Information Asymmetry and Relevance of Sponsored Listings in Online Marketplaces

Vibhanshu Abshishek
University of California, Irvine

Kinshuk Jerath
Columbia Business School

Siddharth Sharma
Carnegie Mellon University

August 2019

Information Asymmetry and Relevance of Sponsored Listings in Online Marketplaces

Abstract

Promoting listings in search results has become a common mechanism of advertising for sellers on e-commerce marketplaces like Amazon and eBay. Although sponsored product listings (ads) create an additional source of revenue for the marketplace, they might hurt the core business of the platform due to an increase in the rankings of potentially inferior listings, diluting the relevance of products shown to consumers. This effect may be more salient if the marketplace faces information asymmetry (IA) regarding the sellers' product quality or relevance. We first use a parsimonious analytical model to develop specific hypotheses that we empirically test in the paper. We show that marketplace advertising can lead sellers of low-relevance listings to advertise under no information asymmetry, but high-type sellers to advertise when there is a great deal of information asymmetry. Therefore, advertising has a potential to act as a screening mechanism for the platform when it has asymmetric information. We also show how low (high) relevance of ads positively (negatively) affects the organic listings at the succeeding positions, and negatively (positively) affects the overall search performance.

We then analyze data from a field experiment with over 2 million users on a marketplace, using two most popular product categories exhibiting different levels of information asymmetry–clothing (high IA) and electronics (low IA). In this experiment, four random groups of users are served varying number of sponsored listings at different sets of positions. Our identification relies on the the exogenous variation in both the positions and the number of ads served across the buckets. Consistent with our theory that ads are of high quality in clothing and of low quality in electronics, we first find that in the clothing category, sponsored listings perform (weakly) better than the organic listings they displace, whereas ads in the electronics category get clicked and converted with 5-25% and 25-35% lower probability, respectively. However, this leads to a positive effect on the organic listings at the succeeding positions in the case of electronics. Overall, at the search level, we find a small positive difference (1-2%) in clicks for buckets with higher proportion of ads in clothing, but do not find any statistically significant difference in clicks or conversion among the four experimental buckets in electronics, implying that the positive effect on succeeding positions in electronics compensates almost all the lost conversions at the sponsored slot.

Key words: E-commerce; Sponsored search advertising; Product heterogeneity; Platform; Asymmetric information

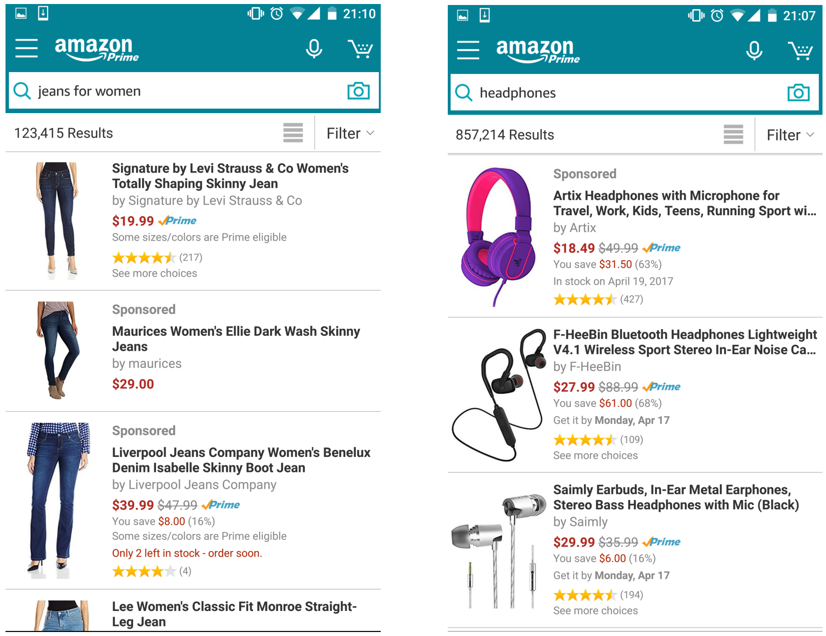
1 Introduction

Over the last decade, the e-commerce industry has been growing rapidly. In the US, the $400 billion industry now accounts for 8.5% of total retail sales compared to only 3.5% in 2007 (U.S. Department of Commerce, 2017). Amazon, the world's biggest e-commerce company by market capitalization, alone sold $136 billion worth of units in 2016. What is noteworthy is that only half of these sales were made by Amazon itself, while the rest were made through 2 million third-party sellers who participate in the Amazon marketplace platform by listing their products. Concurrently, this participation of third-party sellers has also been a growing trend, as they accounted for only a quarter of its sales in 2007.[1] eBay, another multi-billion dollar e-commerce company, sells only through third-party sellers. Thus, e-commerce marketplaces, where millions of sellers and consumers transact online, have been growing in importance in the retail industry. In fact, traditional brick-and-mortar retailers such as Best Buy and Walmart have also adopted the marketplace model (Hagiu and Wright, 2014).

Marketplaces typically charge a commission to the third-party sellers for the transactions taking place through them (Hagiu and Wright, 2014; Abhishek et al., 2015). Allowing third-party sellers to list their products on the platform free of cost allows marketplaces to offer an extensive assortment of products to their consumers. A marketplace thus has hundreds of thousands of products at its disposal for most product categories. Given that a typical user views only a limited set of products (due to search costs), a marketplace faces the daunting task of matching hundreds of related product listings to a search query in order maximize its revenue. Most e-commerce marketplaces rely on historical data and state-of-the-art machine learning algorithms to perform this matching. For every consumer query, they return a list of products sorted by predicted relevance. These listings are referred to as *organic* listings.

Since 2015, many e-commerce marketplaces like Amazon and eBay have started allowing sellers to promote their listings in the search results by artificially increasing their ranking

[1]https://www.statista.com/statistics/259782/third-party-seller-share-of-amazon-platform/



(a) (b)

Figure 1: Sponsored Listings on Amazon Mobile App

in exchange for an additional price per click or conversion.[2] These promoted listings are referred to as *sponsored* listings or *product listing ads*. Sponsored product listings appear in the midst of organic listings just like any other listing, but with a sponsored or ad identifier (see Figure 1). The proportion of sponsored slots in the search results varies between 10-20%, depending on the marketplace. The motive behind enabling these ads is to give new sellers and products more visibility, while creating an additional source of revenue.[3] It is estimated that Amazon earned USD 1.6 billion from ads in 2016, and this number is expected to grow significantly in the future.[4] Our discussion with managers at several e-commerce companies reveals that advertising is an extremely lucrative business as the margins from product sales in e-commerce are typically very thin.[5] Although Amazon's advertising revenue

[2] https://www.thestreet.com/story/13163593/1/ebay-launches-promoted-listings-ads-to-help-sellers-reach-buyers.html
https://www.wsj.com/articles/amazon-and-pinterest-threaten-to-shake-up-the-search-ad-market-1488798004

[3] https://services.amazon.com/services/sponsored-products-overview.htm

[4] http://www.newsmax.com/finance/streettalk/amazon-advertising-industry-bezos/2017/01/30/id/771034/

[5] https://www.forbes.com/sites/greatspeculations/2017/02/06/can-online-advertising-drive-revenues-for-

is a tiny fraction of its total sales, the ad revenue constitutes a significant portion of the net operating income of USD 4.2 billion.[6]

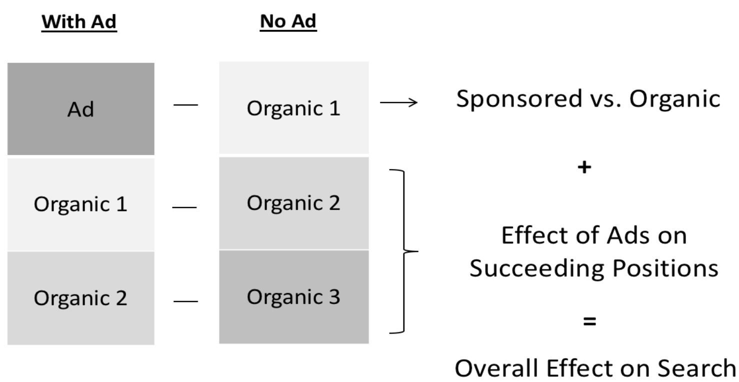


Figure 2: Individual Effects of Sponsored Listings

Although sponsored products create a new source of revenue for these marketplaces, they are 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 product search, 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 probability of a transaction from a search taking place on these platforms. The effect of sponsored listings on search performance can therefore be seen to be driven by two mechanisms. First, it is determined by the effect the sponsored listing has on the performance of the ad slot, i.e., the difference in the performances of the sponsored listings and the organic listings that would have appeared at the same position had ads been absent. Second, it is also driven by the difference in the performances of the organic listings that are preceded by a sponsored listing and the organic listings at the same positions but are not preceded by ads. We refer to this second effect as the effect of ads on organic listings at succeeding positions. See Figure 2 for a simple illustration of these two effects. The weighted sum of these two individual effects determines the extent of the overall effect on the platform's core business of transaction

amazon-in-the-long-term/58aac77d3844

[6]http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-reportsannual

commissions.

Given this setting, we try to answer the following research questions in this paper. First, how relevant are sponsored listings compared to the organic listings they displace, and how this difference varies by product category? Second, what effect sponsored listings have on the organic listings at succeeding positions? Finally, what is the effect of sponsored listings on the overall search performance?

To develop specific hypotheses, we use a simple analytical model with product listings that are heterogeneous in relevance to a consumer's search query, and allow for information asymmetry between the marketplace and the sellers, who have private information regarding the relevance of the products. Our main finding indicates that marketplace advertising results in a separating equilibrium where low-type sellers have a higher incentive to advertise under no information asymmetry, whereas high-type sellers have a higher incentive to advertise when there is a great deal of information asymmetry. In the presence of such information asymmetry, advertising serves as a screening mechanism for the marketplace, as this gives an opportunity to the low-ranked high-type sellers to promote their listings in search results. When this information asymmetry does not exist, advertising only serves the purpose of promoting an inferior listing.

To test the above theoretical predictions, we use data from a large-scale randomized field experiment on the biggest e-commerce marketplace in India. In this experiment, random groups of more than 500,000 users are served different numbers of ads at different sets of positions. The exogenous variation in the positions and the number of the ads allows us to identify the effect of sponsored listings. Specifically, we are able to randomly observe the sponsored listings and the displaced organic listings at the same position across searches from different experimental buckets. In our analysis, we consider two of the most popular product categories, *electronics* and *clothing* – with very low and high number of digital attributes, respectively (Lal and Sarvary, 1999). For products that have mostly digital attributes (e.g., electronics), the marketplace is able to sort them effectively based on the product information and historical data and the information asymmetry is relatively low. On the other hand, the marketplace finds it difficult to rank products with mostly non-digital attributes such as clothing, which results in a high level of information asymmetry.

We first find that, in the clothing category, compared to the displaced organic listings, the click-through rate of sponsored listings is higher by about 25%, and there is no statistically significant difference in the conversion probability. In contrast, ads in the electronics category perform much worse than the displaced organic listings, receiving 5-20% lower clicks, and 25-35% lower conversions. However, we also find that the poor quality ads have a positive effect on the organic listings at succeeding positions, while there is no such effect in clothing. The analysis at the search level suggests very small impacts (1-2%) on clicks in clothing and no effect on probability of conversion or clicks in electronics. This implies that most of the lost conversions at the sponsored slot in electronics are recovered due to the positive effects on succeeding positions, as users continue to search after encountering low-relevance ads.

