first slot displayed an organic listing.*

From Proposition 1, we know that if information asymmetry is high, then the H-type seller obtains the first slot (which is the sponsored slot). If this slot were not for a sponsored ad, an L-type seller would have high likelihood of being displayed in this slot. Therefore, if the first slot displays a sponsored ad listing, this listing will be of the H-type, which implies that it will get more expected clicks and conversions than if there were no sponsored listings. This also implies that fewer consumers proceed to the second slot, which will now display an L-type seller. With low information asymmetry, the reasoning is exactly the reverse — an L-type seller obtains the first slot when this is a sponsored slot, and is less likely to obtain it if there are no sponsored slots, leading to the predictions on clicks and conversions in the two positions in the corollaries.

Next, we state a corollary regarding the total click and conversion performance with and without a sponsored listing at the top slot.

Corollary 3 ***Total effect on clicks and conversions.*** *Under high (low) information asymmetry, if the first slot displays a sponsored listing, the overall click and conversion performance at two* *positions together increases (decreases). In both cases, the magnitude of the effect becomes smaller as the consumer search cost, s, decreases.*

When there is high information asymmetry, the high-relevance listing occupies the sponsored slot, which leads to more clicks and conversions. However, this also lowers the clicks and conversions from the succeeding position, thus nullifying some of the positive effect at the sponsored slot. Conversely, when there is low information asymmetry, the sponsored listing can hurt the search performance due to the selection of the low-relevance listing at the top position. However, if the consumers have low marginal search costs, then they continue with their search and purchase at a higher rate (compared to the case with no ads) from the next position. Note that Corollaries 1 and 2 show countervailing effects on the first and second slots, which implies that they might nullify each other in the overall effect of sponsored listings on clicks and conversions from a search, and so the overall effect can be small based on the search cost.

3 Empirical Setting

For our empirical analysis, we collaborate with *Flipkart*, a leading e-commerce marketplace operating in India (Shelley, 2019). Flipkart users can access the platform via both a website and a mobile application (hereinafter app). We note that mobile commerce accounted for more than 65% of all e-commerce activity in 2016 in India.[9] We conduct our study using data from Flipkart's mobile app.

At the top of the app's homepage, a user sees a search tab to type in her search query. Just below the search tab, Flipkart also lists different product categories like *Electronics*, *Fashion*, *Home*, etc. to aid consumer search. In this paper, we focus on sessions where a user voluntarily searches for a product by typing in a query. Given a query by a user, a ranked list of listings (i.e., products) gets generated. Around 20% of these listings are sponsored and appear at fixed positions, and the rest are organic. Users can browse through these listings by scrolling through them.

We define a *search* as the list of listings that get displayed in response to a query by a user, and an *impression* as the displaying of a listing by the platform after a search, i.e., a search consists of multiple impressions. A new search is generated either when a user searches for a new product

[9]Source: `https://www.statista.com/statistics/244025/india-m-commerce-share`.

or when she adds a certain filter to the current search. The maximum position viewed by the user during a search determines its *depth*. Typical features or information available on the search results page regarding a listing are the name of the product, an image, price, discount, average rating, number of ratings, and an "AD" identifier if the listing is sponsored. Based on this information, a user can *click* (tap) on the listing to enter the specific product's (listing's) page to learn more details, such as more images, seller details, reviews, etc. On this page, a user can either *convert* by tapping on "Buy Now," "Add to Cart" or "Add to Wishlist," or go back to the search results page.

Flipkart charges sellers a fixed price per click on sponsored listings. Given the price, sellers can choose the products (listings) that they want to advertise, total budget, and the campaign duration. We note that this price is set at the product vertical level, e.g., mobile phone, headphones, television, etc., in *Electronics*, and shirt, top, shoes, etc., in *Fashion* (and not at the product level or the query level). Note that if a seller indicates that it wants to advertise a listing, i.e., if a seller indicates that a listing should be considered for a sponsored slot, then that listing is not considered by Flipkart to be placed in the organic slots. When a user enters a search query, using its internal algorithms, Flipkart picks listings from the pool of possible sponsored listings and ranks them in the sponsored slots in decreasing order of predicted relevance. Note that irrespective of positions in which the sponsored listings show up, the advertisers are charged the same cost per click. Similarly, Flipkart picks the most relevant listings from the pool of organic listings (which, by default, are all listings that sellers have not indicated to be included in the sponsored slots), and ranks them in the organic slots in decreasing order of predicted relevance. If the number of listings available for sponsored slots (i.e., number of ads) is lower than the number of sponsored slots, then some sponsored slots are not utilized and are simply used for organic listings.

3.1 Field Experiment Details

A randomized experiment was run at Flipkart with about 1.2 million mobile app users for eight days February 2016. In this experiment, users were randomly allocated to one of three "buckets" with ads served in the following sets of positions (see Figure 4 for a graphical depiction) in the ladder of results:[10]

[10]Although a fourth bucket of users was also a part of this experiment, we remove it from our analysis as some execution problems were detected later for that bucket.

Position	Bucket 1	Bucket 2	Bucket 3
1	Organic 1	Organic 1	Organic 1
2	Organic 2	Ad 1	Organic 2
3	Organic 3	Organic 2	Ad 1
4	Organic 4	Organic 3	Organic 3
5	Organic 5	Organic 4	Organic 4
6	Organic 6	Ad 2	Organic 5
7	Organic 7	Organic 5	Ad 2
8	Ad 1	Organic 6	Organic 6
9	Organic 8	Organic 7	Organic 7
10	Organic 9	Ad 3	Organic 8

Figure 4: Allocation of ads in buckets at the top 10 positions

Bucket 1: 8, 12, 18, 22, 28 ...

Bucket 2: 2, 6, 10, 14, 18, 22 ...

Bucket 3: 3, 7, 13, 17, 23, 27 ...

Neither the sellers nor the consumers were aware of this experiment. The experiment design allows us to observe the sponsored listings and the displaced organic listings at the same positions across searches in different buckets. Using this we can identify the differences in the performance of sponsored and organic listings at the same positions. We can also leverage the exogenous variation in the proportion of ads users see to find the effect of serving sponsored listings on the overall clicks and conversions from a search. As Bucket 1 users do not get served any ads in the top seven positions, we use this bucket as our control group, and the other two buckets as treatment groups.[11]

3.2 What We Expect to Observe — Hypotheses and Predictions

Based on the theoretical results in Section 2, the description of our empirical setting and the experiment design, we expect to observe following results in our empirical analysis.[12]

[11] Although it would have been preferable to have a control group with no ads, Flipkart did not want to lose significant amount of revenues during the experiment.

[12] We note that the auction mechanism considered in the theoretical model deviates from the idiosyncratic flat pricing mechanism of Flipkart. Hence, we also analyzed a model with flat pricing mechanism which yields similar theoretical results.

