Sponsored Search: Do Organic Results help or hurt the Performance and under what conditions?

Completed Research Paper

Ashish Agarwal  
Assistant Professor  
McCombs School of Business  
University of Texas, Austin  
Austin, TX 15213  
ashish.agarwal@mccombs.utexas.edu  

Kartik Hosanagar  
Associate Professor of Operations and Information Management  
The Wharton School  
University of Pennsylvania  
552 Jon M. Huntsman Hall  
3730 Walnut Street  
Philadelphia, PA 19104  
kartikh@wharton.upenn.edu  

Michael D. Smith  
Professor of Information Technology and Marketing  
H. John Heinz III School of Public Policy and Management and Tepper School of Business  
Carnegie Mellon University  
4800 Forbes Avenue, HBH 3028  
Pittsburgh PA, 15217  
mds@cmu.edu  

Abstract  
We study the relative impact of competing links in organic and sponsored search results on the performance of sponsored search ads. We use data generated through a field experiment for several keywords from the ad campaign of an online retailer. Using a hierarchical Bayesian model, we measure the impact of competition on both click-through rate and conversion rate of sponsored search ads for these keywords. We find that the competitor links in organic results have a higher impact on the performance as compared to the competitor links in sponsored results. We also find that competition has a greater influence on the conversion performance as compared to the click through performance. Our results inform advertisers on the impact of organic results on their performance. Our results reveal inefficiency in the current auction mechanism as the click performance may not reveal the true quality of advertisers.  

Keywords: Sponsored search, Organic search, ad placement, hierarchical Bayesian estimation, online advertising, online auctions, search engine marketing
Introduction

Internet advertising spend is currently growing faster than any other form of advertising and is expected to grow from $23.4 billion in 2008 to $34 billion in 2014 (eMarketer 2009). 40% of this ad spend occurs on sponsored search, where advertisers pay to appear alongside the regular search results of a search engine. Most search engines, including Google, Yahoo, and MSN, use auctions to sell their ad space inventory. In these auctions, advertisers submit bids on specific keywords based on their willingness to pay for a click from a consumer searching on that (or a closely related) keyword. Search engines use a combination of the submitted bid and past click performance to rank order the ads. When a customer uses a search engine for a product, advertiser's competitors can appear in both sponsored and organic results. For example, figure 1 shows search results for user keyword 'shirts'. While 'nordstrom.com' appears in sponsored results, its competitors appear in both sponsored results (epromos.com) and organic results (macys.com, oldnavy.com) An important question is whether the presence of competing links in organic results impact the sponsored search performance.

![Search Results](image)

**Figure 1: Search Results**

Search engines claim to manage these results separately.¹ In that case, one possibility is that consumers view these results independent of each other. However, consumers are used to viewing regular (organic) results to satisfy their information needs. Advertisers can attain top rank in sponsored search by submitting very high bids. Organic results are determined by the search engine based on the relevance and the popularity of the page for the given keyword, and are thus less likely to be influenced by advertisers. Because of this, consumers may trust the organic results more than the sponsored search results. This can have a negative effect on the sponsored search performance. Eye tracking studies have

¹ [http://www.google.com/support/forum/p/AdWords/thread?tid=223f4519f6b545a7&hl=en](http://www.google.com/support/forum/p/AdWords/thread?tid=223f4519f6b545a7&hl=en)
shown that consumers tend to focus more on organic results as compared to the sponsored search results.² While there is a growing literature on sponsored search (e.g., Edelman & Ostrovsky, 2007; Weber and Zheng; 2007; Athey et al 2007; Ghose and Yang 2009) there have been very few papers in the literature studying the interplay between these two types of search results. Yang and Ghose (2009) is the only empirical paper we are aware of to analyze this interplay. These authors find that clicks on organic links have a complementary effect on the clicks in sponsored links i.e. when an advertiser’s link appears in both organic and sponsored results users are more likely to click on these links. However, there are several follow up questions that emerge and are not answered by their study. For example, it is not known what is the relative impact of competitor links in both organic and sponsored search results on the performance of sponsored search results. Note that this applies irrespective of the presence or absence of advertiser’s own organic link. Additionally, Yang & Ghose (2009) do not explicitly study the effect of organic links on conversions. This is important because there are different kinds of users who could be driving the results for clicks and conversions as shown by Agarwal, Hosanagar, and Smith (2011). In that case the relevant question is whether there are differences in the impact of competition on the click performance and the conversion performance.

In this paper we address these questions by empirically analyzing how organic search results impact the click and conversion performance in sponsored search. We use a field experiment to generate a unique panel dataset of daily clicks, orders, and cost for multiple keywords in the sponsored search ad campaign of an online retailer. We also use search results data generated for the sample keywords using a web crawler during the period of our experiment. We use a hierarchical Bayesian model to analyze the probabilities for clicking and ordering in this environment. We find that the competing organic links have a higher impact on the conversion performance of sponsored search as compared to the competing sponsored links. We also find that the competing links have a higher negative effect on conversions as compared to their effect on clicks.

Our paper makes several contributions. First, our paper provides key managerial insights for advertisers. Many websites do not appear in organic results and need to determine their marketing strategy on how to respond to competing organic links. Our results show that advertisers should pay more attention to the competition in organic results. Additionally, competition is more important if they are interested in transactional revenue as compared to the branding effort which is primarily driven by clicks (Rutz and Bucklin, 2010).