Our paper contributes to the marketplace advertising literature that tends to focus on the demand-side signaling effects of ads (Feng and Xie, 2012; Sahni and Nair, 2016). Here we focus on the supply-side effect of sponsored products and the screening mechanism they enable (Liu and Viswanathan, 2014). Our rich e-commerce marketplace data allows us to examine the effectiveness of the mechanism across different product types, which has not been explored before in this context. In this paper, we also extend the literature on competitive and spillover effects of other organic (Agarwal et al., 2015) and sponsored (Jeziorski and Segal, 2015; Agarwal and Mukhopadhyay, 2016) listings on the performance of focal sponsored products. Here we study the impact of presence and relevance of sponsored listings on the performance of competing organic listings.

Our findings have some direct implications for e-commerce platform managers. We first identify that sponsored search program can serve as a tool for marketplace firms to screen relevant listings in categories where they face high degree of information asymmetry. In fact, promoting listings seems to be an efficient way for these marketplaces to solve the classic *cold start*[7] problem of recommender systems for products with non-digital attributes. Our study also finds that marketplaces seem to promote inferior listings as sponsored listings in product categories like electronics which can hurt consumer experience. One way to counter this problem is to lower the price of the ads which might incentivize some high-type sellers to also advertise. We last inform the managers that there is no significant opportunity cost of showing ads in their search results, irrespective of the product category, at least in the short run, as users tend to just purchase at a higher rate from other positions. One implication of this is that the revenue from sponsored products only adds to the revenue made from their core business of transaction commissions.

The rest of the paper proceeds as follows. In the next section, we discuss the related literature. In Section 3, we present a parsimonious analytical model of marketplace advertising. Section 4 describes our research setting, the experiment design, and the data used for analysis. We present our empirical strategy and findings in Section 5. We rule out alternative explanations in Section 6, and do some robustness checks in Section 7, before concluding in Section 8.

2 Relevant Literature

The paper relates to two streams of literature: (i) sponsored search advertising, and (ii) information asymmetry.

Starting from Ghose and Yang (2009), a large stream of sponsored search advertising literature has focused on studying the impact of position on the click-through rate and the conversion rate of sponsored search advertising using observational data (Rutz et al., 2012; Narayanan and Kalyanam, 2015; etc.). The general conclusion is that position has a significant negative impact on click-through rate and, therefore, on purchase probability. Recent literature on sponsored search has studied the effectiveness of paid search ads through large-scale randomized field experiments (Blake et al., 2015; Johnson et al., 2017). There are

[7] https://en.wikipedia.org/wiki/Cold_start

also studies that have looked at the effect of competition from other organic (Agarwal et al., 2015) and sponsored results (Jeziorski and Segal, 2015; Agarwal and Mukhopadhyay, 2016) on the performance of sponsored links. The latter studies conclude that high quality ads have a negative impact on the performance of listings at succeeding positions due to information satiation. In this paper, we study the impact of sponsored listings on performance of other organic listings, and how top sponsored listings perform compared to the top organic listings.

We contribute to this literature by studying the usefulness of sponsored listings from the platform's perspective, i.e., how they help in increasing a platform's overall revenue. Earlier papers that studied optimal auction designs for Internet search platforms (Edelman et al., 2007; Athey and Ellison, 2011; Gomes, 2014) did not consider the effect of sponsored results on the search engines' other links because, unlike online marketplaces, Internet search engines typically do not earn any revenue or commission from the organic listings. Choi and Mela (2016) use a structural model of search on an online marketplace to find the effect of sponsored listings on the platform's and sellers' revenue, and the optimal ad pricing mechanism for the platform.

A great deal of advertising may indirectly signal quality if there exist market mechanisms that produce a positive relationship between product quality and advertising expenditures (Nelson 1974). Especially in markets where ex-ante product uncertainty is high, a separating equilibrium is achieved when the gain from the repeat purchase relative to the cost of advertising is higher for high quality firms at optimal prices, and lower quality firms do not gain from mimicking the strategies of high quality firms (Milgrom and Roberts, 1986; Kihlstrom and Riordan, 1984). Although a number of papers (Caves and Greene, 1996; Ackerberg, 2003; etc.) have empirically examined the signaling effect of traditional TV or billboard advertising, there is very little work that has tested the theory in the context of online sponsored search where advertising costs can be variable, and advertised and unadvertised products appear together (Feng and Xie, 2012). Animesh et al. (2010) show the existence of low-quality advertisers on Internet search engines in categories where ex-ante quality uncertainty is high. In their setting, the assumption is that the true quality of the advertiser is known to the platform, but not known to the consumers. This is why platform can resolve the adverse selection problem by doing quality-score adjustment. Using a randomized field experiment, Sahni and Nair (2016) find that users have a positive perception towards sponsored listings in the context of restaurant-search platform. More specifically, a restaurant has a higher conversion rate when it is identified as an ad. Some recent work (Lewis, 2011; Belleflamme et al., 2014) has suggested that platforms are able to solve the adverse selection problems by reducing the information asymmetry between the sellers and consumers by providing a wide variety of information to the consumers. Extensive use of reinforcement learning, i.e., exploration and exploitation of listings, by Internet platforms further suggests that they heavily depend on consumers to learn about the true relevance of a listing (Banerjee et. al., 2014).

In this paper, we discuss how sponsored listings can work as a screening mechanism for an online marketplace to identify good products when private information regarding the true relevance of a product lies with the sellers (Liu and Viswanathan, 2014). In this way, we focus on the information asymmetry between the platforms and the sellers, and not between the consumers and the sellers. We then find evidence for this mechanism using our randomized field experiment data from product categories exhibiting different levels of information asymmetry. To the best of our knowledge, we are the first to compare the performance of sponsored listings to that of the non-sponsored (organic) listings at the same position, and also across different product categories. This is helpful in understanding whether sponsored listings indeed dilute the search quality for the users.

In summary, we combine the above two streams of literature to causally estimate the effect of sponsored listings on search relevance or quality of online marketplaces. Due to our rich e-commerce marketplace data, we are also able to see how this effect differs between different product categories, and provide evidence for the plausible mechanisms driving the results.

3 Theoretical Model

In this section we present a simple analytical model to provide the intuition for the empirical analysis presented later in the paper. First, we discuss the manner in which a marketplace displays organic listings in search results. We then discuss the effect of serving sponsored listings in search results.

An online marketplace comprises of several third-party sellers that sell products to consumers. When a consumer searches for a product, e.g., “digital camera”, the platform identifies the relevant products, predicts their performance using a machine learning approach and arranges them in a list of search results ordered by the predicted performance. The platform estimates the relevance of the listings based on historical performance data. However, several new and existing sellers introduce new products every day in the marketplace. Under such circumstances, it is extremely difficult for the platform to correctly evaluate every listing and match them to a user's search query. This creates the problem of information asymmetry between the platform and the sellers that reduces the accuracy of the relevance algorithm. The platform may fail to identify the good listings especially when a listing's relevance is very subjective and cannot be easily quantified. Typically, this happens for product categories with fewer vertically (and more horizontally) differentiated or digital attributes. Lack of such parameters makes it difficult for the platform to use the available data to learn about relatively new listings and estimate a listing's true quality.

A marketplace may decide to sell some of its positions in search results as sponsored slots to the sellers to promote their listings in exchange of an additional cost per click. In our empirical context, this cost is set by the platform, and is the same for any position. Given this price, the sellers make the advertising decision, which creates the sponsored pool of listings. The marketplace then ranks the sponsored pool of listings in the sponsored slots and non-sponsored pool of listings in the organic slots in the order of relevance. If the number of listings available for sponsored slots is lower than the number of sponsored slots, then some sponsored slots are not utilized and used for organic listings. Therefore, no listing appears twice in the search results, i.e., they only appear as organic or as sponsored. We develop a simplified model below to demonstrate the information asymmetry that platform faces, and its influence on the effect of introducing sponsored listings.

Consider a platform that ranks three products listings $\{H, L_1, L_2\}$ in the search results for a given search query. This ranking is based on the deterministic utility a listing provides to a consumer, as measured by the platform's ranking algorithm. This deterministic utility measures the average relevance of the listing with respect to the search query, and should be seen as measuring both the fit and the quality of the product listing. Let V_H denote the predicted deterministic utility of listing H, and V_L denote the predicted deterministic utility of listings L_1 and L_2, where $1 > V_H > V_L > 0$. Thus, we have two types of listings. A representative consumer's value of clicking on listing $i \in \{H, L\}$ is given by $u_i = V_i + \epsilon_i$, where ϵ_i is an idiosyncratic shock that follows i.i.d. $U(-1, 1)$. Let her utility from the outside option be normalized to zero. The consumer perfectly learns the realizations of u_i after viewing (evaluating) the listing. Note that the consumer faces no information asymmetry in our model, i.e., once she views a listing, she is able to correctly identify its relevance.[8]

The consumer's search behavior is characterized as following. 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 a search cost, $s > 0$, to evaluate the listing at the second position. We assume that her expected utility from clicking on the next listing is given by V_L if the first listing is H and by $V = \frac{V_H + V_L}{2}$ if the first listing is L_1 or L_2, as there is only one high-type and two low-type listings. She views the second listing with a probability such that the expected marginal benefit of evaluating the second listing is greater than the marginal search cost. Let δ^i denote the probability of evaluating the second listing if the first listing is $i \in \{H, L\}$. Then, $\delta^H = Pr(V_L - u_H > s)$ and $\delta^L = Pr(V - u_L > s)$, where $\delta'(s) < 0$ and $\delta^L > \delta^H$. We assume that the consumer's marginal search cost of evaluating the listing at the

[8] The assumption is motivated by previous research (Lewis, 2011; Belleflamme et al., 2014) and also the industry practice of experimenting with users to learn about the relevance of new products (Banerjee et. al., 2014).

third position is so high that she never views the third listing. After evaluating the listings, she makes the click and purchase decisions. Let $\theta_i = Pr(u_i > 0)$ denote the probability of click on listing i 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 $\theta_{HL} > \theta_{LL} > \theta_{LH}$.[9] 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. Then we can write down the expressions for probability of conversion of the three listings as follows.

$\pi_1^{HL} = \gamma_H \beta_1^{HL} = \gamma_H((1 - \delta^H)\theta_H + \delta^H \theta_{HL})$, where γ_i is the probability of conversion per click of listing $i \in \{H, L\}$.[10] Similarly,

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

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

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

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

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

We assume $0 < \gamma_L < \gamma_H < 1$ such that $\pi_1^{HL} > \pi_1^{LH}$ and $\pi_2^{HL} < \pi_2^{LH}$.

Information Asymmetry: When there exists information asymmetry (IA) between the platform and the sellers regarding the product relevance, platform's predicted deterministic utility could be different from a listing's true deterministic utility. This creates an error in the rank order of the listings. Let us assume the marketplace ranks H at the first position with probability $\frac{1}{3} \leq \kappa \leq 1$, and the other two listings with probability $\frac{1-\kappa}{2}$. H gets ranked at the second and third position with an equal probability $\frac{1-\kappa}{2}$. $\kappa = 1$ implies there is no information asymmetry while $\kappa = \frac{1}{3}$ implies maximum information asymmetry. Hence, κ can be seen as the knowledge parameter of the marketplace.

Suppose that the marketplace decides to sell the first position (as a sponsored slot) to

[9] Expressions of δ and θ are derived in the Appendix.