First, based on Corollary 1, we hypothesize that in a scenario with low information asymmetry, a sponsored listing is expected to be clicked and converted at a lower rate than the organic listing it replaces, all else equal. On the other hand, in a scenario with high information asymmetry, a sponsored listing is expected to be clicked and converted at a higher rate than the organic listing it replaces, all else equal. Therefore, we state the following hypotheses:

Hypothesis 1

H1-Low: *In a low information asymmetry scenario, we expect that clicks and conversions at Position 2 in Bucket 2 and Position 3 in Bucket 3 (which are ads) will be lower than the clicks and conversions at Positions 2 and 3 in Bucket 1, respectively.*

H1-High: *In a high information asymmetry scenario, we expect that clicks and conversions at Position 2 in Bucket 2 and Position 3 in Bucket 3 (which are ads) will be higher than the clicks and conversions at Positions 2 and 3 in Bucket 1, respectively.*

Second, based on Corollary 2, we hypothesize that in a scenario with low information asymmetry, an organic listing that appears after a sponsored listing is expected to be clicked and converted at a higher rate than the organic listing that appears if the previous listing were not a sponsored listing, all else equal. On the other hand, in a scenario with high information asymmetry, an organic listing that appears after a sponsored listing is expected to be clicked and converted at a lower rate than the organic listing that appears if the previous link were not a sponsored listing, all else equal. Therefore, we state the following hypotheses:

Hypothesis 2

H2-Low: *In a low information asymmetry scenario, we expect that clicks and conversions at Position 3 in Bucket 2 and Position 4 in Bucket 3 (which follow ads) will be higher than the clicks and conversions at Positions 3 and 4 in Bucket 1, respectively.*

H2-High: *In a high information asymmetry scenario, we expect that clicks and conversions at Position 3 in Bucket 2 and Position 4 in Bucket 3 (which follow ads) will be lower than the clicks and conversions at Positions 3 and 4 in Bucket 1, respectively.*

The idea of doing a position-wise comparison as stated in the hypotheses above is the following. Consider Buckets 1 and 2. Everything is identical for these buckets until Position 2, and users are randomly allocated to these buckets. Therefore, any differences at Position 2 between these buckets in clicks and conversions must be due to the fact that a sponsored listing appears at Position 2 in Bucket 2 but an organic listing appears at Position 2 in Bucket 1; similarly, any differences between these buckets at Position 3 (and also Positions 4 and 5) in clicks and conversions can be attributed to the difference at Position 2. The same argument holds for comparison between Buckets 1 and 3 for Position 3 (with ad) and Positions 4, 5 and 6 (following ads).

We note that, in online advertising and online retailing settings, both click-through rate (click per impression) and conversion rate (conversion per impression) are quite small and have a high variance, and this problem is especially severe for conversion. Therefore, we would require hundreds of thousands of observations for an experiment seeking to detect these small effects to have sufficient power, and our dataset is indeed of this order. Nevertheless, we would be more hopeful about detecting the effects on clicks (which are observed more frequently), than about detecting effects on conversions. We also expect that as we test effects at positions that are further down, because they have lesser user traffic, detecting an effect if it exists will be more difficult, i.e., power will be lower. Under conditions where the click-through and conversion rates are relatively lower (such as our high information asymmetry case), detecting the effect if it exists will again be more difficult. The power analyses are presented in Section 4 as applicable.

Finally, if we consider total clicks and conversions after a search, we note that the effects hypothesized in Corollaries 1 and 2 are countervailing effects in terms of clicks and conversions. This implies that if the sponsored listing gets fewer clicks and conversions than the organic listing it replaces, then the following listing gets greater clicks and conversions, and vice versa. In sum, we expect relatively smaller effects (or negligible effects if search costs are low) of showing ads on clicks and conversions in both low and high information asymmetry scenarios. This aligns with the result in Corollary 3 which states that the effect on total clicks and conversions will be small if the search cost is small. We make the following prediction regarding the impact of including sponsored ads on overall search performance.

Prediction 1

In a low (high) information asymmetry scenario, we expect a very small negative (positive) difference in the number of clicks and conversions between Buckets 2 and 1, and Buckets 3 and 1 , if the search cost is relatively low.

3.2.1 Categories with Different Levels of Information Asymmetry

In an ideal setting, to understand the role of information asymmetry faced by a platform, one could compare the results of the experiment described above across markets that are identical in all aspects except in the degree of information asymmetry. However, in practice, it is almost impossible to observe such variation.

In the absence of such variation across markets, we exploit the variation in ranking performance of a marketplace across two highly distinct product categories — electronics and clothing. The levels of information asymmetry for electronics and clothing are significantly different at Flipkart. Electronics includes subcategories (called "verticals" at Flipkart) such as mobile phone, laptop, television, air conditioner, headphone, speakers, etc., while clothing includes subcategories such as shirt, top, saree, t-shirt, fabric, etc., and these two categories together account for approximately 80% of Flipkart's business. It is typically easier to evaluate the relevance of products with more attributes that are digitally encodable (Lal and Sarvary 1999), and electronics and clothing have relatively higher and lower proportions of digital attributes, respectively. Therefore, information asymmetry is expected to be lower in the category of electronics than in the category of clothing. Another source of information asymmetry in the clothing category is the huge number of products available, compared to the electronics category, making it harder for the platform to evaluate the true relevance of all the listings. For example, in our sample, the assortment size of clothing-related products is approximately 20 times the size of electronics-related products.[13] As pointed out earlier, discussions with executives at Flipkart reveals that their ranking mechanisms perform better for electronics related searches as compared to searches for clothing. Given the difficulty of obtaining data for the ideal empirical test, exploiting the variation in the level of ranking difficulty between the two categories is a feasible way to empirically investigate how information asymmetry affects

[13]We analyze the cumulative distribution of conversions across positions in the two categories. We find that approximately 75% of all conversions from the top 50 positions in electronics-related searches take place in the top 10, while the same proportion of conversions happens at position 27 in clothing-related searches. This highlights the fact that it is much easier for a platform to find relevant products in electronics than in clothing, indicating that it has more information regarding the relevance of electronic products.

		B1	B2	B3	p-value
Total	Users	391,866	390,393	390,260	.133
	Searches per user	3.32	3.33	3.32	.535
Electronics	Users	289,228	288,564	288,245	.418
	Searches per user	2.99	2.99	2.98	.250
Clothing	Users	173,890	172,839	173,162	.188
	Searches per user	2.53	2.53	2.53	.589

(a) Users and Searches Per User

		B1	B2	B3	p-value
Electronics	Click (%)	6.41	6.45	6.44	.02
	Conversion (%)	.345	.348	.349	.54
Clothing	Click (%)	2.93	2.93	2.92	.42
	Conversion (%)	.214	.216	.212	.45

(b) Click-through and Conversion Rates (Two-week Period Before Experiment)

Table 1: Balance Across Buckets

the relevance of sponsored listings.