Current auction mechanism is advertised to be independent of organic results and a tool for small advertisers to reach out to consumers. However, our results show that organic results impose externality on sponsored search results and maybe more important that the sponsored results. As a consequence, advertisers who only appear in sponsored results and are interested in transactional revenue, are at a disadvantage as their ROI can be significantly lower as compared to advertisers who appear in both places even if they get the same click performance. Such advertisers will find it very expensive to participate in sponsored search and this lower participation can lead to lowering of revenue for search engines. Our results also suggest that there may be inefficiencies in the current mechanism used by search engines which uses only click information to compare performance across advertisers. As consumers interested in buying are more likely to pay attention to quality, an advertiser maybe labeled as high quality even if it is getting very few conversions. Thus, the search engine maybe generating lower revenue by labeling these advertisers as high quality. For an advertiser, interested in transactional revenue, the clicks can be misleading as it may not get a proportional share of conversions.

Finally, we also provide insight into consumer behavior in sponsored search environments. Our results suggest that quality of competition plays a more significant role during conversions as compared to clicks. This can be attributed to the differences in the type of users which drive the click and conversion performance. While the click performance is primarily driven by consumers who are information seekers, the conversion performance is driven by the serious buyers (Agarwal et al. 2011). Our results suggest that, consumers who are serious buyers pay more attention to the quality of competition. Additionally,

² http://www.webcitation.org/5FmwyPgDv
consumers give higher importance to the quality of competition in organic results. This suggests that consumers trust organic results more than the sponsored results.

**Literature Review**

The literature most relevant to our study includes past research on consumers’ online search behavior, with a special emphasis on the impact of quality of competition, and the research focused on advertisers' performance in sponsored search markets.

**Consumers’ online search behavior**

An important consideration in evaluating the performance of the sponsored search advertisements is the quality perception of the advertiser relative to its competitors. In a study of the effect of competition on advertising memory recall, Kent and Allen (1994) show that consumers are more likely to recall familiar brands. Dodds et al. (1991) show that consumer perceive familiar brands with higher quality. Kardes et al. (1993) show that pioneering brand is more likely to be retrieved, considered, and selected. Consumers are known to make quality and price tradeoffs in the choice stage (Nedungadi, 1990). Erdem and Swait (2006) show that the brand credibility is important for both consideration and choice. Degeratu, Rangaswamy and Wu (2002) show that in online choice, brand name is more important when less information is available about other attributes. In the sponsored search context, this would suggest that consumers would perceive familiar sites as of higher quality and this would influence how they select the sponsored search ad. Additionally, if consumers are looking at both organic and sponsored search results they would be influenced by the quality perception of links appearing in organic search. As both organic and sponsored search results are ordered lists, an important question is how the position of a link in these results influences its performance.

Prior work in traditional media has demonstrated that message ordering influences ad persuasion (Rhodes et al. 1973, Brunel and Nelson, 2003). Similar results have been shown in online environments. In fact, Hoque and Lohse (1999) find that consumers are more likely to choose advertisements near the beginning of an online directory than they are when using paper directories. Ansari and Mela (2003) have found that the higher position of links in email campaign can lead to higher probability of clicking. There is some evidence for this in the context of search engine. Using eye tracking analysis, Granka et al. (2004) find that users generally investigate search results sequentially and do so top down. Feng et al. (2007) find evidence of an exponential decrease in the number of clicks for an ad with its rank, and attribute this to decay in user attention as one proceeds down a list. Similarly, Ghose and Yang (2009) find that click through rate of an ad decreases with position. This would suggest that position of an advertiser plays an important role in forming a perception about the quality of the advertiser. Another important consideration is whether the quality perception depends on the type of consumers.

Online consumers include both buying consumers and information seekers (Moe 2003, Moe and Fader, 2004; Montgomery, Li, Srinivasan, and Lietchy, 2004). Moe (2003) shows that consumers with high purchase intent tend to be very focused in their search, targeting a few products and categories versus consumers with low purchase intent, who have broad search patterns. Using path analysis, Montgomery, Li, Srinivasan, and Lietchy (2004) show that consumers with directed search have higher probability of purchase. A similar pattern can be expected in sponsored search: consumers may be heterogeneous in terms of their purchase intent, and resulting search behavior. This heterogeneity in consumer search can result in different performance outcomes. As the conversion rate in sponsored search is very low, the clicking performance primarily depends on the information seekers while the conversion performance depends on the serious buyers (Agarwal, Hosanagar, and Smith 2011). Agarwal, Hosanagar, and Smith (2011) show that buying consumers are more likely to visit lower ad positions in sponsored search as compared to the information seekers and this results in a increase in conversion rate with a decrease in position for the top few positions. These different patterns of search may also result in the differential impact of competing links on the click and conversion performance in sponsored search. Broad pattern of search by the information seekers may reflect less sensitivity to the quality. Moe (2006) shows that consumers tend to use simplified decision rules and may not consider certain attributes in the initial stage of purchase. On the other hand, focused search on part of buying consumers may reflect more selectivity.
in clicking and buying from ads. Thus, the effect of competing links in organic results can be different for clicks and conversions as different users are driving the overall performance.