[10] For simplicity, we assume that the preceding or the succeeding listing only affects the click-through rate, but not the conversion post click.

the sellers to promote a listing in exchange of an additional cost per click CPC that is set by the platform. Given the CPC, the sellers make the advertising decision. If multiple sellers decide to pay the CPC, the marketplace chooses the listing which it believes to be the most relevant, otherwise the first position is offered to the only bidding seller. Assuming the profit margin to be equal to 1 for both types of listings, we can then write down the valuation of advertising per click for the two types of listings as follows.[11]

$$v_H = \gamma_H \Big(1 - \frac{(1+\kappa)}{2} \frac{\beta_2^{LH}}{\beta_1^{HL}}\Big), \tag{1}$$

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

In the following subsections, we analyze which listing has a higher incentive to advertise when there is no information asymmetry and when there exists some degree of information asymmetry.

3.1 When $\kappa = 1$ (No Information Asymmetry)

When there is no discrepancy between the platform's ranking order and the true ranking order, H gets ranked at the first position with probability 1, and $\{L_1, L_2\}$ get ranked at the second and the third position with probability $\frac{1}{2}$. Due to the introduction of the sponsored slot, there could be the following possible ranking outcomes- $\{H, L_i, L_j\}$ if H gets the sponsored slot, or $\{L_i, H, L_j\}$ if an L gets the sponsored slot, where $i \neq j$.

Proposition 1 *(i) When $\kappa = 1$, we have $v_L > v_H$ if*

$$\frac{1 - \frac{\beta_2^{HL}}{\beta_1^{LH}}}{1 - \frac{\beta_2^{LH}}{\beta_1^{HL}}} > \frac{\gamma_H}{\gamma_L}. \tag{3}$$

(ii) A low type listing gets the sponsored slot when $v_L > CPC > v_H$.

[11] See Appendix for the derivation, and also for the proofs of the upcoming propositions and corollaries.

The inequality (3) means that the ratio of percentage change in the probability of click from the first to the second position of the low type to the high type is larger than the ratio of the conversion per click rate of the high type to the low type.[12] Therefore, the result implies that if the position effects are stronger for the low type, then the low type listing has a higher value for advertising than the high type.[13] In other words, when search costs are high enough, the sponsored slot gives an opportunity to the low type to occupy the first position, and get much higher exposure and profits than without the sponsored slot.

3.2 When $\kappa < 1$ (With Information Asymmetry)

When there exists some degree of information asymmetry, $\{L_i, L_j, H\}$ is also a possible ranking outcome when L occupies the sponsored slot, where $i \neq j$. We first look at how κ affects the incentives of the two types of listings.

Lemma 1 $\frac{dv_i}{d\kappa} < 0 \; \forall \; i \in \{H, L\}$

Lemma 1 says that both types have a higher valuation for advertising when the marketplace faces IA. This happens for the high type because as κ decreases, 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, two low types can occupy the first two positions more frequently, which increases the returns to the first position because of lower competition. The result implies that the platform can charge a higher CPC for the sponsored slot when there is asymmetric information. If the CPC remains the same as in the no information asymmetry case, then both types bid for the sponsored slot (pooling equilibrium). This also implies that

[12] In our parsimonious framework, both types of listings consider an increase in rankings from the same (second) position while making the advertising decision. However, in reality, a low type listing is ranked much lower than a high type listing when there is no IA. So the gain in number of clicks by getting a top slot for a low type is typically much higher than what we have in our simple model.

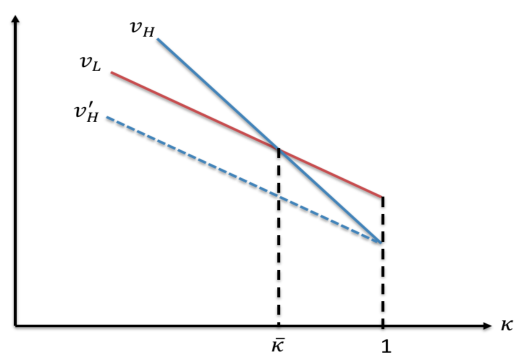
[13] Although we exclude the case where same listing appears twice, e.g., $\{H, H, L\}$, the results qualitatively remain the same in that setting as well. This is because the marginal gain for H by removing competition from L at the second position can still be lower than the marginal gain for L by appearing at the first position from the third position.

there exists a CPC such that the high type listing is advertised when there is information asymmetry, while no listing is advertised when there is no information asymmetry.

Proposition 2 *(i) Suppose inequality (3) holds such that $v_L > v_H$ when $\kappa = 1$. Then, we have $v_H > v_L$ when $\kappa < \bar{\kappa} < 1$ and*

$$\frac{\gamma_H \beta_2^{LH}}{\beta_1^{HL}} - \frac{\gamma_L \beta_2^{HL}}{\beta_1^{LH}} > max\{\frac{\gamma_L(\beta_2^{LL} - \beta_2^{HL})}{\beta_1^{LL}}, \gamma_H - \gamma_L\}. \tag{4}$$

(ii) The high type listing always gets the sponsored slot when $v_H > CPC > v_L$.



Value of Advertising w.r.t. κ

From Lemma 1 we know that, as κ decreases (i.e., IA increases), sellers of both types of listings value the sponsored slot more. When the rate of increase in v_H is greater than the rate of increase in v_L due to a decrease in κ (i.e., when inequality (4) holds), then the high type can have a higher incentive to advertise when there is a great deal of information asymmetry, i.e. when κ is low enough. The above figure depicts such a scenario. v_H and v'_H denote the value of advertising for the high type in a case where inequality (4) holds and does not hold, respectively. The greater the relative returns to the second position for the high type, the bigger the loss due to demotion to the third position (i.e., due to lower κ). This results in a higher rate of increase in the incentive to advertise for H due to a decrease in κ when $\frac{\beta_2^{LH}}{\beta_1^{HL}}$ is relatively large.

The above proposition implies that, when position effects are strong enough, sponsored products work as a screening mechanism for the platform to identify high-relevance listings when there is a high level of asymmetric information between the platform and the sellers regarding the true relevance of a listing. The intuition behind the result is that when there is no information asymmetry, the high type is always at a relatively higher position, i.e., it never gets ranked at the third position. Therefore, it does not have as much incentive as a low-ranked low type to acquire the sponsored slot. On the other hand, when there is information asymmetry, the high type is sometimes ranked low in the search results. Under such a scenario, the sponsored product listing allows the seller of the high type to increase its ranking as their valuation for advertising can be greater than the low-ranked low type.

Based on the above two propositions, we derive the following corollaries regarding the effect of introducing sponsored listing (i) at the ad slot, (ii) at the second position, and (iii) on the two positions together. We will assume that a low-relevance (high-relevance) listing occupies the sponsored slot when information asymmetry is low (high).

Corollary 1 *When the marketplace faces no information asymmetry, the introduction of the sponsored listing reduces the click and conversion performance at the sponsored (first) slot, and vice versa.*

From Proposition 2, we know that the presence of information asymmetry creates a bigger incentive for the sellers of high-relevance listings to advertise than for those selling low-relevance listings. Therefore, a low-type listing replaces the high-type listing at the sponsored slot when there is no information asymmetry, whereas a high-type listing occupies the first position more often with the sponsored slot when there exists a great deal of information asymmetry.

Corollary 2 *When the marketplace faces no information asymmetry, the introduction of the sponsored listing increases the click and conversion performance at the succeeding (second) position, and vice versa.*

When a low-relevance listing gets the sponsored slot, as in the case of no information asymmetry, it has a positive effect on the organic listing at the succeeding position in two ways. First, the high-type listing replaces a low-type listing at the second position. Second, it increases the number of views and clicks the listing at the second position gets due to lack of satiation. Whereas in the case where high-relevance listing occupies the sponsored slot more frequently, as in the case with information asymmetry, the lower ranked listing gets lower clicks and conversion. This is caused by the competitive effects of preceding listings when there is satiation due to high-quality preceding listing.

Corollary 3 *(i) When the marketplace faces low information asymmetry, the introduction of the sponsored listing decreases the overall click and conversion performance at two positions together, but the magnitude of the effect decreases as s decreases.*

(ii) When the marketplace faces high information asymmetry, the introduction of the sponsored listing increases the overall click and conversion performance at two positions together, but the magnitude of the effect decreases as s decreases.

Based on the first two corollaries, we arrive at the corollary regarding the overall effect of sponsored listings on the search performance. When there is no information asymmetry, the sponsored listing can hurt the search performance due to the selection of low-relevance listings at a top position. However, if the consumers have low marginal search costs at the ad slot, then they continue with their search and purchase at a higher rate (compared to the case with no ads) from the next position. Conversely, when high-relevance listings occupy the sponsored slot, as it is likely to happen in the case of asymmetric information, then search performance improves due to ads. But it can also lower the demand from other succeeding positions, thus nullifying some of the positive effect at the sponsored slot.

4 Empirical Setting

4.1 Background

In this research, we collaborate with *Flipkart*, the largest e-commerce marketplace in India with a 43% market share. Flipkart users can access the platform via both a website and a mobile application (hereinafter app). When considering only m-commerce, which accounts for 70% of India's e-commerce traffic, Flipkart's market share increases to 63%.[14] As the majority of e-commerce activity in India happens through mobile devices, we use data from their 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, either by typing in a query or by using the search menu. Given a query by a user, a list of listings (search results) 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 down. Depending on the size of the screen, a user can view 2-4 listings at a time on the search results page.

We define a *search* as the list of listings that get viewed, and each *impression* as the view of a listing by a user during a search. A new search is generated either when a user searches for a new product or when she adds a certain filter to the current search. The maximum position viewed during a search determines its *depth*, i.e., number of impressions per search. 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 check out more details, such as more images, seller details, reviews, etc. On this page, a user can either *convert* by

[14] http://www.business-standard.com/article/companies/flipkart-controls-63-of-app-based-e-commerce-traffic-116012000614_1.html

tapping on “Buy Now”, “Add to Cart” or “Add to Wishlist”, or go back to the search result page.

On the supply side, Flipkart charges sellers a flat price per click on sponsored listings. The price is set at the product vertical level, and is, therefore, the same for every position. Given the price, sellers can choose the listings that they want to advertise, total budget, and the campaign duration. Flipkart picks the most relevant listings from the pool of listings chosen for advertisement, and also makes sure that they do not show irrelevant ads, i.e., ads are served to a user only if the relevance of listings chosen for an advertisement crosses a certain threshold. Typically, a user is less likely to see an ad as her search query becomes more specific. As most of the queries in the clothing category are generic, we see a higher *fill-rate*[15] in clothing-related searches compared to that in the electronics category. The users can, therefore, see different numbers of ads or no ad in some of the searches. However, one cannot use this variation to identify the effect of ads on user behavior, because this variation is endogenous, due to either demand-side (e.g., different search queries) or supply-side (e.g., lack of advertisers, pricing, etc.) reasons. To create exogenous variation in the number of ads shown, we use data from an experiment described in the next subsection.

4.2 Experiment Design

Our identification strategy relies on impression-level data from a randomized field experiment during February 2016 with around 2 million mobile app users. In this experiment, four random groups of users were served different numbers of ads in the following sets of positions (see Figure 3 for a snapshot). Neither the sellers nor the users were aware of this experiment.

Bucket 1: 8, 12, 18, 22, 28 ... **Bucket 2**: 2, 6, 10, 14, 18, 22 ...

Bucket 3: 3, 7, 13, 17, 23, 27 ... **Bucket 4**: 4, 14, 20, 24, 30 ...