3.3 Data

We have data on a total of 391,866, 390,393 and 390,260 users in Buckets 1, 2 and 3 (denoted by B1, B2 and B3), respectively. First, to make sure that users were properly assigned into each condition, we conduct a randomization check by testing the balance in observations across the three buckets. Table 1(a) presents the statistics for number of users and searchers per user in our full sample. The chi-square tests reveal that buckets are well balanced as we fail to reject the equality of observations at overall 5% significance level. We also analyze the impression-level click and conversion behavior of users in each bucket for the two weeks before the experiment. We present the results in Table 1(b). We again fail to reject the null that all click and conversion means are equal at overall 5% significance level.[14]

We now provide a general understanding of user behavior in the electronics and clothing categories as reflected in our data. In Table 2 we present the means (with standard deviations in parentheses) for search-level statistics. First, in the electronics category, there are a total of 613,404

[14]This is after Bonferroni correction — all the individual z-tests' p-values are larger than .0083 ($= \frac{.05}{2k}$, where $k = 3$) for electronics click-through rate — to control for familywise error rate due to multiple hypothesis testing.

	N	Depth	Clicks	Conversions
Electronics	613,404	14.3	.56	.04
		(15.4)	(.93)	(.21)
Clothing	659,936	29.3	.69	.05
		(18.5)	(1.33)	(.27)

Table 2: Search-level Summary Statistics

	N	Click	Conversion
Electronics	8,004,443	.043	.0030
		(.204)	(.055)
Clothing	17,795,318	.026	.0019
		(.158)	(.044)

Table 3: Impression-level Summary Statistics

searches.[15] Across these searches, the mean depth of search is 14.3 with a standard deviation of 15.4, the average number of clicks is 0.56 (56%) with a standard deviation of 0.93, and the mean conversions is 0.04 (4%) with a standard deviation of 0.21. In the clothing category, there are a total of 659,936 searches. The mean depth of search across these searches is 29.3 with a standard deviation of 18.5, the mean clicks is 0.69 (69%) with a standard deviation of 1.33, and the mean conversions is 0.052 (5.2%) with a standard deviation of 0.27. We note the large standard deviations for search depth, clicks and conversions for both categories. We also present the impression-level summary statistics in Table 3. We again note very low conversion probabilities in both categories, resulting in a very high coefficient of variation (which is, approximately, 20).

Table 4 presents summary statistics for the rating and number of ratings for sponsored and organic listings in the two categories. We note that, in electronics the average rating for the sponsored listings is lower than the average rating for the organic listings, while it is the reverse in clothing; this is in line with our theoretical result that in a low information asymmetry scenario lower value products will be advertised, while in a high information asymmetry scenario higher value products will be advertised. We can also see that ads in both the electronics and clothing

[15] We exclude searches where ads are not served from our analysis, as we only want to exploit the variation in the number of ads due to being in different buckets, and not due to some outside demand or supply side reason. About 32.7% of all searches remain in each bucket. All searches are truncated at position 50. Around 32% of all clothing searches are of depth 50 or above, whereas the same number is 8% for electronics.

		Rating		Rating Count	
		Mean	SD	Mean	SD
Electronics	Sponsored	3.7	0.8	142	600
	Organic	3.8	0.7	2468	5106
Clothing	Sponsored	3.7	1.1	4.0	18.2
	Organic	3.3	1.1	14.1	44.4

Table 4: Sponsored vs. Organic Characteristics

	B1	B2	B3
Overall	.090	.224	.174
Electronics	.059	.214	.165
Clothing	.119	.233	.183

Table 5: Average Proportion of Ads Shown in each Bucket

categories have on average lower numbers of ratings than their organic counterparts, indicating that sponsored products are relatively new.

Next, we consider the proportions of ads received by the users in different buckets, which is shown in Table 5. We see that at the overall level (i.e., combining the two categories) the difference between the lightest treatment (9% in Bucket 1) and the heaviest treatment (22.4% in Bucket 2) is approximately 13.5% points. This difference is bigger (15% points) for electronics and smaller (11% points) for clothing; this is due to the difference in search depth distribution across the two categories.

4 Empirical Analysis

In this section, we empirically analyze the hypotheses and predictions that we developed in Section 3.2.

4.1 Empirical Results

4.1.1 Effect at the Sponsored Slot

We consider how the presence of sponsored listings affects clicks and conversions at that position. Hypothesis 1 states that for low (high) information asymmetry clicks and conversions at a sponsored listing will be lower (higher) than at an organic listing that it displaces. To test this, we first compare the mean clicks at Position 2 in Buckets 2 and 1, and at Position 3 in Buckets 3 and 1.

Position	Click	Conversion	*DF*
2 (B2−B1)	−.0174*** [-21%]	−.0024*** [-37%]	363,600
3 (B3−B1)	−.0026*** [-5%]	−.0010*** [-25%]	323,780

(a) Electronics

Position	Click	Conversion	*DF*
2 (B2−B1)	.0079*** [19%]	.0004* [11%]	407,255
3 (B3−B1)	.0076*** [25%]	.0003	400,896

(b) Clothing

Table 6: Differences at Sponsored Slots

Tables 6(a) and 6(b) present the differences in mean clicks and conversions at these positions for electronics (low information asymmetry) and clothing (high information asymmetry), respectively. (Throughout the paper, *, ** and *** denote a p-value of < 0.1, < 0.05 and < 0.01, respectively.) The first column shows the difference in mean clicks, while the second column shows the difference in mean conversions. For instance, the first column and the first row (titled 2) shows that the mean click on the sponsored listing at Position 2 in Bucket 2 is lower than the mean click on the organic listing at Position 2 in Bucket 1 by $-.0174$ and this is significant at the 1% level; the number -21% shows that the mean click is lower by 21%. The numbers in the first column of Tables 6(a) and 6(b) show that the prediction from our theoretical model is verified for clicks. The second column reports the difference in the mean conversions, in the same format, for the electronics and clothing categories. Table 6(a) shows that our prediction holds for conversions in electronics, while Table 6(b) shows that in clothing it only holds for Bucket 2; for Bucket 3, the number is directionally in agreement though not statistically significant at the 10% level. However, as we had noted earlier, conversions are very rare in our data, especially in the clothing category, which makes these patterns difficult to detect for this category.