**Sponsored search markets**

Existing work in sponsored search has focused on auction design, consumer behavior, and advertiser strategy. In terms of work on auction design, Edelman et al. (2007) and Varian (2006) compute the equilibria of the generalized second price sponsored search auction and demonstrate that the auction, unlike the Vickrey-Clarke-Groves (VCG) mechanism, is not incentive compatible. Thus, advertisers will bid strategically in these auctions. Edelman and Ostrovsky (2007) examine data on paid search auctions and find evidence of strategic bidder behavior. Feng et al. (2007) and Weber and Zheng (2007) compare the performance of various ad-ranking mechanisms, finding that a yield-optimized auction, with ranking based on a combination of the submitted bid and ad relevance, provides the highest revenue to the search engine. However, none of these studies consider the impact of organic results on the advertiser performance and the auction mechanism. Jansen and Spink (2006) find a negative bias for sponsored search ads versus organic search results. This suggests that consumers are less likely to click on sponsored results. Katona and Sarvary (2008) and Xu et al. (2009) have assumed this behavior on part of the consumer in their theoretical work. Ghose and Yang (2009) and Agarwal, Hosanagar, and Smith (2011) study the impact of ad position on the click and conversion performance. These papers do not explicitly study the impact of organic results on sponsored search. Yang and Ghose (2009) find complementarity between clicks on organic results and clicks on sponsored results. However, they do not consider the impact on competing links on sponsored search results. So they cannot establish whether or not organic links of competitors can hurt the click performance. Additionally, while they model conversions, Yang and Ghose (2009) do not explicitly study the effect of organic results on the conversion performance.

Thus, the prior literature suggests that consumers may consider the quality of links while making choices. Additionally, the search behavior depends on the type of consumer. However, it is not known as to how consumers respond to competing links in organic search results and how does this depend on whether they are just clicking or also buying from the advertiser.

**Field Experiment and Data**

Our main dataset were generated through a field experiment for a sponsored search ad campaign on Google for an online retailer for pet products and is similar to the dataset used by Agarwal, Hosanagar, and Smith (2011). The data were generated by submitting randomized bids for several keywords and measuring the consumer response in terms of clicks and orders for different positions of the ads corresponding to the keywords. These keywords were randomly chosen from a set of keywords in the campaign related to the food product category that had generated orders in the past for the retailer. We used an automated web crawler to determine the organic results that consumers would see in response to their search queries corresponding to the experimental keywords. Google allows advertisers to use ‘broad’, ‘exact’ or ‘phrase’ match for their keywords. An ‘exact’ match ensures that the search query exactly matches keywords, and to ensure replicability, we have used only keywords with an ‘exact’ match in our sample. The resulting data includes the competing sites in the organic search and their relative position.

Following the literature, we use Alexa rank obtained from alexa.com as the measure of the perceived quality of competitor links appearing in sponsored search results and organic search results (Brynjolfsson and Smith 2000, Palmer 2002, Animesh et al 2010). Web usage parameters such as internet traffic to a website has been found to have strongly positive relationship with the market value of firms (Trueman et al. 2000, Demers and Lev 2001). Note that not all listings in the organic results are competitors. For example, ‘wikipedia.com’ appears regularly in the results for many search queries. Similarly, other information portals as well as news links also appear in organic results. We take this into account while determining the average position of competitors. Additionally, we have verified the list of competitors with our advertiser. In order to determine the quality perception for competitors we consider the cumulative log value of the Alexa rank of competitors in sponsored results relative to Alexa rank of our advertiser. We calculate the perceived quality separately for organic results and sponsored search results. This is to account for the differential impact of the two types of results. In order to account for other factors that may influence the consumer click behavior on sponsored results, we also have the quality score measure maintained by the search engine and available to the advertiser. This measure represents
the click propensity of an advertiser and is based on several metrics such as the relative click performance of the advertiser for the keyword, the relative overall click performance of the advertiser, the relative quality of the ad and the relevance of the ad for the keyword. Search engine is known to use a sliding window to determine the value of the quality score. However, this value remained unchanged for all our keywords during the course of the experiment. In order to calculate the quality of competing links we only consider the first page results as our advertiser always appears on the first page.

The data set consists of 1228 observations of daily impressions, clicks, and orders for 33 keywords over a 45-day period from June 2009 to July 2009. Summary statistics are provided in Table 1. Note that the observations represent daily aggregate data for ads corresponding to the sample keywords for our advertiser and the dataset is typical of the information received by advertisers in sponsored search. We do not have information on the performance of competing ads or detailed information on how an individual consumer makes a choice during a search session. The variation in quality of competing links is comparable for both organic and sponsored results.

Table 1: Keyword Performance Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impressions</td>
<td>72.8</td>
<td>159</td>
<td>1</td>
<td>1666</td>
</tr>
<tr>
<td>Clicks</td>
<td>1.1</td>
<td>2.2</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Orders</td>
<td>0.03</td>
<td>0.2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>AdPos</td>
<td>3.47</td>
<td>1.7</td>
<td>1</td>
<td>9.78</td>
</tr>
<tr>
<td>OrganicPos</td>
<td>17.8</td>
<td>5</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Comp_OrganicQuality</td>
<td>7</td>
<td>5.5</td>
<td>0.8</td>
<td>25</td>
</tr>
<tr>
<td>Comp_AdQuality</td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>53</td>
</tr>
<tr>
<td>LQScore</td>
<td>8</td>
<td>1.5</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Brand</td>
<td>0.6</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Specificity</td>
<td>1</td>
<td>0.7</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Bid</td>
<td>0.5</td>
<td>0.3</td>
<td>0.08</td>
<td>2</td>
</tr>
<tr>
<td>CompBid</td>
<td>0.7</td>
<td>0.4</td>
<td>0.08</td>
<td>2</td>
</tr>
</tbody>
</table>

Simultaneous Model of Clicks, Conversion and Position

Consider an advertiser placing bids for a keyword in order to ensure its ads are visible in the list of sponsored results for a query related to that keyword. The search engine uses this bid and expected ad performance to determine the ad position in the list of sponsored search ads. Consumers see the ads and decide to click on the ads, and subsequently decide whether to make a purchase. We simultaneously model consumers’ click-through and conversion behavior, and the search engine’s keyword ranking decision for the ad.