[15] $\text{fill-rate} = \frac{\text{number of search result pages where ads are served}}{\text{total number of search result pages}}$

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

Figure 3: Experiment Design at the top 10 positions

The experiment design allows us to randomly observe the sponsored listings and the displaced organic listings at the same position across searches. Thus we can identify the difference in the performance of ads and the organic listings that would have appeared at the same position had ads been absent. We can also leverage the exogenous variation in the proportion of ads users see to find the effect of serving sponsored listings on the probability of click and conversion from a search. As Bucket 1 users do do not get served ads at the top 7 positions, we treat the searches under this condition as our control group for answering our research questions. Each of the other three buckets becomes our treatment group.

To empirically analyze how information asymmetry mediates the impact of sponsored listings, we identify two product categories in which platforms exhibit very different levels of information asymmetry. It is typically easier to evaluate the quality or relevance of products with more vertically differentiated attributes (say, digital attributes). So we focus on clothing and electronics that have very low and high numbers of digital attributes, respectively. As it is easier to evaluate electronic goods than clothing goods based on product attributes, we assume that the level of information asymmetry is lower in this product category. The clothing products have implicit attributes that the platform finds hard to infer. Another source of information asymmetry in the clothing category is the huge number of products available, compared to the electronics category. As clothing products are more horizontally differentiated than electronics products, the number of products available in the clothing category usually far exceeds the number in 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 almost 20 times the size of electronics-related products.

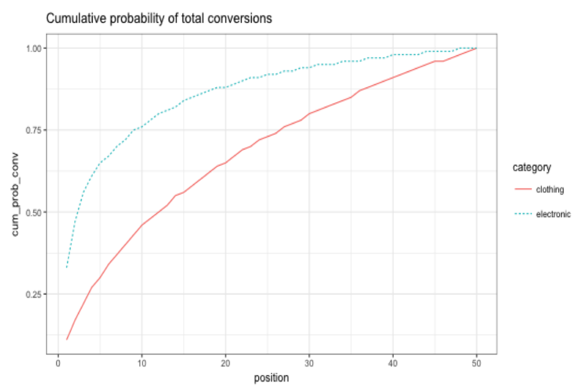


Figure 4: Cumulative Probability of Total Conversions till Position 50

In Figure 4, we plot the cumulative distribution of conversions across positions in the two categories, given generic queries like "white tops", "blue shirt", "women jeans", "laptop", "android phone", "led tv", etc. We find that 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 until 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 true relevance of electronic products.

4.3 What We Expect to See

Based on the theoretical results of section 3, and the description of our empirical setting and the experiment design, we expect to observe following results in our empirical analysis.

First, based on Corollary 1, we hypothesize that in a product category like electronics, where the level of asymmetric information is low, sponsored listings would be of lower relevance than the organic listings they displace. Specifically, the electronics ads are expected to be clicked and converted at a lower rate than the organic listings they replace. On the other hand, in a category like clothing, where the marketplace has relatively little information regarding the product's relevance, higher relevance listings should appear as sponsored listings, leading to higher click-through and conversion rates for ads.

Second, based on Corollary 2, as a result of the previous hypothesis, the click and conversion performance of organic listings appearing next to the sponsored listing would increase in the case of electronics, and decrease in the case of clothing-related searches. However, there could also be a case that relatively high-relevance listings are being displaced in clothing. In that case, we do not expect to see a strong positive effect on the click or conversion at the sponsored slot or, therefore, a negative effect on the organic listings at succeeding positions in clothing.

Finally, based on Corollary 3, as a weighted effect of the two individual effects, we expect a small but negative effect on the clicks and conversion at the search level in the case of electronics, and a small but positive effect on the search clicks and conversion in clothing. As performance metrics, especially conversion rate, typically have a large coefficient of variation, we would require hundreds of thousands of observations to detect small effects with enough power. Therefore, although we might be able to detect the effect on clicks (which are observed more frequently), detecting effects of similar order on conversion metrics might be difficult. We provide the details of the power analyses in section 5 wherever applicable.

4.4 Data

As mentioned earlier, we focus on searches related to two very distinct product categories in terms of number of digital attributes- clothing and electronics. While electronics includes verticals like mobile phone, laptop, television, air conditioner, headphones, speakers, clothing verticals include shirt, top, jeans, t-shirt, fabric, etc. These two categories account for about

80% of Flipkart's business. We exclude searches where ads are not served, 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 reason. There are over 600 thousand unique products listed by around 17000 sellers. Almost 95% of these listings belong to the clothing category. This tells how big the assortment is in the clothing category, compared to the electronics category. The proportion of sellers who advertise in the clothing category is also higher than the fraction of sellers who advertise in the electronics category.

We first try to understand how users search in the presence of ads. More specifically, we determine how the number of products they view changes when they see different numbers of ads in the search results. Table 1 below provides the summary statistics of search depth for clothing and electronics across the different experimental buckets. There are two key observations to be made from Table 1. First, the mean depth of a search is higher in clothing than in electronics. In other words, on average, users view more listings in clothing than in electronics. Second, the distribution of the depth is similar across the four buckets in both the categories, with buckets with more ads having only 1.5-5% higher search depth on average.

Table 1: Distribution of Search Depth

Category	Bucket	Searches	Mean	Std Dev.	Median
Clothing	1	149356	18.9	14.0	16
Clothing	2	149951	19.7	14.4	16
Clothing	3	147665	19.2	14.1	16
Clothing	4	151296	19.6	14.5	16
Electronics	1	187278	10.9	11.4	6
Electronics	2	190115	11.4	11.7	6
Electronics	3	184226	11.1	11.4	6
Electronics	4	178078	11.5	11.8	7

Table 2: Click and Conversion per Impression

Category	Impressions	Click Mean	Click SD	Conversion Mean	Conversion SD
Clothing	23737101	0.0257	0.158	0.00193	0.044
Electronics	10635407	0.0431	0.203	0.00302	0.055

Table 3: Click and Conversion per Search

Category	Searches	Click Mean	Click SD	Conversion Mean	Conversion SD
Clothing	598268	0.49	0.99	0.041	0.22
Electronics	739697	0.53	0.83	0.038	0.20

Table 4: Average Proportion of Ads Impressed per Bucket

Bucket	Overall	Clothing	Electronics
1	0.085	0.113	0.062
2	0.270	0.260	0.277
3	0.201	0.209	0.195
4	0.125	0.152	0.103

Tables 2 and 3 provide the summary statistics regarding click and conversion per impression and per search, respectively. In clothing, the mean click-through rate (CTR) is around 2.5%, and the mean conversion rate is 0.2%. In electronics, both the variables are higher on average (see Table 2). Table 3 shows that the mean probability of click and conversion per search is around 50% and 4%, respectively. Note that conversion has a very high standard deviation; the coefficient of variation is of the order 20 at the impression level and of 6 at the search level. Thus, we require hundreds of thousands of observations to be able to detect effects of even high economic significance with enough power in our tests.

We now look at the level of ads received by the users of different buckets. Table 4 provides a summary of the proportion of ads served to the users in different buckets. One can see that the difference between the smallest and the largest treatment is close to 18.5% points overall, while it is smaller (15% points) for clothing and bigger (21% points) for electronics due to the difference in search depth distribution across the two categories.

Table 5 presents the summary statistics for listing characteristics like rating, number of ratings, price and discount of sponsored and organic listings from the top 8 positions. We can see that ads in both clothing and electronics generally have lower number of ratings than their organic counterparts, hinting that sponsored products are relatively new. The difference in the average rating of sponsored and organic listings in clothing is higher than the difference between the two in electronics. Also, the price of sponsored clothing products is higher than the price of the organic products on average. Overall, the characteristics of sponsored listings seem to be different from that of organic listings in both the categories.

Table 5: Sponsored vs. Organic Characteristics

Category	Source	Rating		Rating Count		Price (Rs.)		Discount(%)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Clothing	Sponsored	3.7	1.07	5	17	938	967	57	18.2
Clothing	Organic	3.2	1.00	24	60	878	1601	60	18.7
Electronics	Sponsored	3.7	0.60	137	650	13452	18497	14	23.7
Electronics	Organic	3.8	0.59	2565	4329	13540	14914	11	23.3

We also do a randomization check for users by looking at their click and conversion behavior two weeks before the experiment to make sure they were randomly assigned to each condition (see Table 20 in the Appendix). We also do a randomization check to see whether similar sets of sponsored listings were served to the different groups of users (see Table 21 in the Appendix).

5 Empirical Analysis

In this section, we analyze the experimental data to test the theoretical predictions we generate in section 3 regarding the relevance of sponsored listings, and its the effect on the succeeding positions, and on the probability of click and conversion in a search.

5.1 The Effect of Ads at the Sponsored Slot

We commence our empirical analysis by testing our first corollary that sponsored listings in categories like clothing, where the marketplace faces a high degree of information asymmetry, are of relatively higher relevance than the sponsored listings in categories like electronics, where the marketplace faces a low degree of information asymmetry. In order to determine this, we see how sponsored listings perform compared to the organic listings that would have appeared at the same position had there been no ads. Let $\hat{\alpha_c}$ and $\hat{\alpha_e}$ denote the the estimated difference between the performance of the sponsored and the organic listing in clothing and electronics, respectively. If $\hat{\alpha_e} < \hat{\alpha_c}$, then that would support our hypothesis.

To estimate this effect, we split the sample into impressions from different positions, and then run a separate regression for three positions {2,3,4} among the top 7 positions where an ad is served. In all of our tests, Bucket 1 forms the control group, where no ad is served at the top 4 positions.

Table 6: Difference in CTR at Sponsored Slots in Electronics

Position	B2-B1	B3-B1	B4-B1
2	$-.0174^{***}$ [-21%]		
3		$-.0026^{***}$ [-5%]	
4			$-.0021^{***}$ [-5%]
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 7: Difference in CTR at Sponsored Slots in Clothing

Position	B2-B1	B3-B1	B4-B1
2	$.0079^{***}$ [19%]		
3		$.0076^{***}$ [25%]	
4			$.0089^{***}$ [28%]
Note:			*p<0.1; **p<0.05; ***p<0.01

Tables 6 and 7 present the results for the mean click-through rate in electronics and clothing, respectively, while Tables 8 and 9 present the difference in the mean conversion rates for the two categories. We find statistically significant negative effects in the case of

Table 8: Difference in Conversion at Sponsored Slots in Electronics

Position	B2-B1	B3-B1	B4-B1
2	$-.0024^{***}$ [-37%]		
3		$-.0010^{***}$ [-25%]	
4			$-.0006^{***}$ [-24%]
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 9: Difference in Conversion at Sponsored Slots in Clothing

Position	B2-B1	B3-B1	B4-B1
2	.0004* [11%]		
3		.0003	
4			.0003
Note:	*p<0.1; **p<0.05; ***p<0.01		

electronics, but not in the case of clothing, where the effects are either positive or statistically insignificant. More specifically, the ads in electronics searches get clicked and converted with 5-21% and 24-35% lower probability, respectively, compared to the organic listings they displace. On the other hand, the click-through rate of ads in the clothing category is (19-28%) higher at all positions than that of the organic listings they substitute, while the conversion rate is either higher (at 10% significance level in Bucket 2) than or at par with organic listings they displace.[16] Thus we find strong empirical support for our first hypothesis.