To summarize, the performance of electronics ads is lower than organic listings, as they are clicked and convert with 5%-21% and 25%-37% lower probability, respectively, compared to the organic listings at the same position. On the other hand, the click-through rate of ads in the clothing category is (19%-25%) higher at both positions than that of the organic listings they

Position	Click	Conversion	*DF*
3 (B2−B1)	.0147*** [29%]	.0015*** [38%]	367,503
4 (B3−B1)	.0034*** [8%]	.0005***[19%]	284,549

(a) Electronics

Position	Click	Conversion	*DF*
3 (B2−B1)	.0018*** [6%]	.0001	415,451
4 (B3−B1)	.0007	−.00009	412,292

(b) Clothing

Table 7: Differences at Organic Slots that Follow Sponsored Slots

substitute, while the mean conversion is either higher than (at 10% significance level in Bucket 2) or not different from organic listings in the same positions.[16] Overall, we find strong empirical support for Hypothesis 1. Note that even though ads do relatively worse than organic listings they displace, the advertisers are better off due to inflation of their ranks in the search list which significantly increases their visibility and conversions.

4.1.2 Effect at the Position After the Sponsored Slot

We now consider how the presence of sponsored listings affects clicks and conversions at the organic listings at the succeeding positions. According to Hypothesis 2, we expect this effect to be positive (negative) for low (high) information asymmetry scenarios. To test this, we first compare clicks at Position 3 in Buckets 2 and 1, and at Position 4 in Buckets 3 and 1. Tables 7(a) and 7(b) present the differences in mean click and conversion at these positions for electronics (low information asymmetry) and clothing (high information asymmetry) categories, respectively.

The mean clicks of organic listings appearing after sponsored listings in electronics searches are higher by 8%-29% and the mean conversions are higher by 19%-38% in Buckets 2 and 3. This provides strong support for Hypothesis H2-Low. However, except for clicks at the third position in Bucket 2, we detect essentially no effect on clicks or conversions on the succeeding position

[16] Power analysis, which can be found in the appendix (Table A1), reveals that we do not have sufficient power ($< 50\%$) to detect 10% effects in conversion probability, given our sample size. However, we have close to 100% power in our tests done in Table 6 for clicks.

	Clicks	Conversion	DF
Overall	.0038* [.6%]	.0003	1,273,338
Electronics	.0021	.0003	659,934
Clothing	.0055* [.8%]	.0003	613,402

Table 8: Effect of Ads on Search Performance

in clothing-related searches for both the buckets (the difference is not statistically significantly different from 0), i.e., we do not find support for Hypothesis H2-High. As we discussed earlier, effects are difficult to detect in the high information asymmetry category of clothing due to a weak signal as click activity is distributed over more positions as compared to the low information asymmetry category of electronics, and effects are even more difficult to detect for conversions due to the small probability of conversion.

4.1.3 Effect on Overall Search Performance

We now test Prediction 1, which implies that the number of clicks and conversions will be similar across the three buckets in both the low and the high information asymmetry scenarios. The fact that search depth is not significantly different and relatively high across the three buckets in both categories (Table A2 in Appendix) is already indicative of this. More formally, we run t-tests to estimate the differences in mean performance metrics between the treated groups (searches from users in Buckets 2 and 3) and the control group (searches from users in Bucket 1).

We report the results for both clicks and conversion in Table 8. We find that, compared to Bucket 1, the other buckets on average have no statistically significant difference in clicks in the electronics (low information asymmetry) category, while there is only a small positive difference (significant at the 10% level) in clicks in the clothing (high information asymmetry) category, as expected. Overall, there is a small and positive effect (significant at the 10% level). When considering the mean conversion, we find no statistically significant differences between treated and control groups for both the categories (and overall).

The results show that even serving close to 15% points higher proportion of ads does not significantly affect the performance at the search level in electronics. One of the reasons behind this could be the low search costs of users who continue to search after encountering poor quality ads, and purchase from succeeding (or preceding) listings. The positive effect on the organic listings at succeeding positions, as identified in Hypothesis 2-Low, provides a basis for this argument. These positive effects compensate almost all the loss in conversion and clicks at the sponsored slots. We also note that if the positions where ads appear are high, it is more likely that users continue with the search, as the marginal advantage of search is typically higher at higher positions. Instead, if ads appear further down the list, a lower fraction of users are expected to continue with the search. In the latter case, however, the loss to the platform due to the presence of poor ads is also low, as these ads now substitute only relatively less relevant listings. Having said that, it is important to note that the overall effect also depends on the number of ads in a search. If the proportion of ads becomes extremely high, the negative effect of poor quality ads (e.g., in the case of electronics) may start dominating, as the positive spillover on fewer organic listings at succeeding positions might not be enough. As noted earlier, the impact also depends on consumers' search costs. If the search cost is very high, then it might be difficult to overcome the negative effect of ads in low information asymmetry cases.

4.2 Robustness Checks

Here, we conduct a couple of robustness checks for our empirical results.

4.2.1 No Ad Situation

In the analysis in Section 4.1.3, we obtain the impact of showing different numbers of ads on clicks and conversions from a search. An important scenario to consider would be the difference in conversions between searches with ads and without ads. Although we do not have an experimental condition in which a group of users was not served any ad, we can treat searches till Position 7 in Bucket 1 as a proxy for that condition. (The top seven listings account for approximately 60% and 30% of the total conversions in electronics and clothing, respectively.) Users in Bucket 1 did not see any ads in the top seven listings, while users in Buckets 2 and 3 saw two ads. Therefore, restricting to Position 7 for all buckets, we treat Bucket 1 as our control group, and Buckets 2 and 3 as our treatment group. Essentially we run the same analysis as in Section 4.1.3, but with data restricted to the top seven positions. We find a relatively small positive effect on clicks in clothing and a small negative but statistically insignificant effect on clicks in electronics. For conversion, we find no statistically significant impact in either clothing or electronics (see Table A3). These results are identical to those in Section 4.1.3.

4.2.2 Analysis at Product Vertical Level

In our main analysis, we focused on category-level analyses. Here, we present the comparison of the performances of sponsored and organic listings at product vertical (subcategory) level to show the robustness of our result at a more granular level (see Table A4). We do the analysis at Position 2 between Buckets 1 and 2, where we have maximum statistical power. We also analyze shoes, a clothing accessory, in this analysis to see if the results are similar to clothing. One can see that the difference between sponsored and organic is negative for electronics verticals like mobile, headphone, computer and television, whereas it is either significantly positive or insignificant for clothing-related verticals.

4.3 Alternative Explanations

Here, we discuss phenomena other than information asymmetry between the platform and the sellers, which can possibly explain our results. We do the analysis at Position 2 between Buckets 1 and 2, where we have maximum statistical power.

4.3.1 Consumer Heterogeneity

It is plausible that the heterogeneous effects that we find across the two product categories in Section 4.1 are not due to different product category characteristics, but due to consumer heterogeneity. This could be especially true if users who browse clothing products are very different from those who buy electronic products. To gain some insight into this, we repeat our analysis with only those users who browsed both electronic and clothing products during the experiment period (see Table A5). The results remain robust, as we again find a significantly negative difference for electronics, and a positive (or insignificant) difference for clothing impressions. This suggests that it is the product heterogeneity, and not consumer heterogeneity, which is driving the result.