Click through Rate per Impression (CTR)

Consumer choice of selecting an advertisement can be modeled in terms of the latent utility of clicking. This depends on the position of the ad and the quality of advertiser as well as the quality of its competitors. Competing links can appear in both organic results and sponsored search results. In order to delineate the effects of these two different types of results we assume that overall quality perception of competition is additively separable into the relative search quality and relative ad quality. We represent
these in terms of the quality of the competing links in the organic results as well as the sponsored results. As mentioned earlier, we consider the cumulative effect of the log of alexa rank of the competing links to the log of alexa rank of our advertiser as a measure of the quality of competition both in organic and sponsored results. Later on we also show the results with alternate measures for the competition quality. In order to account for other factors such as the ad relevance and the expected quality of the landing page relative to that of competition, we use the quality score measure provided by the search engine. Advertiser’s own link can appear in the organic results for certain keywords. We control for that using a variable representing the actual rank of advertiser’s link. If the link is not present then we use a large number to represent the rank.

Our unit of analysis is a keyword as search engine auction is keyword specific. Keyword characteristics are an indication of the underlying search behavior which varies across consumers. For example, keyword ‘shirt’ is less specific and indicates initial stage of information search while more specific keywords like ‘levi shirt’, ‘formal blue shirt’ indicate a more advanced and directed stage of information search. To account for these differences across keywords, we capture how specific a keyword is using two different measures ‘specificity’ and ‘brand’. Specificity of a keyword is based on the number of words in the keyphrase. A keyword can also represent the national brand preference of the consumer. For example, the keyword “Levi’s jeans” would indicate that the consumer has a preference for the brand Levi’s and is further along in his search. We use a dummy variable to represent the presence of brand information. This approach of representing keyword heterogeneity is similar to the one adopted by Agarwal, Hosanagar, and Smith (2011), Ghose and Yang (2009) and Yang and Ghose (2009).

We use a hierarchical model to capture the effect of keyword characteristics. This provides a flexible random component specification that allows us to incorporate both observable and unobservable keyword-specific heterogeneity given the small number of observations for each keyword. Hierarchical models are commonly used to draw inferences on individual level characteristics (Rossi and Allenby, 2003). HB models have also recently been applied to study sponsored search data with keyword as a unit of analysis (Rutz and Bucklin 2006; Ghose and Yang 2009, Yang and Ghose 2009).

We assume an i.i.d. extreme value distribution of the error term for individual choices and use a logit model to represent the click probability for a keyword k at time t as follows

$$\Lambda_{k,t}^{\text{CTR}} = \frac{\exp(\theta_{k,t}^{\text{CTR}})}{1+\exp(\theta_{k,t}^{\text{CTR}})}$$

where $\theta_{k,t}^{\text{CTR}}$ is the latent utility of clicking. For a keyword k at time t, latent utility of clicking can be expressed as

$$(2) \quad \theta_{k,t}^{\text{CTR}} = \theta_0^k + \theta_1^k \text{AdPos}_{kt} + \theta_2^k \text{Comp}_{\text{OrganicQuality}}_{kt} + \theta_3^k \text{Comp}_{\text{AdQuality}}_{kt} + \theta_4^k \text{OrganicPos}_{kt} + \theta_5 LQScore_{kt} + \theta_6 Time_{kt} + \epsilon_{k,t}^\theta$$

$$\theta^k = \Delta^\theta z_k + u_k^\theta \quad u_k^\theta \sim N(0,V^\theta)$$

where AdPos represents the position of the ad in sponsored search results, OrganicPos is the position of the advertiser’s organic link for keyword k and time t and is the actual position when the link is present or a large number otherwise, Comp_OrganicQuality is the quality of the competition in organic results, Comp_AdQuality is the quality of the competition in sponsored results, LQScore is the quality score of the ad, $z_k$ represents keyword specific characteristics: brand and specificity. $\Delta^\theta$ is a matrix which captures the relationship between the keyword characteristics and the mean values of coefficients,
\( u_k^\theta \) represents the unobservable heterogeneity for each keyword, which we assume is normally distributed with a mean 0 and covariance matrix \( V^\theta \).

We also control for the time dynamics of the auction using a time variable Time

\( \varepsilon_k^\theta \) represents the time varying unobserved keyword attributes which are common for all consumers.

**Conversion Rate per Click (CONV)**

Assuming an i.i.d. extreme value distribution of the error term for individual choices, we can express the conversion probability as

\[
\Lambda_{kt}^{\text{CONV}} = \frac{\exp (U_{kt}^{\text{CONV}})}{1+\exp (U_{kt}^{\text{CONV}})}
\]

where \( U_{kt}^{\text{CONV}} \) is the latent utility of conversion. This may depend on the position of the ad. Similar to the click probability, the relative quality of competition in the sponsored search and organic search results can influence the conversion probability. For a keyword \( k \) at time \( t \), this latent utility can be expressed as

\[
U_{kt}^{\text{CONV}} = \beta_0^k + \beta_1^k \text{AdPos}_{kt} + \beta_2^k \text{Comp_OrganicQuality}_{kt} + \beta_3^k \text{Comp_AdQuality}_{kt} + \\
+ \beta_4^k \text{OrganicPos}_{kt} + \beta_t^k \text{Time}_{kt} + \varepsilon_{kt}^\theta
\]

\( \beta_k = \Delta^\theta z_k + u_k^\theta \quad u_k^\theta \sim N(0, V^\theta) \)

We also have controls for advertiser’s own link in organic results as well as the time dynamics.