We also look into how the difference in the characteristics of sponsored listings and organic listings varies across the two product categories. Table 23 in the appendix presents the comparison for sponsored and organic listings at positions {2,3,4}. The interaction term with electronics gives the difference in the two categories for the differences in sponsored and organic listings. We observe that sponsored listings in clothing, compared to the organic listings, have a higher rating and price than what sponsored listings have compared to the organic listings in electronics. If we assume rating and price to be credible signals of quality, then sponsored listings indeed seem to select good quality listings in clothing, even in terms of characteristics.

We further do a regression analysis by using a binary logit model to represent the click

[16]Power analysis, which can be found in the appendix (Table 22), reveals that we do not have sufficient power ($< 50\%$) to detect 10% effects in conversion probability, given our sample size. However, we have 100% power in our tests done in Table 7 for CTR.

or the conversion probability on being impressed at any fixed position.

$$Pr(Click_i) = \frac{exp(\bar{U}_i^{Click})}{1 + exp(\bar{U}_i^{Click})} \tag{5}$$

$$Pr(Conversion_s) = \frac{exp(\bar{U}_i^{Conv})}{1 + exp(\bar{U}_i^{Conv})} \tag{6}$$

where U_i is the latent utility of a user on converting or clicking, and can be expressed as following.

$$U_i^{Click} = \underbrace{b_0 + b_1 Sponsored_i}_{\bar{U}_i^{Click}} + \epsilon_i^{Click} \tag{7}$$

$$U_i^{Conv} = \underbrace{\beta_0 + \beta_1 Sponsored_i}_{\bar{U}_i^{Conv}} + \epsilon_i^{Conv} \tag{8}$$

where $Sponsored_i$ identifies the bucket where ad is served at the position in consideration, ϵ represents the listing characteristics or features unobservable to the researcher, and follows i.i.d. extreme value distribution. Here, b_1 (β_1) identifies the difference in the click-through (conversion) rates of sponsored and the corresponding organic listings; positive and statistically significant $\hat{b_1}$ ($\hat{\beta_1}$) implies that ads perform better than the organic listings, and vice versa. Tables with the regression results are provided in the Appendix (see Tables 24-27).

5.2 The Effect of Ads at Succeeding Positions

We next look at how the presence of sponsored listings affects the performance of the organic listings at the succeeding positions. According to our second corollary, we expect this effect to be positive in searches where ads perform much worse than the displaced organic listings, such as in electronics. In order to estimate the effect, we compare the click-through and the conversion rates of organic listings at the succeeding position of sponsored slots with the click-through and conversion rates of organic listings at the same position but with no preceding sponsored listing.

Tables 10 and 11 present the results for the click-through rate in electronics and clothing,

Table 10: Difference in CTR at Next Position in Electronics

Position	B2-B1	B3-B1	B4-B1
3	.0147*** [29%]		
4		.0034*** [8%]	
5			.0032*** [9%]
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 11: Difference in CTR at Next Position in Clothing

Position	B2-B1	B3-B1	B4-B1
3	.0018*** [6%]		
4		.0007	
5			0.0009
Note:	*p<0.1; **p<0.05; ***p<0.01		

respectively, while tables 12 and 13 present the conversion rate for the two categories. Again, we find strong support for our hypothesis. We see that the click-through rate of organic listings appearing next to ads in electronics searches are higher by 8-29% and the conversion rate is higher by 20-37% in Buckets 2 and 3. Whereas there is no effect on click or conversion of the succeeding position in clothing-related searches for most of the buckets (the difference is not statistically significantly different from 0), thus we do not observe any negative effect on the succeeding position after relatively high-quality ads in clothing. The differences we observe happens because of two reasons. First, the organic listings succeeding the ad could be of higher relevance than the ones in the control group, as their original rank is higher. Secondly, some of the lost demand at the sponsored slots in electronics-related searches gets carried over to the succeeding listings. In other words, poor sponsored listings improve the performance of the listing at succeeding positions due to low competition (lack of satiation). The result is also consistent with the one previously found in Jeziorski and Segal (2015) where it is conversely shown that the presence of higher quality ads negatively affects the

Table 12: Difference in Conversion at Next Position in Electronics

Position	B2-B1	B3-B1	B4-B1
3	.0015*** [37%]		
4		.0005*** [20%]	
5			.0001
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 13: Difference in Conversion at Next Position in Clothing

Position	B2-B1	B3-B1	B4-B1
3	.0001		
4		−.00009	
5			.0004
Note:			*p<0.1; **p<0.05; ***p<0.01

click-through rate of succeeding ads due to satiation.

We further conduct a regression analysis to find the average effect of ads on the performance of organic listings at multiple positions preceded by a sponsored listing. We run the same regression as in the previous section, but now $Sponsored_i$ identifies whether an ad is served before position j in search i. Specifically, we consider three succeeding positions– $\{3,4,5\}$ for Bucket 2, $\{4,5,6\}$ for Bucket 3, and $\{5,6,7\}$ for Bucket 4. Again, Bucket 1 forms the control group, where no ad is served between positions 1 and 7. We also control for the position of the impression to compare the performance with or without a preceding ad at each position, and not across positions. We use a binary logit model to represent the click or conversion probability, on being viewed at position j during search i.

$$Pr(Click_{ij}) = \frac{exp(\bar{U}_{ij}^{Click})}{1 + exp(\bar{U}_{ij}^{Click})} \tag{9}$$

$$Pr(Conversion_{ij}) = \frac{exp(\bar{U}_{ij}^{Conv})}{1 + exp(\bar{U}_{ij}^{Conv})} \tag{10}$$

where U_{ij} is the latent utility of a user on clicking or converting at position j during search i, and can be expressed as following.

$$U_{ij}^{Click} = \underbrace{b_0 + b_1 Sponsored_i + b_2 Position_j}_{\bar{U}_{ij}^{Click}} + \epsilon_{ij}^{Click} \tag{11}$$

$$U_{ij}^{Conv} = \underbrace{\beta_0 + \beta_1 Sponsored_i + \beta_2 Position_j}_{\bar{U}_{ij}^{Conv}} + \epsilon_{ij}^{Conv} \tag{12}$$

where ϵ represents the listing characteristics or features unobservable to the researcher, and follows i.i.d. extreme value distribution. Here, b_1 (β_1) identifies the average effect on click-through (conversion) rate of organic listings across the three succeeding positions; positive $\hat{b_1}$ ($\hat{\beta}_1$) implies that the presence of ads increases the click-through (conversion) rate of organic listings at succeeding positions. The following tables (14-17) provide coefficient estimates of the logistic regression. All the results for the average effect on multiple succeeding positions are consistent with out hypothesis.

Table 14: Effect of Ads on CTR at Succeeding Positions in Electronics

	Dependent variable:		
	Click		
	{3,4,5}	{4,5,6}	{5,6,7}
	(1)	(2)	(3)
Preceded by Ad	0.193***	0.058***	−0.014
	(0.010)	(0.012)	(0.014)
Position	−0.247***	−0.143***	−0.030***
	(0.006)	(0.008)	(0.008)
Constant	−2.160***	−2.574***	−3.148***
	(0.025)	(0.038)	(0.051)
Observations	890,031	732,382	647,020
Log Likelihood	−169,993.600	−117,631.600	−96,912.620
Akaike Inf. Crit.	339,993.200	235,269.200	193,831.200
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 15: Effect of Ads on CTR at Succeeding Positions in Clothing

	Dependent variable:		
	Click		
	{3,4,5}	{4,5,6}	{5,6,7}
	(1)	(2)	(3)
Preceded by Ad	0.033***	0.013	0.012
	(0.011)	(0.011)	(0.012)
Position	−0.059***	−0.073***	−0.059***
	(0.007)	(0.007)	(0.007)
Constant	−3.241***	−3.135***	−3.238***
	(0.027)	(0.034)	(0.044)
Observations	1,191,117	1,137,005	1,069,354
Log Likelihood	−162,759.200	−151,658.700	−133,041.800
Akaike Inf. Crit.	325,524.300	303,323.400	266,089.700
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 16: Effect of Ads on Conversion at Succeeding Positions in Electronics

	Dependent variable:		
	Conversion		
	{3,4,5}	{4,5,6}	{5,6,7}
	(1)	(2)	(3)
Preceded by Ad	0.258***	0.156***	−0.004
	(0.036)	(0.045)	(0.051)
Position	−0.324***	−0.108***	−0.083***
	(0.023)	(0.028)	(0.032)
Constant	−4.538***	−5.473***	−5.559***
	(0.089)	(0.139)	(0.190)
Observations	890,031	732,382	647,020
Log Likelihood	−21,297.620	−13,600.950	−10,716.340
Akaike Inf. Crit.	42,601.250	27,207.900	21,438.690
Note:			*p<0.1; **p<0.05; ***p<0.01

5.3 The Effect of Ads on the Overall Search Performance

In this section, we test our third corollary. As ads are of very low relevance in electronics and of high relevance in clothing, we expect the overall effect of ads at the search level to be negative for electronics and positive for clothing, especially with respect to the click performance, as we might not have enough observations to detect a small effect on conversion

Table 17: Effect of Ads on Conversion at Succeeding Positions in Clothing

	Dependent variable:		
	Conversion		
	{3,4,5}	{4,5,6}	{5,6,7}
	(1)	(2)	(3)
Preceded by Ad	0.025	−0.008	0.083**
	(0.037)	(0.039)	(0.042)
Position	−0.114***	−0.132***	−0.044*
	(0.023)	(0.024)	(0.026)
Constant	−5.555***	−5.432***	−5.924***
	(0.093)	(0.120)	(0.156)
Observations	1,191,117	1,137,005	1,069,354
Log Likelihood	−20,809.960	−18,343.030	−16,359.060
Akaike Inf. Crit.	41,625.910	36,692.060	32,724.120
Note:			*p<0.1; **p<0.05; ***p<0.01

with sufficient power. We estimate the effect of serving different proportion of ads on the number of clicks and the probability of conversion in a search. We aggregate the data at the search level, and count the number of clicks ($Clicks_i$) and see whether conversion took place during search i. There are 1.34 million searches of varying depths- clothing (0.6 million) and electronics (0.74 million)- by over 600 thousand users.[17]

$Conversion_i = 1$ if there was a conversion during search i, otherwise it equals 0. We then regress clicks or conversion on $Treated_i$, which equals 1 if search i is from a user in Buckets {2,3,4} and equals 0 if search i is from a user in Bucket 1. The exogenous variation in $Treated_i$ identifies the effect of serving higher proportion of ads on the search performance (see Table 4). To estimate the effect on search clicks, we make the standard assumption of exponential mean parametrization and estimate a negative binomial regression model (Cameron and Trivedi, 1998):

$$E[Clicks_i] = \mu_i = exp(b_0 + b_1 Treated_i + e_i), \tag{13}$$

[17]We consider searches of depth 49 or less in our analysis. Around 32% of all clothing related searches are of depth 50 or above, whereas the same number is 8% in electronics related searches.

where μ follows Gamma distribution.

We use a binary logit model to represent the conversion probability during search i.

$$Pr(Conversion_i) = \frac{exp(\bar{U}_i)}{1 + exp(\bar{U}_i)}, \tag{14}$$

where U_i is the latent utility of a user on converting from search i, and can be expressed as follows.

$$U_i = \underbrace{\beta_0 + \beta_1 Treated_i}_{\bar{U}_i} + \epsilon_i, \tag{15}$$

where ϵ represents the search characteristics or features unobservable to the researcher, and follows independent and identically distributed (i.i.d.) extreme value distribution. Here, b_1 (β_1) identifies the average effect of being in different on the probability of conversion from a search.