4.3.2 Specificity of Search Queries

It is possible that the difference we observe in the relative performance of ads between clothing and electronics is because search queries in electronics are usually more specific, though this could itself be another source of difference in information asymmetry. This increases the likelihood of sponsored listings being less relevant compared to organic listings, as the pool of sponsored listings is smaller. However, Flipkart does not serve ads if the relevance of the listings in the sponsored pool does not cross a threshold relevance, and we only analyze the searches where ads were served. To further make sure that our results are not driven by the specificity of search queries, we analyze the data from searches with only less specific (more generic) search queries — we select non-branded and non-descriptive search queries that have at least 200 searches — in electronics to see if the poor performance of ads still persists. Our main results continue to hold in this situation (see Table A6).

4.3.3 Role of Observable Features

In this subsection, we investigate whether quantifiable features like price, discount, rating and number of ratings can explain the differences in the performances of sponsored and displaced organic listings found in Section 4.1. This will help us understand how important are unobservable (to the platform or to an analyst) or unquantifiable features, a major source of the information asymmetry, in determining the true relevance of a listing. In order to do this, we use logistic regression to regress the binary variable click or conversion on the binary indicator *Sponsored* which equals 1 if the impression belongs to Bucket 2 (treatment group), otherwise it equals 0 (for impressions in Bucket 1) at Position 2, while controlling for listing characteristics, such as rating, price, discount and number of ratings. See Table A7 in the Appendix for results on click and conversion. We find a statistically significant and negative difference between mean clicks and conversions of sponsored and organic listings in the electronics category, a statistically significant and positive difference between the click rates of sponsored and organic listings in the clothing category, and no statistically significant result between the mean conversion of sponsored and organic listings in the clothing category. All of these results are as expected. This implies that sponsored listings are not better or worse than the substituted organic listings only due to observable and quantifiable features.

5 Conclusions and Discussion

Advertising on e-commerce marketplaces, wherein sponsored listings are interleaved with organic listings, is a large and growing phenomenon. This serves as an additional source of revenue for the marketplace beyond sales commissions, but it also influences the products that consumers see when they reach the platform. In this paper, we study how useful or hurtful these sponsored ads are to consumers on the marketplaces.

Our theoretical analysis reveals that the relevance of the ads to the consumers' search queries depends on the degree of information asymmetry between the marketplace owner and the sellers selling products with different degrees of relevance. Specifically, when information asymmetry is low and the platform can distinguish high-relevance sellers from low-relevance sellers, then low-relevance sellers have a greater incentive to advertise because they know that otherwise they will not be able to obtain slots high enough in the ranked list of results. On the other hand, when information asymmetry is high and the platform in unable to distinguish high-relevance sellers from low-relevance sellers, then high-relevance sellers have a greater incentive to advertise because they know that otherwise the platform will determine the organic results essentially randomly and they have a sizable chance of being placed low in the ranked list of results, an event under which they have much to lose.

Using unique data from a large field experiment at Flipkart, one of the largest online marketplaces, we empiricially test the analytical findings. We take one low information asymmetry category, electronics, and one high information asymmetry category, clothing, and find that the empirical patterns implied by our theoretical results hold in the data. In the case of electronics, the sponsored listings have a significantly lower probability of click and conversion than that of the substituted organic listings. However, the lost conversions at the ad slots are compensated by the positive effect of ads on succeeding organic listings. In the case of clothing, sponsored ads get clicked or converted at a higher or an equal rate compared to that of the organic listings they displace. We find that, at the search level, the outcomes with and without sponsored ads are not very different — it is interesting to note that although we find the overall effect on search to be negligible for both clothing and electronics, the underlying mechanisms behind this result are very different (in fact, completely opposite) for the two categories.

Our paper has important implications for e-commerce platform managers. First, we show that sponsored ads can work very differently in different settings. Interestingly, we show that the degree of information asymmetry may determine whether ads increase or decrease the quality of the overall set of results presented to a consumer — specifically, if the platform finds it difficult to distinguish between different types sellers, then introducing ads can improve the set of results, and vice versa. Therefore, for goods with non-digital attributes, allowing for sponsored listings may help marketplace firms towards solving the classic "cold start" problem (i.e., determining the quality and relevance of new products for which there is not much data), while also increasing monetization. Second, we find that, irrespective of the product category, there is no significant downside from showing ads in the search results, at least in the short term. This is because as consumers search multiple product listings, they are able to find well-matching products even if ill-matching products are displayed in the sponsored ads (as would happen in a low information asymmetry setting); at the same time, the sponsored ads bring additional ad fees. This implies that sponsored products are supplementing the core business of platforms by generating more revenue. Third, platform managers who are concerned about the user experience because of poor quality ads in searches related to digital products (low IA categories), may adjust bids with the quality score in deciding the winner of auctions. In case of flat pricing, they may charge lower prices to incentivize even high-relevance sellers to advertise.

Our paper has several limitations that can be addressed in future research. First and foremost, we could strengthen our empirical analysis if we were able to obtain more variation in information asymmetry across multiple settings. Second, our paper looks at only the short-term impact of showing ads in the search results. It will be interesting to examine how the relevance of sponsored listings impacts user behavior and the platform in the long term, where the results may be different. Third, we do not explore the aspects of optimal positions for ads or the optimal number of sponsored ads. Choi and Mela (2019) study this, and advancing on this front would be an important addition to the as yet small literature on advertising on online marketplaces.

References

[1] Abhishek, V., Jerath, K., and Zhang, Z. J. (2015). Agency selling or reselling? Channel structures in electronic retailing. *Management Science*, 62(8), 2259-2280.

[2] Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience goods markets: an empirical examination. *International Economic Review*, 44(3), 1007-1040.

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

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

[5] Blystone, D. (2019). Understanding the Alibaba Business Model. `https://www.investopedia.com/articles/investing/062315/understanding-alibabas-business-model.asp`

[6] Borison, R. (2015). EBay Launches Promoted Listings Ads to Help Sellers Reach Buyers. `https://www.thestreet.com/story/13163593/1/ebay-launches-promoted-listings-ads-to-help-sellers-reach-buyers.html`

[7] Choi, H., and Mela, C. F. (2019). Monetizing online marketplaces. *Marketing Science*.

[8] Economist (2018). Amazon's ambitious drive into digital-advertising. `https://www.economist.com/business/2018/10/27/amazons-ambitious-drive-into-digital-advertising`

[9] Edelman, B., Ostrovsky, M., and Schwarz, M. (2007). Internet advertising and the generalized second-price auction: selling billions of dollars worth of keywords. *American Economic Review*, 97(1), 242-259.