**Ad Position**

The search engine determines the position of an advertisement for a keyword based on the product of the current bid and the quality of the advertisement relative to competing ads. As mentioned earlier, this relative quality measure is called ‘quality score’ and is available to the advertisers as the listed quality score. The dependence of ad position on bid and past performance introduces two sources of endogeneity related to the advertiser’s decision and the search engine’s decision. Advertisers can influence the position by changing their bids. In particular, advertisers might choose bids to obtain positions that yield the best performance for them. As a consequence, position is endogenously determined. Further, search engines might assign advertisers to specific positions that yield the search engine the highest revenues.

In order to correct for the resulting bias, we have to account for the advertiser’s bid choices as well as the position assigned by the search engine. In our setup, bids were randomized for the sample keywords. Thus the advertiser did not control the bids during the field experiment, taking away any strategic effect of our advertiser. Using a wide range of random bids also ensures that even if other advertisers are bidding using their own objective functions, the ads in our experiment are exposed to consumers over a wide range of positions.

Endogeneity is also introduced, because search engines use ad performance data to compute an ad’s position. In order to account for this, we explicitly model the search engine’s decision. Similar approach has been taken by Ghose and Yang (2009) and Agarwal, Hosanagar, and Smith (2011). The ad position for a keyword \( k \) at time \( t \) it can be expressed as
(5) \( \ln(AdPos_{kt}) = \gamma_0^k + \gamma_1^k \ln(bid_{k,t}) \) \( + \gamma_2^k \ln(LQscore_{k,t}) \) \( + \gamma_3^k \text{CompBid}_{kt} + \gamma_{Time} \text{Time}_{kt} \) + \( \varepsilon_{kt}^\gamma \)

with \( \gamma^k = \Delta^a z_k + u_k^\gamma \) and \( u_k^\gamma \sim N(0, V^\gamma) \)

Note that the position of the ad is the daily average position and is a continuous variable. The functional form ensures that bid and listed quality score, LQscore, are required to determine the ad position and explicitly incorporates the fact that the ad position is not randomized even if advertiser bids are random. In order to account for the effect of competition we also use the maximum competitive bid, CompBid, for each keyword which can be obtained from Google’s keyword tool.\(^3\)

Finally, as the position of the ad depends on the search engine’s decision and is endogenous, the unobservable time varying keyword attributes for the equations representing consumer decisions will be correlated with the error term for the equation representing the search engine decision. As such, we use the following distribution to account for correlation between the error terms for clickthrough rate, conversion rate, and position equations:

\[
\begin{bmatrix}
\varepsilon_{kt}^\theta \\
\varepsilon_{kt}^\beta \\
\varepsilon_{kt}^\gamma
\end{bmatrix}
\sim N(0, \Omega)
\]

where \( \Omega = \begin{bmatrix}
\Omega_{11} & \Omega_{12} & \Omega_{13} \\
\Omega_{21} & \Omega_{22} & \Omega_{23} \\
\Omega_{31} & \Omega_{32} & \Omega_{33}
\end{bmatrix} \)

**Identification**

Our identification of the effects of competition quality comes from the variation in the search results for keywords in our sample. For a keyword, the relative position of links in the organic and sponsored results can change with time due to changes to websites. However, given the short period of our experiment it is not possible for websites to strategically change webpages at a keyword level. These changes are more likely at an aggregate level which in turn would affect all the keywords. Our variable for time dynamics captures this aggregate effect. It is also possible that the Alexa rank incorporates some of the traffic resulting from sponsored search. However, we use the average value of Alexa rank for each website during the period of the experiment. Additionally, the traffic due to a specific keyword in sponsored search would have a minimal impact as advertisers have a large number of keywords in the advertising portfolio and the sponsored search traffic on an advertiser’s website is a cumulative effect of the entire keyword portfolio. Search engine maintains that the organic results are independent of the sponsored results and does not strategically adjust the organic results to maximize revenue on sponsored search.\(^4\) As a result, similar to Yang and Ghose (2009), we treat the changes to the organic results as exogenous to our setup. Similarly, as the quality score is not changing for our keywords, we do not expect the search engine to manipulate the sponsored search competing links systematically and treat these changes as exogenous.

The above set of simultaneous equations represents a triangular system with and has been addressed by authors in classical econometrics (Lahiri and Schmidt 1978, Hausman 1975, Greene 1999) and bayesian econometrics (Zellner 1962). It can be represented as follows

---

\(^3\) http://adwords.google.com/select/KeywordToolExternal

\(^4\) http://www.google.com/support/forum/p/AdWords/thread?tid=223f4519f9d645a7&hl=en


\[
U_{kt}^{CTR} = f(\text{Position}, X1, \varepsilon_{kt}^0)
\]

\[
U_{kt}^{CONV} = f(\text{Position}, X2, \varepsilon_{kt}^\beta)
\]

\[
\text{Position} = f(X3, \varepsilon_{kt}^Y)
\]