Table 18: Effect of Ads on Clicks at Search Level

	Dependent variable:		
		Clicks	
	Overall	Electronics	Clothing
	(1)	(2)	(3)
Treated	0.008**	0.005	0.014**
	(0.003)	(0.004)	(0.006)
Constant	−0.682***	−0.644***	−0.732***
	(0.003)	(0.004)	(0.005)
Observations	1,337,965	739,697	598,268
Log Likelihood	−1,290,025.000	−721,153.900	−558,332.300
θ	1.124*** (0.005)	2.377*** (0.023)	0.583*** (0.003)
Akaike Inf. Crit.	2,580,054.000	1,442,312.000	1,116,669.000
Note:			*p<0.1; **p<0.05; ***p<0.01

In Tables 18 and 19, we present the results for number of clicks and the probability of conversion, respectively, in a search. We find that, compared to Bucket 1, the other buckets on average have about 1% higher number of clicks in clothing, but there is no statistically significant difference in electronics with any bucket. When looking at the conversion rates

Table 19: Effect of Ads on Conversion at Search Level

	Dependent variable:		
	Conversion		
	Overall	Electronics	Clothing
	(1)	(2)	(3)
Treated	-0.0002	-0.010	0.011
	(0.011)	(0.014)	(0.016)
Constant	-3.258^{***}	-3.261^{***}	-3.255^{***}
	(0.009)	(0.012)	(0.014)
Observations	1,337,965	739,697	598,268
Log Likelihood	$-211,904.600$	$-116,316.800$	$-95,584.540$
Akaike Inf. Crit.	423,813.200	232,637.600	191,173.100
Note:			*p<0.1; **p<0.05; ***p<0.01

(Table 19), we find no statistically significant difference with any bucket for both the categories. The results seem to suggest that even serving close to 20% 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 section 5.2, provides the basis for this hypothesis. 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 marginal search costs are typically low at top positions. Instead, if ads appear further down the list, a lower fraction of users continue with the search. In the latter case though 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 also important to note that the overall effect also depends on the number of ads in a search. If the proportion of ads becomes very 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.

We conduct power analyses to see if we have enough power in our tests to statistically detect the effects we notice, given our sample size. The analysis, which can be found in the appendix (see Table 28), reveals that we have very low power ($\sim$ 10%) to detect statistically significant effects on conversion even after having around 150,000 observations or more per bucket in each category. We also do a regression analysis where we regress clicks and conversion on each bucket id, instead of all 3 buckets together, and find similar results (see Tables 29-30 in the appendix).

6 Alternative Explanations

In this section, we discuss phenomena, other than the information asymmetry between the platform and the sellers, which can explain our results.

6.1 Consumer Heterogeneity

It is plausible that the heterogeneous effects we find across the two product categories in section 5.1 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 check if this is the case or not, we repeat our analysis with only those users who browsed both electronic and clothing products during the experiment period (see Table 31). The results remain robust, suggesting that it is the product heterogeneity, and not consumer heterogeneity, which is driving the result.

6.2 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. This increases the likelihood of sponsored listings being more irrelevant compared to organic listings, as the pool of sponsored listings is smaller. But, as we mentioned earlier in section

4.1, 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[18] in electronics to see if the poor performance of ads still persists. We find that though the click-through rates of ads can be a little higher, they get converted at a much lower rate than the corresponding organic listings (see Table 32).

6.3 Effect of Unobservable Features

In this subsection, we try to see whether quantifiable features like price, discount, rating and number of ratings can explain all the difference in the performances of sponsored and displaced organic listings found in section 5.1. 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 compare the conversion rates of Bucket 1 (control group) with the other 3 experimental buckets at the second, third and fourth position, respectively, while controlling for listing characteristics, such as rating, price, discount and number of ratings (see Table 33 in the Appendix). Once again we find no difference in the conversion performance between the sponsored and the corresponding organic listings in the clothing category, whereas there is still a statistically significant negative difference between the conversion rates of sponsored and organic listings in the electronics category. This implies that sponsored listings are not better or worse than the substituted organic listings just due to observable and quantifiable features.

[18] We select non-branded and non-descriptive search queries that have at least 100 searches.

7 Robustness Checks

7.1 No Ad Situation

In our main analysis, we measure the impact of different numbers of ads on the probability of conversion from a search. One important scenario to consider is the difference in conversions between searches with ads and without ads. Although we do not have an experimental condition where 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.[19] Users in this bucket did not see any ad in the top 7 listings, while users in the other buckets saw one or two ads. Therefore, we again treat searches till position 7 or less in *Bucket 1* as our control group, and the other 3 buckets as our different treatment groups. We find positive effects on clicks in clothing and small negative effects in the case of clicks in electronics. For conversion, we find no statistically significant impact on search in clothing, but an effect of negative 4% in electronics for Bucket 4 users (see Tables 34-35). We are able to see significant effects here probably because the coefficients do not capture the positive or negative effect of ads on succeeding positions after position 7.

7.2 More Product Verticals

In our main analysis, we focused on only clothing and electronics product categories. In this subsection, we also analyze the data for clothing accessories like watches and shoes to see if the results are similar to clothing. We again find that sponsored listings perform as good as the displaced organic listings (see Table 38). Furthermore, we present the comparison of the performance of organic and sponsored listings at product vertical level for eight verticals in Tables 39-40 to show the robustness of our result at a more granular level.

[19] The top 7 listings account for more than 60% of the total conversions in electronics.

8 Summary and Implications

In this paper, we study how sponsored listings in the search results of an online marketplace perform compared to the relevantly ranked organic listings. We then study how the relevance of sponsored listings impacts the organic listings at succeeding positions, and therefore, also the search performance.

Although we find the overall effect on search to be negligible for both clothing and electronics product categories, the underlying mechanisms behind the result are very different for the two categories. In the case of clothing, ads that appear get clicked or converted at a higher or an equal rate compared to that of the organic listings they displace, and, therefore, there is no major change due to the presence of advertised listings. In the case of electronics, the sponsored listings have a much 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. We show that one of the reasons behind such an observation is the difference in the level of information asymmetry between the marketplace and the sellers regarding the true relevance of clothing and electronic products. When the information gap is high, as in the case of clothing products, more relevant listings get selected as ads, and vice versa. However, even when the information gap is low and ads perform poorly, the platform can recover the lost conversions from the succeeding organic listings due to the positive effects of poor quality ads, given user search costs are low enough.

Our paper has some important implications for e-commerce platform managers. First, we inform the managers that, irrespective of the product category, there is no significant opportunity cost of showing ads in their search results. This implies that sponsored products are only supplementing their core business of transactions by generating more revenue. Second, allowing for sponsored listings seems to be an efficient way for the firms to solve the classic *cold start* problem of recommender systems for goods with non-digital attributes. Instead of running costly experiments with a large number of users to identify good products (exploration-exploitation), sponsored search serves as a tool for managers not only to find good new products, but also to monetize this exploration. Third, based on our findings, managers who are concerned about the consumer experience and long-run value of the platform, may want to use relevance, in addition to the bid, i.e., do quality-score adjustment, in deciding the winner of the auctions in searches related to product categories with more digital attributes. In the case of pricing, they can charge lower prices to sellers of more relevant listings to do some quality control.

Our paper has several limitations that can be addressed in future research. First, our paper looks at only the short term impact of showing ads in the search results on online marketplaces. It might be interesting to examine how the relevance of sponsored listings impacts user behavior and the platform in the long term. Second, our paper does not say much about the optimal positions or the number of sponsored slots for the marketplace. Given that we now know how search ads affect the search performance, the next set of research questions should deal with how many ads to show, and where to show them. The focus of this paper has been on the marketplace's payoffs due to the presence of sponsored products. In future, one can look at the advertisers' short-run and long-run value from using sponsored products 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 good markets: an empirical examination. *International Economic Review*, 44(3), 1007-1040.

[3] Agarwal, A., Hosanagar, K., and Smith, M. D. (2015). Do organic results help or hurt sponsored search performance?. *Information Systems Research*, 26(4), 695-713.

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

[5] Animesh, A., Ramachandran, V., and Viswanathan, S. (2010). Research note- Quality uncertainty and the performance of online sponsored search markets: An empirical investigation. *Information Systems Research*, 21(1), 190-201.

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

[7] Banerjee S., Zhou, Z., and Johari, R. (2014). The importance of exploration in online marketplaces, *Decision and Control (CDC) 2014 IEEE 53rd Annual Conference on*, 3499-3504.

[8] Belleflamme, P., Lambert, T., and Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing*, 29(5), 585-609.

[9] Blake, T., Nosko, C., and Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness: A largescale field experiment. *Econometrica*, 83(1), 155-174.

[10] Cameron, C., and Trivedi, P. (1998). Models for count data.

[11] Caves, R. E., and Greene, D. P. (1996). Brands' quality levels, prices, and advertising outlays: empirical evidence on signals and information costs. *International Journal of Industrial Organization*, 14(1), 29-52.

[12] Choi, H., and Mela, C. F. (2016). Online marketplace advertising. *Available at SSRN.*

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

[14] 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.

[15] Ghose, A., and Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, 55(10), 1605-1622.

[16] Gomes, R. (2014). Optimal auction design in twosided markets. *The RAND Journal of Economics*, 45(2), 248-272.

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

[18] 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.

[19] Johnson, G. A., Lewis, R. A., and Nubbemeyer, E. I. (2017). Ghost ads: Improving the economics of measuring ad effectiveness. *Journal of Marketing Research*, forthcoming.

[20] Katona, Z., and Sarvary, M. (2010). The race for sponsored links: Bidding patterns for search advertising. *Marketing Science*, 29(2), 199-215.

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

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

[23] Lewis, G. (2011). Asymmetric information, adverse selection and online disclosure: The case of eBay motors. *The American Economic Review*, 101(4), 1535-1546.

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

[25] Milgrom, P., and Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy*, 94(4), 796-821.

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

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

[28] Rutz, O. J., Bucklin, R. E.,and Sonnier, G. P. (2012). A latent instrumental variables approach to modeling keyword conversion in paid search advertising. *Journal of Marketing Research*, 49(3), 306-319.

[29] Sahni, N. S., and Nair, H. (2016). Does advertising serve as a signal? Evidence from field experiments in mobile search.

[30] Yao, S., and Mela, C. F. (2011). A dynamic model of sponsored search advertising. *Marketing Science*, 30(3), 447-468.

A Appendix A

A.1 Expressions for δ and θ

$\delta^H = \frac{V_L - V_H - s + 1}{2}$ and $\delta^L = \frac{\frac{V_H - V_L}{2} - s + 1}{2}$. We assume $s < 1 - (V_H - V_L)$ so that $\delta^H > 0$.
$\theta_i = \frac{1+V_i}{2}$, $\theta_{HL} = \frac{(1+V_H)(3+V_H-2V_L)}{8}$, $\theta_{LH} = \frac{(1+V_L)(3-2V_H+V_L)}{8}$ and $\theta_{LL} = \frac{(1+V_L)(3-V_L)}{8}$, as ϵ_i and $\epsilon_i - \epsilon_j$ has joint density $\frac{1}{4}$, with support $\epsilon_i - 1 < \epsilon_i - \epsilon_j < 1 + \epsilon_i$, where $i \neq j$.