[10] Feng, J., and Xie, J. (2012). Research Note— Performance-Based Advertising: Advertising as signals of product quality. *Information Systems Research*, 23(3-part-2), 1030-1041.

[11] Gabaix, X., and Laibson, D. 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, 121(2), 505-540.

[12] Hagiu, A., and Wright, J. (2014). Marketplace or reseller?. *Management Science*, 61(1), 184-203.

[13] Ho, T.-H., Lim, N. and Camerer, C.F. (2006). How "psychological" should economic and marketing models be? *Journal of Marketing Research*, 43(3), 341-344.

[14] Jerath, K., Ma, L., Park, Y.-H. and Srinivasan, K. (2011). A "position paradox" in sponsored search auctions. *Marketing Science*, 30(4), 612-627.

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

[16] Jiang, B., Jerath, K. and Srinivasan, K. (2011). Firm strategies in the "mid tail" of platform-based retailing. *Marketing Science*, 30(5), 757-775.

[17] Keyes, D. (2019). 3rd-party sellers are thriving on Amazon. `https://www.businessinsider.com/amazon-third-party-sellers-record-high-sales-2019-5`

[18] Kihlstrom, R. E., and Riordan, M. H. (1984). Advertising as a Signal. *Journal of Political Economy*, 92(3), 427-450.

[19] Lal, R., and Sarvary, M. (1999). When and how is the Internet likely to decrease price competition?. *Marketing Science*, 18(4), 485-503.

[20] Liu, D., and Viswanathan, S. (2014). Information asymmetry and hybrid advertising. *Journal of Marketing Research*, 51(5), 609-624.

[21] Long, F., Jerath, K., and Sarvary, M. (2019). The Informational Role of Sponsored Advertising on Online Retail Marketplaces, Working paper, University of North Carolina, Chapel-Hill.

[22] Lunden, I. (2018). Amazon's share of the US e-commerce market is now 49%, or 5% of all retail spend. `https://techcrunch.com/2018/07/13/amazons-share-of-the-us-e-commerce-market-is-now-49-or-5-of-all-retail-spend/`

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

[24] Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, 82(4), 729-754.

[25] Sahni, N. S., and Nair, H. (2019). Does advertising serve as a signal? Evidence from field experiments in mobile search. *Review of Economic Studies*, forthcoming.

[26] Shelley, M. (2019). India Ecommerce Market: $65 Billion By 2023, Flipkart To Leapfrog Amazon. https://dazeinfo.com/2019/03/30/india-ecommerce-market-flipkart-amazon-2023/

[27] Shields, M. (2017). Amazon and Pinterest Threaten to Shake Up the Search Ad Market. https://www.wsj.com/articles/amazon-and-pinterest-threaten-to-shake-up-the-search-ad-market-1488798004

[28] Soper, T. (2019). Amazon's big new business: Here's how much advertising revenue the company generated in 2018. https://www.geekwire.com/2019/amazons-big-new-business-heres-much-advertising-revenue-company-generated-2018/

Appendices

A Appendix: Derivations and Proofs

A.1 Expressions for δ, θ, β and π

We first derive the expressions for δ_H and δ_L.

$\delta^H = \Pr(V_L - u_H > s) = \Pr(\epsilon_H < V_L - V_H - s) = \frac{V_L - V_H - s + 1}{2}$, as $\epsilon_H \sim U[-1,1]$. Similarly,

$$\delta^L = \Pr(\hat{V} - u_{L_i} > s) = \Pr(\epsilon_{L_i} < \hat{V} - V_L - s) = \begin{cases} \frac{(\frac{V_H - V_L}{2}) - s + 1}{2} & \text{without the sponsored listing,} \\ \frac{(\frac{1+3\kappa}{4})(V_H - V_L) - s + 1}{2} & \text{with the sponsored listing} \end{cases}$$

$\forall\, i \in \{1, 2\}$. We assume $s < 1 - (V_H - V_L)$ so that $\delta^H > 0$.

Next, we derive the expressions for θ_H, θ_L, θ_{HL}, θ_{LH} and θ_{LL}.

$\theta_i = \Pr(u_i > 0) = \Pr(\epsilon_i > -V_i) = \frac{1+V_i}{2}\ \forall\, i \in \{H, L_1, L_2\}$.

$\theta_{HL} = \Pr(u_H > \max\{u_{L_i}, 0\}) = \Pr(\epsilon_H - \epsilon_{L_i} > V_L - V_H, \epsilon_H > -V_H) = \frac{(1+V_H)(3+V_H-2V_L)}{8}$,

$\theta_{LH} = \Pr(u_{L_i} > \max\{u_H, 0\}) = \Pr(\epsilon_{L_i} - \epsilon_H > V_H - V_L, \epsilon_{L_i} > -V_L) = \frac{(1+V_L)(3-2V_H+V_L)}{8}$ and

$\theta_{LL} = \Pr(u_{L_i} > \max\{u_{L_j}, 0\}) = \Pr(\epsilon_{L_i} - \epsilon_{L_j} > 0, \epsilon_{L_i} > -V_L) = \frac{(1+V_L)(3-V_L)}{8}\ \forall\, i, j \in \{1, 2\}$ and $i \neq j$, as ϵ_k and $\epsilon_k - \epsilon_l$ has a joint density $\frac{1}{4}$ with support $[-1, 1]$ and $[\epsilon_k - 1, \epsilon_k + 1]$, respectively, where $k, l \in \{H, L_1, L_2\}$ and $k \neq l$.

Now we can derive the following expressions.

$$\pi_1^{HL} = \gamma_H \cdot \beta_1^{HL} = \gamma_H((1-\delta^H)\theta_H + \delta^H \theta_{HL})$$

$$\pi_2^{HL_1} = \pi_2^{HL_2} = \pi_2^{HL} = \gamma_L \cdot \beta_2^{HL} = \gamma_L \cdot \delta^H \cdot \theta_{LH}$$

$$\pi_2^{LH} = \gamma_H \cdot \beta_2^{LH} = \gamma_H \cdot \delta^L \cdot \theta_{HL}$$

$$\pi_1^{LH} = \gamma_L \cdot \beta_1^{LH} = \gamma_L((1-\delta^L)\theta_L + \delta^L \cdot \theta_{LH})$$

$$\pi_1^{LL} = \gamma_L \cdot \beta_1^{LL} = \gamma_L((1-\delta^L)\theta_L + \delta^L \cdot \theta_{LL})$$

$$\pi_2^{LL} = \gamma_L \cdot \beta_2^{LL} = \gamma_L \cdot \delta^L \cdot \theta_{LL}$$

Note that all the βs and πs are positive.