Position is endogenous while variables X1-X3 are exogenous. The identification comes from the fact that position is completely determined by exogenous variables bid and LQScore. Bid for each keyword is randomized in our setup. LQScore is a value internally calculated by the search engine for each keyword and remains stable for the short period unless the advertisers change their ads or landing pages to influence the quality score. Position in turn influences the click and conversion performance. Thus the rank and order conditions are satisfied for identification purposes (Greene, 2003). Lahiri and Schmidt (1978) have shown that the parameter estimates for a triangular system can be fully identified using GLS. Hausman (1975) shows that the likelihood function for a triangular system is the same as for seemingly unrelated regressions. Zellner (1962) has addressed triangular systems from a Bayesian point of view, and shows that the posterior probability distribution function is the same as in a seemingly unrelated regressions setting. Triangular systems have been estimated using the classical approach (Alberse and Eliashberg 2003; Godes and Mayzlin 2004) and more recently in sponsored search using the Bayesian approach (Ghose and Yang 2009, Yang and Ghose 2009).

We estimate the model using a Bayesian approach, applying Markov chain Monte Carlo sampling due to the non-linear characteristics of our model (Rossi and Allenby 2005). For a discussion of the priors and conditional posteriors of this model, please refer to the Technical Appendix A1. For the HB Models, we run the MCMC simulation for 80,000 draws and discard the first 40,000 as burn-in. In order to ensure that our parameter estimates are accurate we have simulated the clicks, orders, bids and positions using our estimates. By repeating the estimation with this simulated dataset we were able to recover our parameter estimates. This indicates our parameters are fully identified.

**Results**

**Click through rate (CTR)**

Table 2 provides the mean values for the posterior distribution of the \(\Delta^0\) matrix and the covariance matrix \(V^0\) from equation 2. The coefficient for pos is negative and significant indicating that the click performance decays with position. This is similar to the finding of the previous literature (Ghose and Yang, 2009; Yang and Ghose, 2009) and suggests that click through rate decays with the position of the ad. The coefficient for Comp_OrganicQuality and Comp_AdQuality are not significant. This suggests that majority of consumers clicking on our advertiser’s ads are not influenced by the variation in the quality of links. Yang and Ghose (2009) show that there is complementarity between an advertiser’s link in organic and sponsored search results. Our results suggest that the even if there is complementarity, it may not be having a strong adverse effect on other advertisers in sponsored search who don’t have a corresponding organic link. This is because consumers are not paying so much attention to the relative quality of the competition. This can be attributed to the search mode of the majority consumers driving the clicks. Most of these consumers are in the information seeking mode and may not care so much about the relative quality differences between links. The coefficient for OrganicPos is not significant. This is due to the fact that very few keywords in our sample have advertiser’s own link in the organic search result and the small variation in the position of organic link for our advertiser is not sufficient to influence the consumer click behavior.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Intercept</th>
<th>Brand</th>
<th>Specificity</th>
</tr>
</thead>
</table>

Table 2: Estimates for the CTR
Agarwal et. al. / Sponsored Search: Do Organic Results help or hurt?

Conversion rate (CONV)

Table 3 provides the mean values for the posterior distribution of the $\Delta \beta$ matrix and the covariance matrix $V_\beta$ in equation 4. The coefficient for pos is +ve and significant indicating that on an average conversion rate increases with position. This result is similar to the finding by Agarwal, Hosanagar, and Smith (2011) and suggests that serious buyers are visiting the lower positions more than the information seekers and are buying from these positions. Note that this holds for only top few positions which is the case for our dataset where maximum number of positions in only 10. The coefficient for Comp_OrganicQuality is +ve and significant. Our results suggest that the conversion rate decreases as the quality of competing organic links increases. This shows that the serious buyers do pay attention to the quality of organic links. The coefficient for Comp_AdQuality is not significant except for very specific keywords. This indicates that buying consumers may be paying less attention to the quality of competing links in sponsored links as compared to the quality of competing links in organic results.

Table 3: Estimates for the CONV

<table>
<thead>
<tr>
<th>Variables</th>
<th>Intercept</th>
<th>Brand</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>-1.51 (0.5)***</td>
<td>-1.27 (1.02)</td>
<td>1.69 (0.47)***</td>
</tr>
<tr>
<td>AdPos</td>
<td>0.28 (0.08)***</td>
<td>0.08 (0.15)</td>
<td>0.14 (0.09)</td>
</tr>
<tr>
<td>Comp_OrganicQuality</td>
<td>-0.19 (0.08)**</td>
<td>0.16 (0.17)</td>
<td>-0.05 (0.09)</td>
</tr>
<tr>
<td>Comp_AdQuality</td>
<td>-0.03 (0.07)</td>
<td>-0.1 (0.15)</td>
<td>-0.17 (0.09)**</td>
</tr>
<tr>
<td>OrganicPos</td>
<td>0.02 (0.07)</td>
<td>0.06 (0.15)</td>
<td>-0.05 (0.08)</td>
</tr>
<tr>
<td>Time</td>
<td>-0.03 (0.001)***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Estimates for the covariance matrix $V_\beta$

<table>
<thead>
<tr>
<th>Variables</th>
<th>Const</th>
<th>AdPos</th>
<th>Comp_OrganicQuality</th>
<th>Comp_AdQuality</th>
<th>OrganicPos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>1.58 (0.44)***</td>
<td>-0.06 (0.08)</td>
<td>-0.03 (0.08)</td>
<td>0.0 (0.07)</td>
<td>-0.04 (0.07)</td>
</tr>
<tr>
<td>AdPos</td>
<td>0.27 (0.05)***</td>
<td>0.01 (0.03)</td>
<td>-0.01 (0.03)</td>
<td>-0.01 (0.03)</td>
<td></td>
</tr>
<tr>
<td>Comp_OrganicQuality</td>
<td>0.23 (0.04)**</td>
<td>-0.01 (0.03)</td>
<td>-0.02 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp_AdQuality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OrganicPos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As mentioned earlier, the variation in the quality of competing links for each keyword is comparable for both sponsored and organic results. Thus our results suggest that for the given variation in the quality of competitor links, consumers are more responsive when they are buying. Our results also suggest that these consumers are more responsive to the quality of competing links in organic results. It is possible that a very large variation in the quality of competing links in sponsored results will lead consumers to evaluate quality of competing links in sponsored results as well.