A.2 Derivation of v_H and v_L

The platform ranks the three listings with the following probabilities. $Pr(HL_iL_j) = \frac{\kappa}{2}$, $Pr(L_iHL_j) = \frac{1-\kappa}{4}$ and $Pr(L_iL_jH) = \frac{1-\kappa}{4}$ $\forall\ i, j \in \{1, 2\}$ and $i \neq j$. Let v_i denote the maximum willingness to pay per click of the seller of listing type $i \in \{H, L\}$ for the sponsored slot. Note that when $CPC > max\{v_H, v_L\}$, then no listing is advertised, leading to no effect of having a sponsored slot. On the other hand, when $CPC < min\{v_H, v_L\}$, there will be no change in the ranking order due to introduction of the sponsored slot, as all the listings are in the advertising pool. Therefore, in order to concentrate on cases where ranking order can potentially alter due to introduction of the sponsored slot, we will assume that $v_i > CPC > v_j$, where $i, j \in \{H, L\}$ and $i \neq j$.

If H gets the sponsored slot, its payoff is given by π_1^{HL}. Instead, if L gets the sponsored slot, H's payoff is $(\kappa+\frac{1-\kappa}{2})\pi_2^{LH}$, as $\kappa+\frac{1-\kappa}{2}$ is the probability with which H is ranked higher than at least one listing. Then,

$$v_H = \frac{\pi_1^{HL} - (\kappa+\frac{1-\kappa}{2})\pi_2^{LH}}{\beta_1^{HL}} = \gamma_H\Big(1-(\kappa+\frac{1-\kappa}{2})\frac{\beta_2^{LH}}{\beta_1^{HL}}\Big) > 0, \tag{16}$$

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

Similarly, if H gets the sponsored slot, L's payoff is given by $\frac{\pi_2^{HL}}{2}$. When an L gets the sponsored slot, L_i occupies the first position with probability $\frac{1}{2}$, H occupies the second position with probability $\kappa+\frac{1-\kappa}{2}$ and an L occupies the second position with probability $\frac{(1-\kappa)}{2}$, where $i \neq j$. Then,

$$\begin{aligned} v_L &= (\kappa+\frac{1-\kappa}{2})\frac{(\frac{\pi_1^{LH}}{2}+\frac{0}{2}-\frac{\pi_2^{HL}}{2})}{\frac{\beta_1^{LH}}{2}} + \frac{(1-\kappa)}{2}\frac{(\frac{\pi_1^{LL}}{2}+\frac{\pi_2^{LL}}{2}-\frac{\pi_2^{HL}}{2})}{\frac{\beta_1^{LL}}{2}} \\ &= \gamma_L\Big(1+\frac{(1-\kappa)}{2}\frac{(\beta_2^{LL}-\beta_2^{HL})}{\beta_1^{LL}} - (\kappa+\frac{1-\kappa}{2})\frac{\beta_2^{HL}}{\beta_1^{LH}}\Big) > 0, \end{aligned} \tag{17}$$

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

A.3 Proof of Proposition 1

When $\kappa = 1$, we have $v_L > v_H$ under the following condition.

$$\frac{\gamma_H\beta_2^{LH}}{\beta_1^{HL}} - \frac{\gamma_L\beta_2^{HL}}{\beta_1^{LH}} > \gamma_H - \gamma_L, \text{ or}$$

$$\frac{1-\frac{\beta_2^{HL}}{\beta_1^{LH}}}{1-\frac{\beta_2^{LH}}{\beta_1^{HL}}} > \frac{\gamma_H}{\gamma_L} \tag{18}$$

Therefore, there exists a CPC such that $v_L > CPC > v_H$, implying only L gets advertised.

A.4 Proof of Proposition 2

We know, at $\kappa = 1$ and when $\frac{\gamma_H \beta_2^{LH}}{\beta_1^{HL}} - \frac{\gamma_L \beta_2^{HL}}{\beta_1^{LH}} > \gamma_H - \gamma_L$, $v_L > v_H$. The rate of increase in v_H is greater than the rate of increase in v_L due to a decrease in κ if the following holds.

$$\frac{\gamma_H \beta_2^{LH}}{\beta_1^{HL}} - \frac{\gamma_L \beta_2^{HL}}{\beta_1^{LH}} > \frac{\gamma_L(\beta_2^{LL} - \beta_2^{HL})}{\beta_1^{LL}}. \tag{19}$$

If the RHS of inequality (19) is lower than $\gamma_H - \gamma_L$, then the rate of increase in v_H is always higher. Otherwise, it is only higher when the LHS is big enough. We have $v_H = v_L$ when $\kappa = \bar{\kappa}$, where

$$\bar{\kappa} = 1 + \frac{2\beta_1^{LL}(\gamma_H(\beta_1^{HL} - \beta_2^{LH})\beta_1^{LH} + \gamma_L \beta_1^{HL}(\beta_2^{HL} - \beta_1^{LH}))}{\gamma_H \beta_2^{LH}\beta_1^{LH}\beta_1^{LL} - \gamma_L \beta_1^{HL}(\beta_1^{LL}\beta_2^{HL} + \beta_1^{LH}(\beta_2^{LL} - \beta_2^{HL}))}. \tag{20}$$

When inequality (19) holds, then the high type has a higher valuation for the sponsored slot than the low type for $\kappa < \bar{\kappa}$. Therefore, there exists a CPC such that $v_H > CPC > v_L$ when $\kappa < \bar{\kappa}$. Otherwise, when $\kappa \geq \bar{\kappa}$ we have $v_L > v_H$.

A.5 Proof of Corollary 1

When L occupies the sponsored slot, the effect at the first position in the case with no IA is given by $\pi_1^{LH} - \pi_1^{HL} < 0$. When H occupies the sponsored slot, the effect at the first position in the case with IA is given by following.

$\pi_1^{HL} - (\kappa \pi_1^{HL} + \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_1^{LL})) = (1-\kappa)\pi_1^{HL} - \frac{(1-\kappa)}{2}(\pi_1^{LH} + \pi_1^{LL}) > 0 \; \forall \; \kappa < \bar{\kappa}$, 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 no IA is given by $\pi_2^{LH} - \pi_2^{HL} > 0$. When H occupies the sponsored slot, the effect at the first position in the case with IA is given by following.

$\pi_2^{HL} - (\kappa\pi_2^{HL} + \frac{(1-\kappa)}{2}(\pi_2^{LH} + \pi_2^{LL}) = (1-\kappa)\pi_2^{HL} - \frac{(1-\kappa)}{2}(\pi_2^{LH} + \pi_2^{LL}) < 0 \;\forall\; \kappa < \bar{\kappa}$, 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 no IA, the overall effect on the two positions in the search is given by $\Delta = \pi_1^{LH} + \pi_2^{LH} - \pi_1^{HL} - \pi_2^{HL} < 0$. We have $\frac{\partial\Delta}{\partial s} < 0$, as $\frac{\partial(\pi_1^{LH}+\pi_2^{LH})}{\partial s} < 0$ and $\frac{\partial(\pi_1^{HL}+\pi_2^{HL})}{\partial s} > 0$.

When H occupies the sponsored slot in the case with IA, the overall effect is given by $\Delta_\kappa = \pi_1^{HL} + \pi_2^{HL} - \Big(\kappa(\pi_1^{HL} + \pi_2^{HL}) + \frac{(1-\kappa)}{2}(\pi_2^{LH} + \pi_1^{LH}) + \frac{(1-\kappa)}{2}(\pi_2^{LL} + \pi_1^{LL})\Big) > 0$. Δ_κ is always positive as the introduction of the sponsored slot removes the possibility of two low types in the first two positions. Let us denote the second term of Δ_κ as π_κ. We have $\frac{\partial(\pi_1^{HL}+\pi_2^{HL})}{\partial s} > 0$ and $\frac{\partial(\pi_1^{HL}+\pi_2^{HL})}{\partial s} > \frac{\partial\pi_\kappa}{\partial s} \;\forall\; \kappa < 1$. Therefore, $\frac{\partial\Delta_\kappa}{\partial s} > 0$.

Table 20: Randomization Check from 2 weeks before the experiment

Category	Bucket	Impressions	Click-through Rate (%)	Conversion per Impression (%)
Clothing	1	6397940	2.88	0.203
Clothing	2	6347051	2.89	0.204
Clothing	3	7639873	2.88	0.202
Clothing	4	7647181	2.88	0.206
Electronics	1	6390098	6.09	0.334
Electronics	2	6359573	6.12	0.336
Electronics	3	7641974	6.11	0.337
Electronics	4	7645129	6.10	0.341

Table 21: Randomization Check- Sponsored Listings

Category	Bucket	Rating		Rating Count		Price (Rs.)		Discount(%)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Clothing	1	3.65	1.07	5	17	935	948	57	18.0
Clothing	2	3.67	1.07	5	17	938	967	57	18.2
Clothing	3	3.66	1.07	5	17	937	970	57	18.2
Clothing	4	3.67	1.07	5	17	934	958	57	18.3
Electronics	1	3.72	0.64	138	660	12563	18114	16.8	25.6
Electronics	2	3.72	0.60	137	650	13542	18497	13.8	23.7
Electronics	3	3.72	0.60	134	631	13404	18474	14.2	23.9
Electronics	4	3.72	0.63	143	673	12678	18197	15.9	25.0

Table 22: Power Analysis for Table 9 ($\alpha = 0.05$)

Position	Difference	Power
2	11%	49%
3	10%	37%
4	10%	34%

Table 23: Characteristics of Sponsored Listings at Positions {2,3,4}

	Dependent variable:
	Sponsored
Rating	0.268***
	(0.002)
LogPrice	0.411***
	(0.005)
Discount	−0.008***
	(0.0001)
LogRatingCount	−0.281***
	(0.002)
Rating*Electronics	−0.192***
	(0.004)
LogPrice*Electronics	−0.590***
	(0.005)
Discount*Electronics	−0.011***
	(0.0002)
LogRatingCount*Electronics	−0.002
	(0.002)
Electronics	5.548***
	(0.040)
Constant	−3.914***
	(0.035)
Observations	3,242,829
Log Likelihood	−1,526,187.000
Akaike Inf. Crit.	3,052,394.000
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 24: Difference between the CTR of Sponsored and Organic Listings in Electronics

	Dependent variable:				
	Click				
	(2)	(3)	(4)	(6)	(7)
Sponsored	−0.256***	−0.056***	−0.054***	−0.168***	−0.191***
	(0.013)	(0.016)	(0.020)	(0.026)	(0.026)
Constant	−2.408***	−2.931***	−3.155***	−3.369***	−3.290***
	(0.008)	(0.011)	(0.013)	(0.017)	(0.017)
Observations	363,602	323,782	257,160	202,393	185,729
Log Likelihood	−96,351.030	−63,735.420	−43,139.140	−27,901.010	−26,965.020
Akaike Inf. Crit.	192,706.100	127,474.800	86,282.270	55,806.010	53,934.040
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 25: Difference between the CTR of Sponsored and Organic Listings in Clothing

	Dependent variable:				
	Click				
	(2)	(3)	(4)	(6)	(7)
Sponsored	0.186***	0.231***	0.255***	0.178***	0.214***
	(0.015)	(0.017)	(0.017)	(0.020)	(0.021)
Constant	−3.155***	−3.465***	−3.407***	−3.533***	−3.671***
	(0.011)	(0.013)	(0.012)	(0.014)	(0.015)
Observations	407,257	400,898	401,000	346,934	332,451
Log Likelihood	−74,333.050	−59,393.710	−62,553.550	−47,881.630	−41,927.670
Akaike Inf. Crit.	148,670.100	118,791.400	125,111.100	95,767.260	83,859.340
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 26: Difference between the Conversion of Sponsored and Organic Listings in Electronics