A.2 Derivation of v_H and v_L

If H wins the auction, its payoff is given by π_1^{HL}. Instead, if L_i wins the auction, H's payoff is $(\kappa + \frac{1-\kappa}{4})\pi_2^{LH}$, as $\kappa + \frac{1-\kappa}{4}$ is the probability with which H is ranked higher than L_j, where $i, j \in \{1, 2\}$ and $i \neq j$. Then, taking the difference between these two payoffs we obtain:

$$v_H = \frac{\pi_1^{HL} - (\kappa + \frac{1-\kappa}{4})\pi_2^{LH}}{\beta_1^{HL}} = \gamma_H\Big(1 - \left(\kappa + \frac{1-\kappa}{4}\right)\frac{\beta_2^{LH}}{\beta_1^{HL}}\Big) > 0, \tag{A1}$$

as $\kappa \leq 1$ and $\frac{\beta_2^{LH}}{\beta_1^{HL}} < 1$.

When the L-types win the auction, each L gets the sponsored slot with probability $\frac{1}{2}$, H occupies the second position with probability $\kappa + \frac{1-\kappa}{4}$ and the other L occupies the second position with probability $\frac{3(1-\kappa)}{4}$. If the L-types lose the auction to H, then the payoff of each L is given by $\frac{\pi_2^{HL}}{2}$, as they occupy the second position with probability $\frac{1}{2}$. If they win the auction, L gets a payoff of $\frac{\pi_1^{LH}}{2}$ when H occupies the second position. Instead, when an L occupies the second position, L gets a payoff of $\frac{\pi_1^{LL}}{2} + \frac{\pi_2^{LL}}{2}$. Then, v_L is given by the weighted sum of the values with H and L at the second position.

$$\begin{aligned} v_L &= (\kappa + \frac{1-\kappa}{4})\frac{(\frac{\pi_1^{LH}}{2} - \frac{\pi_2^{HL}}{2})}{\frac{\beta_1^{LH}}{2}} + \frac{3(1-\kappa)}{4}\frac{(\frac{\pi_1^{LL}}{2} + \frac{\pi_2^{LL}}{2} - \frac{\pi_2^{HL}}{2})}{\frac{\beta_1^{LL}}{2}} \\ &= \gamma_L\Big(1 - \left(\frac{1+3\kappa}{4}\right)\frac{\beta_2^{HL}}{\beta_1^{LH}} + \frac{3(1-\kappa)}{4}\frac{(\beta_2^{LL} - \beta_2^{HL})}{\beta_1^{LL}}\Big) > 0, \end{aligned} \tag{A2}$$

as $\kappa \leq 1$, $\frac{\beta_2^{HL}}{\beta_1^{LH}} < 1$ and $\frac{\beta_2^{LL}}{\beta_2^{HL}} > 1$.

A.3 Proof of Lemma 1

We have

$$\frac{dv_H}{d\kappa} = -\frac{3\gamma_H\beta_2^{LH}}{4\beta_1^{HL}} < 0, \tag{A3}$$

and

$$\frac{dv_L}{d\kappa} = -\frac{3\gamma_L\beta_2^{HL}}{4\beta_1^{LH}} - \frac{3\gamma_L(\beta_2^{LL} - \beta_2^{HL})}{4\beta_1^{LL}} < 0, \tag{A4}$$

as $\beta_2^{LL} > \beta_2^{HL}$.

A.4 Proof of Proposition 1

We already know from Lemma 1 that both v_H and v_L monotonically decrease with κ. Therefore, if we have $v_L > v_H$ at $\kappa = 1$ and $v_H > v_L$ at $\kappa = \frac{1}{3}$, then there always exists $\overline{\kappa} \in (\frac{1}{3}, 1)$ such that $v_H \geq v_L \ \forall\ \kappa \in [\frac{1}{3}, \overline{\kappa}]$ and $v_L > v_H \ \forall\ \kappa \in (\overline{\kappa}, 1]$.

We have $v_L > v_H$ at $\kappa = 1$ iff:

$$\gamma_L \left(1 - \frac{\beta_2^{HL}}{\beta_1^{LH}}\right)\Bigg|_{\kappa\,=\,1} > \gamma_H \left(1 - \frac{\beta_2^{LH}}{\beta_1^{HL}}\right)\Bigg|_{\kappa\,=\,1}, \text{ or}$$

$$\frac{\gamma_H}{\gamma_L} < \left.\frac{1 - \frac{\beta_2^{HL}}{\beta_1^{LH}}}{1 - \frac{\beta_2^{LH}}{\beta_1^{HL}}}\right|_{\kappa\,=\,1}. \tag{A5}$$

Similarly, $v_H > v_L$ at $\kappa = \frac{1}{3}$ iff:

$$\gamma_H \left(1 - \frac{\beta_2^{LH}}{2\beta_1^{HL}}\right)\Bigg|_{\kappa\,=\,\frac{1}{3}} > \gamma_L \left(1 - \frac{\beta_2^{HL}}{2\beta_1^{LH}} + \frac{\beta_2^{LL} - \beta_2^{HL}}{2\beta_1^{LL}}\right)\Bigg|_{\kappa\,=\,\frac{1}{3}}, \text{ or}$$

$$\frac{\gamma_H}{\gamma_L} > \left.\frac{\left(1 - \frac{\beta_2^{HL}}{2\beta_1^{LH}} + \frac{\beta_2^{LL} - \beta_2^{HL}}{2\beta_1^{LL}}\right)}{\left(1 - \frac{\beta_2^{LH}}{2\beta_1^{HL}}\right)}\right|_{\kappa\,=\,\frac{1}{3}}. \tag{A6}$$

Without loss of generality, we assume $V_H = \frac{1}{2}$ to obtain analytical solutions. There exist parameter ranges where both (A5) and (A6) hold together. However, there does not exist any parameter range such that $v_H > v_L$ at $\kappa = 1$ and $v_L > v_H$ at $\kappa = \frac{1}{3}$.

A.5 Proof of Corollary 1

When L occupies the sponsored slot, then H occupies the second slot with probability $(\kappa + \frac{1-\kappa}{4})$, and the other L listing occupies with probability $\frac{3(1-\kappa)}{4}$. Therefore, the effect at the first position in the case with low information asymmetry is given by

$$\begin{aligned}&\left(\kappa + \frac{1-\kappa}{4}\right)\pi_1^{LH} + \frac{3(1-\kappa)}{4}\pi_1^{LL} - \left(\kappa\pi_1^{HL} + \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_1^{LL})\right)\\ &= -\kappa(\pi_1^{HL} - \pi_1^{LH}) + \left(\frac{1-\kappa}{4}\right)(\pi_1^{LL} - \pi_1^{LH}) < 0,\end{aligned} \tag{A7}$$

as $\pi_1^{LH} < \pi_1^{LL} < \pi_1^{HL}$.