**Ad Position**

Tables 4 provide the mean values for the posterior distribution of the $\Delta^\gamma$ matrix and $V^\gamma$ from equation 5. In these results, higher bids lead to higher current position. Similarly higher LQscore leads to higher current position. This is reasonable as both bid and LQscore are the primary inputs used to compute the ad position and higher values of these should move the ad higher.
endogenous and the proposed simultaneous equation model helps to capture the effect of this endogeneity.

Table 5: Estimates for the Covariance Matrix \( \Omega \)

<table>
<thead>
<tr>
<th></th>
<th>CONV</th>
<th>CTR</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONV</td>
<td>0.44 (0.05)**</td>
<td>-0.05 (0.01)**</td>
<td>-0.02 (0.01)**</td>
</tr>
<tr>
<td>CTR</td>
<td>0.21 (0.02)**</td>
<td>0.02 (0.01)**</td>
<td>0.09 (0.0)**</td>
</tr>
<tr>
<td>Rank</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Robustness of Results

In this section we outline several steps we have taken to evaluate the robustness of these results.

Holdout Sample Analysis

As one test of robustness, we have attempted to verify the prediction accuracy of our results using a holdout sample. To do this, we consider data for the first 4 weeks as the estimation sample and the data for the remaining two weeks as the holdout sample. We use mean absolute percentage error (MAPE) for daily CTR and CONV values at the aggregate level and at the keyword level. The error values are reported in Table 6 and indicate that the model prediction accuracy is similar for both the estimation and holdout samples. This suggests that our model estimates are robust.

Table 6: Prediction Accuracy for Estimation & Holdout Samples

<table>
<thead>
<tr>
<th>Models</th>
<th>CTR Fit (MAPE)</th>
<th>CONV Fit (MAPE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aggregate</td>
<td>Keyword</td>
</tr>
<tr>
<td>Estimation Sample</td>
<td>0.41</td>
<td>0.40</td>
</tr>
<tr>
<td>Holdout Sample</td>
<td>0.43</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Aggregate MAPE is the average MAPE across all datapoints. Keyword MAPE is the average of the average MAPE for different keywords.

Alternate Measure of Quality

We have verified our results with alternate measures of quality perception of our advertiser by comparing the number of links across different competing websites both in sponsored search and organic results. The number of links is a key input used by search engines to evaluate the relative ranking of websites, and has been used by previous research (e.g. Animesh et al, 2010). A higher number of links is an indication of higher popularity of a website. In order to determine the relative quality perception of a website we consider the sum of the ratio of the log value of the number of links of the competitor and the log value of the number of links of our advertiser. We calculate perceived quality separately for organic results and sponsored search results. Results are shown in Tables 7-8. The qualitative results are similar to our main analysis.

Table 7: Parameter estimates for CTR using a different measure of quality

<table>
<thead>
<tr>
<th>Variables</th>
<th>Intercept</th>
<th>Brand</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>-1.35 (0.7)**</td>
<td>-1.97 (0.96)**</td>
<td>-0.22 (0.5)</td>
</tr>
<tr>
<td>AdPos</td>
<td>-0.4 (0.08)**</td>
<td>-0.07 (0.16)</td>
<td>0.05 (0.09)</td>
</tr>
</tbody>
</table>
### Table 8: Parameter estimates for CONV using a different measure of quality

<table>
<thead>
<tr>
<th>Variables</th>
<th>Intercept</th>
<th>Brand</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>-2.66 (0.71)***</td>
<td>4.7 (1.35)***</td>
<td>0.26 (0.62)</td>
</tr>
<tr>
<td>AdPos</td>
<td>0.42 (0.1)***</td>
<td>0.19 (0.16)</td>
<td>-0.16 (0.09)*</td>
</tr>
<tr>
<td>Comp_OrganicQuality</td>
<td>-0.19 (0.09)**</td>
<td>-0.24 (0.18)</td>
<td>0.03 (0.12)</td>
</tr>
<tr>
<td>Comp_AdQuality</td>
<td>0.08 (0.08)</td>
<td>-0.21 (0.17)</td>
<td>-0.16 (0.09)*</td>
</tr>
<tr>
<td>OrganicPos</td>
<td>0.01 (0.07)</td>
<td>-0.12 (0.15)</td>
<td>0.04 (0.08)</td>
</tr>
<tr>
<td>LQScore</td>
<td>-0.02 (0.0)***</td>
<td>-0.04 (0.13)</td>
<td>-0.03 (0.1)</td>
</tr>
<tr>
<td>Time</td>
<td>-2.66 (0.71)***</td>
<td>4.7 (1.35)***</td>
<td>0.26 (0.62)</td>
</tr>
</tbody>
</table>