	Dependent variable:				
	Conversion				
	(2)	(3)	(4)	(6)	(7)
Sponsored	-0.463^{***}	-0.290^{***}	-0.281^{***}	-0.302^{***}	-0.312^{***}
	(0.047)	(0.061)	(0.084)	(0.103)	(0.106)
Constant	-5.023^{***}	-5.522^{***}	-5.938^{***}	-6.109^{***}	-6.075^{***}
	(0.028)	(0.038)	(0.052)	(0.065)	(0.066)
Observations	363,602	323,782	257,160	202,393	185,729
Log Likelihood	−12,257.130	−7,571.964	−4,242.011	−2,848.444	−2,684.735
Akaike Inf. Crit.	24,518.260	15,147.930	8,488.022	5,700.888	5,373.469
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 27: Difference between the Conversion of Sponsored and Organic Listings in Clothing

	Dependent variable:				
	Conversion				
	(2)	(3)	(4)	(6)	(7)
Sponsored	0.101^{*}	0.097	0.093	−0.004	0.110
	(0.052)	(0.060)	(0.060)	(0.074)	(0.075)
Constant	-5.675^{***}	-5.922^{***}	-5.926^{***}	-6.148^{***}	-6.210^{***}
	(0.037)	(0.043)	(0.043)	(0.051)	(0.054)
Observations	407,257	400,898	401,000	346,934	332,451
Log Likelihood	−9,693.661	−7,726.644	−7,691.773	−5,284.237	−5,037.053
Akaike Inf. Crit.	19,391.320	15,457.290	15,387.550	10,572.480	10,078.100
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 28: Power Analysis for Section 5.3 ($\alpha = 0.05$)

Category	Variable	Difference	Power
Electronics	Clicks	0.002	20%
Electronics	Conversion	-0.0003	10%
Clothing	Clicks	0.007	65%
Clothing	Conversion	0.0004	11%

Table 29: Effect on the Number of Search Clicks

	Dependent variable:	
	Clicks	
	Clothing	Electronics
	(1)	(2)
Bucket 2	0.012*	0.008
	(0.007)	(0.005)
Bucket 3	0.011	0.004
	(0.007)	(0.005)
Bucket 4	0.020***	0.002
	(0.007)	(0.005)
Constant	−0.732***	−0.644***
	(0.005)	(0.004)
Observations	598,268	739,697
Log Likelihood	−558,331.500	−721,153.100
θ	0.583*** (0.003)	2.377*** (0.023)
Akaike Inf. Crit.	1,116,671.000	1,442,314.000
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 30: Effect on the Probability of Search Conversion

	Dependent variable:	
	Conversion	
	Clothing	Electronics
	(1)	(2)
Bucket 2	−0.025	−0.005
	(0.019)	(0.017)
Bucket 3	0.031	0.004
	(0.019)	(0.017)
Bucket 4	0.028	−0.029
	(0.019)	(0.018)
Constant	−3.255***	−3.261***
	(0.014)	(0.012)
Observations	598,268	739,697
Log Likelihood	−95,579.190	−116,315.000
Akaike Inf. Crit.	191,166.400	232,638.000
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 31: Difference between Sponsored and Organic for Users who browsed both Categories

	Dependent variable:	
	Click	Conversion
	(1)	(2)
Sponsored	0.173***	0.012
	(0.018)	(0.067)
Sponsored*Electronics	−0.296***	−0.390***
	(0.027)	(0.103)
Position	−0.232***	−0.252***
	(0.007)	(0.026)
Electronics	0.491***	0.356***
	(0.013)	(0.046)
Constant	−2.633***	−5.129***
	(0.022)	(0.080)
Observations	736,654	736,654
Log Likelihood	−133,426.600	−16,069.150
Akaike Inf. Crit.	266,863.200	32,148.300
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 32: Difference between Sponsored and Organic listings for Generic Queries in Electronics

	Dependent variable:	
	Click	Conversion
	(1)	(2)
Sponsored	0.059***	−0.688***
	(0.020)	(0.080)
Position	−0.131***	−0.162***
	(0.011)	(0.035)
Constant	−2.957***	−5.073***
	(0.033)	(0.103)
Observations	384,880	384,880
Log Likelihood	−58,285.610	−8,840.018
Akaike Inf. Crit.	116,577.200	17,686.040
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 33: Difference between Sponsored and Organic listings after Controlling for Observables

	Dependent variable:	
	Click	Conversion
	(1)	(2)
Sponsored	0.048^{***}	-0.035
	(0.011)	(0.039)
Sponsored*Electronics	-0.240^{***}	-0.295^{***}
	(0.014)	(0.050)
Position	-0.251^{***}	-0.288^{***}
	(0.003)	(0.011)
Rating	0.0003	0.049^{***}
	(0.004)	(0.013)
LogPrice	0.138^{***}	-0.015
	(0.003)	(0.010)
Discount	-0.0003	0.001^{*}
	(0.0002)	(0.001)
LogRatingCount	-0.026^{***}	0.056^{***}
	(0.001)	(0.004)
Electronics	0.203^{***}	0.239^{***}
	(0.011)	(0.034)
Constant	-3.323^{***}	-5.150^{***}
	(0.032)	(0.114)
Observations	3,242,829	3,242,829
Log Likelihood	−635,545.500	−84,072.740
Akaike Inf. Crit.	1,271,109.000	168,163.500
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 34: Effect of Ads on Clicks from Top 7 Listings

	Dependent variable:		
	Clicks		
	Overall	Electronics	Clothing
	(1)	(2)	(3)
Bucket 2/3/4	-0.007^{*}	-0.014^{***}	0.018^{**}
	(0.004)	(0.004)	(0.007)
Constant	-1.161^{***}	-0.935^{***}	-1.543^{***}
	(0.003)	(0.004)	(0.006)
Observations	1,337,965	739,697	598,268
Log Likelihood	−956,196.200	−596,096.700	−341,036.500
θ	2.864^{***} (0.040)	22.339^{***} (2.031)	0.854^{***} (0.011)
Akaike Inf. Crit.	1,912,396.000	1,192,197.000	682,077.000
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 35: Effect of Ads on Conversion from Top 7 Listings

	Dependent variable:		
	Conversion		
	Overall	Electronics	Clothing
	(1)	(2)	(3)
Bucket 2/3/4	-0.028^{**}	-0.029^{*}	−0.020
	(0.013)	(0.016)	(0.022)
Constant	-3.709^{***}	-3.527^{***}	-3.994^{***}
	(0.011)	(0.014)	(0.019)
Observations	1,337,965	739,697	598,268
Log Likelihood	−148,659.500	−94,360.450	−53,541.750
Akaike Inf. Crit.	297,323.000	188,724.900	107,087.500
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 36: Effect of Ads on Clicks from Top 7 Listings

	Dependent variable:	
	Clicks	
	Clothing	Electronics
Bucket 2	0.014	−0.008
	(0.009)	(0.005)
Bucket 3	0.020**	−0.011**
	(0.009)	(0.005)
Bucket 4	0.021**	−0.023***
	(0.009)	(0.005)
Constant	−1.543***	−0.935***
	(0.006)	(0.004)
Observations	598,268	739,697
Log Likelihood	−341,036.100	−596,092.600
θ	0.854*** (0.011)	22.352*** (2.033)
Akaike Inf. Crit.	682,080.200	1,192,193.000
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 37: Effect of Ads on Conversion from Top 7 Listings

	Dependent variable:	
	Conversion	
	Clothing	Electronics
Bucket 2	−0.046*	−0.023
	(0.028)	(0.020)
Bucket 3	−0.015	−0.017
	(0.028)	(0.020)
Bucket 4	0.001	−0.048**
	(0.027)	(0.020)
Constant	−3.994***	−3.527***
	(0.019)	(0.014)
Observations	598,268	739,697
Log Likelihood	−53,540.230	−94,359.190
Akaike Inf. Crit.	107,088.500	188,726.400
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 38: Difference between Sponsored and Organic listings for Clothing Accessories

	Dependent variable:	
	Click	Conversion
	(1)	(2)
Sponsored	0.129^{***}	-0.029
	(0.008)	(0.027)
Position	-0.128^{***}	-0.174^{***}
	(0.005)	(0.014)
Constant	-3.120^{***}	-5.304^{***}
	(0.014)	(0.043)
Observations	2,542,698	2,542,698
Log Likelihood	−344,441.500	−51,102.870
Akaike Inf. Crit.	688,889.000	102,211.700
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 39: Difference between CTR of Sponsored and Organic Listings by Product Vertical

	Dependent variable:							
	Click							
	Mobile	Headphone	Television	Shirt	Top	Sari	Watch	Shoe
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sponsored	-0.158^{***}	-0.308^{***}	-0.255^{***}	0.140^{***}	0.264^{***}	0.316^{***}	0.069^{***}	0.162^{***}
	(0.009)	(0.028)	(0.046)	(0.022)	(0.030)	(0.013)	(0.014)	(0.010)
Position	-0.375^{***}	-0.261^{***}	-0.180^{***}	-0.130^{***}	-0.081^{***}	-0.130^{***}	-0.137^{***}	-0.124^{***}
	(0.005)	(0.014)	(0.024)	(0.012)	(0.017)	(0.008)	(0.008)	(0.006)
Constant	-1.528^{***}	-2.481^{***}	-2.647^{***}	-2.968^{***}	-3.251^{***}	-2.832^{***}	-3.057^{***}	-3.155^{***}
	(0.013)	(0.040)	(0.069)	(0.035)	(0.053)	(0.023)	(0.023)	(0.017)
Observations	1,324,603	252,536	79,138	338,490	168,597	670,476	884,720	1,657,978
Log Likelihood	−329,971.200	−38,670.440	−12,872.180	−51,032.580	−23,685.770	−116,021.900	−121,946.000	−222,469.700
Akaike Inf. Crit.	659,948.500	77,346.870	25,750.360	102,071.200	47,377.550	232,049.900	243,898.000	444,945.400
Note:								*p<0.1; **p<0.05; ***p<0.01

Table 40: Difference between Conversion of Sponsored and Organic Listings by Product Vertical

	Dependent variable:							
	Conversion							
	Mobile	Headphone	Television	Shirt	Top	Saree	Watch	Shoe
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sponsored	-0.436^{***}	-0.473^{***}	-0.588^{***}	0.159^{**}	0.080	0.175^{***}	-0.060	-0.011
	(0.035)	(0.088)	(0.192)	(0.080)	(0.099)	(0.046)	(0.040)	(0.037)
Position	-0.461^{***}	-0.155^{***}	-0.275^{***}	-0.170^{***}	-0.137^{**}	-0.151^{***}	-0.163^{***}	-0.185^{***}
	(0.017)	(0.040)	(0.090)	(0.043)	(0.055)	(0.026)	(0.021)	(0.019)
Constant	-3.953^{***}	-5.011^{***}	-5.052^{***}	-5.536^{***}	-5.412^{***}	-5.259^{***}	-5.062^{***}	-5.458^{***}
	(0.046)	(0.119)	(0.258)	(0.130)	(0.167)	(0.078)	(0.064)	(0.058)
Observations	1,324,603	252,536	79,138	338,490	168,597	670,476	884,720	1,657,978
Log Likelihood	−41,378.210	−6,373.058	−1,427.469	−5,838.000	−3,475.780	−15,505.530	−22,321.750	−28,598.480
Akaike Inf. Crit.	82,762.410	12,752.120	2,860.939	11,682.000	6,957.560	31,017.070	44,649.490	57,202.960

Note: *p<0.1; **p<0.05; ***p<0.01