When H occupies the sponsored slot, then the second slot is always occupied by L. The effect at the first position in the case with information asymmetry is given by

$$\pi_1^{HL} - (\kappa\pi_1^{HL} + \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_1^{LL}))$$
$$= (1-\kappa)\pi_1^{HL} - \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_1^{LL}) > 0, \tag{A8}$$

as $\pi_1^{LH} < \pi_1^{LL} < \pi_1^{HL}$.

A.6 Proof of Corollary 2

When L occupies the sponsored slot, the effect at the second position in the case with low information asymmetry is given by

$$\left(\kappa + \frac{1-\kappa}{4}\right)\pi_2^{LH} + \frac{3(1-\kappa)}{4}\pi_2^{LL} - \left(\kappa\pi_2^{HL} + \frac{(1-\kappa)}{2}(\pi_2^{LH} + \pi_2^{LL})\right) > 0, \tag{A9}$$

as $\pi_2^{LH} > \pi_2^{LL} > \pi_2^{HL}$.

When H occupies the sponsored slot, the effect at the second position in the case with information asymmetry is given by

$$\pi_2^{HL} - (\kappa\pi_2^{HL} + \frac{(1-\kappa)}{2}(\pi_2^{LH} + \pi_2^{LL}) = (1-\kappa)\pi_2^{HL} - \frac{(1-\kappa)}{2}(\pi_2^{LH} + \pi_2^{LL}) < 0 \tag{A10}$$

$\forall\ \kappa < \overline{\kappa}$, as $\pi_2^{HL} < \pi_2^{LL} < \pi_2^{LH}$.

A.7 Proof of Corollary 3

When L occupies the sponsored slot in the case of low information asymmetry, the overall effect on the two positions in the search is given by
$\Delta_\kappa^L = (\kappa + \frac{(1-\kappa)}{4})(\pi_1^{LH} + \pi_2^{LH}) + \frac{3(1-\kappa)}{4}(\pi_1^{LL} + \pi_2^{LL}) - \Big(\kappa(\pi_1^{HL} + \pi_2^{HL}) + \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_2^{LH}) + \frac{(1-\kappa)}{2}(\pi_1^{LL} + \pi_2^{LL})\Big) < 0\ \forall\ \kappa > \overline{\kappa}$, as $\pi_1^{LH} < \pi_1^{HL}$ and $\pi_2^{LH} < \pi_2^{HL}$.

We have $\frac{\partial\Delta_\kappa^L}{\partial s} = \frac{1}{64}(2\gamma_L(\kappa(11 + V_H) - 3 - V_H)(1 + V_L) + \gamma_H(6 + V_H(4 + V_H) + 4V_L - 2V_HV_L +$

$3V_L^2 - \kappa(22 + V_H(20 + V_H) + 4V_H - 2V_HV_L + 3V_L^2))) < 0$.

When H occupies the sponsored slot in the case with high information asymmetry, the overall effect is given by $\Delta_\kappa^H = \pi_1^{HL} + \pi_2^{HL} - \left(\kappa(\pi_1^{HL} + \pi_2^{HL}) + \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_2^{LH}) + \frac{(1-\kappa)}{2}(\pi_1^{LL} + \pi_2^{LL})\right) > 0$. Δ_κ^H is always positive as the introduction of the sponsored slot removes the possibility of two low types in the first two positions.

We have, $\frac{\partial \Delta_\kappa^H}{\partial s} = \frac{1}{32}(1-\kappa)(\gamma_H(1+V_H)(5 - V_H + 2V_L) + \gamma_L(5 - 2V_H + 3V_L)(1 + V_L)) > 0$.

B Appendix: Tables of Supplementary Results

Position	Difference	Power
2	11%	49%
3	10%	37%

Table A1: Power Analysis for Clothing Conversion in Table 6 ($\alpha = 0.05$)

Category	Bucket	Mean	SD	Median
Electronics	1	14.3	15.4	7
	2	14.4	15.3	7
	3	14.4	15.4	7
Clothing	1	29.0	18.6	28
	2	29.4	18.5	30
	3	29.4	18.5	30

Table A2: Search Depth Across Buckets

	Clicks	Conversion	*DF*
Overall	.0029*** [1%]	−.0003	1,273,338
Electronics	−.0028	−.0004	613,402
Clothing	.0082*** [3.6%]	−.0003	659,934

Table A3: Effect of Ads Restricting to Top Seven Positions

Vertical	Clicks	Conversions
Electronics		
Mobile	$-.0171^{***}$ [−18%]	$-.0026^{***}$ [−35%]
Headphone	$-.0149^{***}$ [−30%]	$-.0022^{***}$ [−43%]
Computer	$-.0215^{***}$ [−26%]	.0003
Television	$-.0100^{***}$ [−20%]	$-.0024^{**}$ [−53%]
Clothing		
Saree	$.0150^{***}$ [32%]	$.0007^{*}$ [19%]
Fabric	$.0126^{***}$ [29%]	$.0007^{*}$ [19%]
Shirt	$.0028^{*}$ [7%]	.0001
Top	$.0109^{**}$ [33%]	−.0004
Shoe	$.0036^{***}$ [10%]	−.0002

Table A4: Difference at Sponsored Slot by Product Vertical

	Click	Conversion	*DF*
Electronics	$-.0113^{***}$ [-15%]	$-.0016^{***}$ [-33%]	56,703
Clothing	$.0066^{***}$ [16%]	−.0003	70,913

Table A5: Difference at Sponsored Slot for Users who browsed both Categories

	Click	Conversion	*DF*
B2−B1	$-.004^{***}$ [-8%]	$-.0015^{***}$[-40%]	162,311

Table A6: Difference at Sponsored Slot for Generic Queries in Electronics

	Clicks		Conversions	
	Electronics	Clothing	Electronics	Clothing
Sponsored	-0.205^{***}	0.106^{***}	-0.295^{***}	0.05
	(0.014)	(0.022)	(0.052)	(0.075)
Rating	0.075^{***}	0.018^{*}	0.190^{***}	0.103^{***}
	(0.015)	(0.011)	(0.059)	(0.039)
Price	0.071^{***}	0.031	-0.088^{***}	-0.270^{***}
	(0.008)	(0.022)	(0.028)	(0.077)
Discount	-0.008^{***}	0.007^{***}	-0.009^{***}	0.009^{***}
	(0.001)	(0.001)	(0.002)	(0.002)
Rating Count	-0.009^{***}	0.011	0.087^{***}	0.040
	(0.003)	(0.008)	(0.011)	(0.026
Constant	-3.281^{***}	-3.812^{***}	-5.484^{***}	-4.836^{***}
	(0.084)	(0.161)	(0.310)	(0.569)

Table A7: Difference between Sponsored and Organic listings after Controlling for Observables