### Estimates for the covariance matrix \( V^\theta \)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Const</th>
<th>AdPos</th>
<th>Comp_OrganicQuality</th>
<th>Comp_AdQuality</th>
<th>OrganicPos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>1.54 (0.41)***</td>
<td>0.03 (0.08)</td>
<td>-0.01 (0.11)</td>
<td>-0.07 (0.09)</td>
<td>-0.06 (0.07)</td>
</tr>
<tr>
<td>AdPos</td>
<td>0.29 (0.05)***</td>
<td>-0.02 (0.04)</td>
<td></td>
<td>-0.01 (0.03)</td>
<td>-0.02 (0.03)</td>
</tr>
<tr>
<td>Comp_OrganicQuality</td>
<td></td>
<td>0.34 (0.08)***</td>
<td></td>
<td>0.0 (0.04)</td>
<td>-0.01 (0.04)</td>
</tr>
<tr>
<td>Comp_AdQuality</td>
<td></td>
<td></td>
<td></td>
<td>0.29 (0.05)***</td>
<td>-0.01 (0.03)</td>
</tr>
<tr>
<td>OrganicPos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.16 (0.04)***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Const</th>
<th>AdPos</th>
<th>Comp_OrganicQuality</th>
<th>Comp_AdQuality</th>
<th>OrganicPos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>1.37 (0.96)</td>
<td>0.01 (0.11)</td>
<td>-0.04 (0.13)</td>
<td>-0.03 (0.1)</td>
<td>-0.03 (0.1)</td>
</tr>
<tr>
<td>AdPos</td>
<td>0.3 (0.05)***</td>
<td>0.03 (0.04)</td>
<td></td>
<td>-0.04 (0.03)</td>
<td>-0.01 (0.03)</td>
</tr>
<tr>
<td>Comp_OrganicQuality</td>
<td>0.34 (0.08)***</td>
<td>-0.01 (0.04)</td>
<td></td>
<td>0.25 (0.05)***</td>
<td>-0.02 (0.03)</td>
</tr>
<tr>
<td>Comp_AdQuality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.19 (0.04)***</td>
</tr>
</tbody>
</table>
Discussion and Conclusion

In this paper, we analyze the impact of competing links in organic and sponsored results on the performance of sponsored search advertisements. Sponsored search is viewed as an effective ad mechanism to reach out to the consumers. However, the effect of competing links on the performance of these ads is not very well understood. This is especially important as consumers can see two different lists simultaneously. We analyze the impact of competing links in organic results on sponsored search performance using a unique dataset generated from a field experiment on an online retailer’s ad campaign on Google. This dataset documents the daily impressions, clicks, orders, and costs for a select sample of keywords in the ad campaign for different positions for the corresponding ads. We also use the data from a web crawler which records the search results during the period of our field experiment. Our results also show that the presence of competitor’s links has a higher negative effect on the conversion rate as compared to the click through rate. We also find that competing organic links have a larger effect on conversion rate as compared to the competing sponsored links.

These findings are important to the industry as advertisers are not clear about the impact of organic results on sponsored search. This is especially important for advertisers who may not be able to appear in the high positions for organic results. Even large advertisers may face this as they may not show up in the organic results for certain keywords. For example, Honda website shows up in the top organic results for ‘suv car’ but not for ‘sedan car’. Our results suggest that the advertisers are not as impacted in their branding effort which primarily relies on clicks when competing links are present in organic results. However, they may not get the transactional benefits as the rate of conversion for sponsored search decreases due to the presence of competition in organic links. Thus, depending on their objective for sponsored search, advertisers will have to respond differently to the presence of competition in organic results.

Our result suggests that there may be inefficiencies in the current mechanisms used by search engines. The mechanism assumes that the organic results do not influence sponsored search results. However, our results show that organic results impose externality on sponsored search results. While presence of competition in organic results does not influence the clicks, it has an effect on conversions. Thus, advertisers interested in transactional revenue are getting adversely affected. They end up paying the search engine for the clicks but do not get the appropriate benefit in terms of conversions. This can reduce advertiser participation. Thus, the search engine mechanism should account for the impact of organic links on conversions in its ranking mechanism.

Finally, our study sheds light on consumer behavior in sponsored search environments. Click through rate is not influenced by the quality of competition in organic links. However, conversion rate is influenced by the quality of competition in organic links. This suggests consumer behavior is driven by their mode of purchase. Consumers in the information seeking mode do not pay attention to the quality of competition. Consumers in the buying mode are more careful about the selection of ads and seem to rely on organic links for their decision making.

As with any empirical analysis there are several limitations of our study. While, our results explain some information search behavior of consumers at an aggregate level, the aggregate nature of our data limits our ability to account for the actions of individual consumers. This calls for future research using click stream data to empirically evaluate the behavior of different types of consumers in sponsored search. An additional limitation is that our analysis of orders is based on measurements conducted by the SEM firm wherein consumer action is tracked during the entire search session. This is potentially problematic because, consumers may click on an ad, visit the advertiser’s landing page without converting but return on a later day (even using a different search engine query) to then buy the product. In these instances, the future purchases are not properly attributed to the original keyword. While we are able to evaluate the impact of organic search results, we do not have data to evaluate the reverse effect i.e. the impact of sponsored search on organic results. Yang and Ghose (2009) show that the clicking propensity for sponsored results increases the clicking propensity for organic results. However, they do not explicitly study the conversion performance of organic results. Future research should investigate the combined effect of both organic and search results on the overall performance and suggest strategies to optimize the extent of advertiser participation in sponsored search.
References


