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

Ashish Agarwal
Assistant Professor
McCombs School of Business
University of Texas, Austin
Austin, TX 15213
Phone: (512) 471 5814
Email: 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
Phone: (215) 573 0831
Email: kartikh@wharton.upenn.edu

Michael D. Smith
Associate Professor of Information Technology and Marketing
Heinz College, School of Information Systems and Management, and
Tepper School of Business
Carnegie Mellon University
4800 Forbes Avenue
Pittsburgh, PA 15213
Phone: (412) 268 5978
Email: mds@cmu.edu

ABSTRACT

Sponsored search accounts for 40% of the total online advertising market. These ads appear as ordered lists along with the organic (regular) search results in search engine results pages. There is conflicting evidence regarding the impact of organic search results on the sponsored search especially when there is an overlap in the results.

We evaluate the impact of advertiser's link as well as the competing links in organic results on the its sponsored search performance using 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 organic results on both click-through rate and conversion rate for these keywords. We find that while the advertiser's link in organic result has a positive impact on its click through rate, it has a negative impact on its conversion rate. We find that conversion rate also depends on the quality perception of the advertiser as compared to its competitors' links in organic results. The conversion rate is higher where the advertiser is perceived to be of higher quality among the competing organic links.

Our results inform the advertising strategies of firms participating in sponsored search auctions and provide insight into consumer behavior in these environments. Specifically, we show that organic links may hurt performance for immediate transactions. However, as the click through rate increases, it helps in increasing the brand awareness for the advertiser.

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, an advertiser can show up in both organic and sponsored search results. Additionally, an advertiser's competitors can also appear in organic results. An important question is how the collocation of links in organic results and sponsored results impact the sponsored search performance.

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 searching using 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 less likely to be influenced by advertisers. In that case, 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 shown that consumers tend to focus more on organic results as compared to the sponsored search results.[2]

---

[1] http://www.google.com/support/forum/p/AdWords/thread?tid=223f4519fdb645a7&hl=en
[2] http://www.webcitation.org/5FmwyPgDv

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. However, they evaluate the outcome using weekly aggregate data that does not allow them to analyze the underlying drivers of this complementarity. Addressing this gap, we seek to analyze what is the effect of competing organic links on sponsored search performance, and what is the impact of advertiser's organic link on its sponsored search 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 organic links do influence the performance of sponsored search ads. The effect is different for clicks and conversions. Presence of advertiser's link in organic results improves the clickthrough rate for its sponsored search ads. However, it has a negative effect on the conversion rate for such ads. We also find that organic links of competing firms can hurt the conversion performance if they are perceived to be of higher quality. We also find that the competing links do not impose any externality on the click through rate.

Our paper makes several contributions. First, our paper provides key managerial insights for advertisers. Popular websites already appear in organic results and need to justify their participation in the sponsored search. Our results show that the firm can only gain benefit in

terms of branding achieved through viewing and clicking sponsored search ads. However, they cannot get full benefit of immediate transactions from the sponsored search ads potentially due to cannibalization effect of their own organic links. Advertisers also have to pay attention to the competition from other links in organic results which can negatively impact their conversion performance.

Our results also suggest that there may be inefficiencies in the current mechanisms used by search engines. Current auction mechanism is advertised to be independent of organic results and a tool to reach out to consumers. However, our results show that organic results impose externality on sponsored search results. Familiar brands are likely to appear in organic results and be perceived as higher quality advertisers. As a consequence, relatively unknown advertisers are at a disadvantage as their ROI can be significantly lower as compared to familiar brands even if they appear at higher positions in sponsored search results.

Finally, we also provide insight into consumer behavior in sponsored search environments. Our results suggest that the presence of a firm's organic link is viewed positively by the consumers. This increases the clickthrough rate for its sponsored link. However, the buying consumers tend to ignore the sponsored link when the links co-exist which is evident from a decrease in the conversion rate. One possibility is that these consumers maybe instead selecting the organic link for making their purchase decision. We also find that the quality perception of the advertiser as compared to firms appearing in organic results is important for the buying consumers. This again indicates that the buying consumers pay attention to the organic results.

We organize the rest of the article as follows. We begin with a description of the sponsored search market and describe prior academic work in this area. Next, we describe our field experiment and the resulting dataset. We then develop a model to measure the performance of

sponsored search ads and describe the estimation approach and discuss the empirical results.
Finally we conclude the study, and discuss limitations and areas for future research.

SPONSORED SEARCH BACKGROUND

When a consumer enters a search query, for example "shirts," the search engine displays algorithmic (i.e. regular), and sponsored search results as shown in Figure 1. The algorithmic results are determined based on their relevance to the query. The sponsored results are ranked based on continuous real-time auctions run by the search engines. Advertisers bid on sponsored search keywords of relevance to them. Upon receiving a query, the search engine identifies the advertisers bidding on closely related keywords, and uses data on bids and past click performance of ads to rank order the ads that appear in the list of sponsored results.

An advertiser pays the search engine only when the consumer clicks on the advertiser's ad. The cost per click (CPC) is determined using a generalized second price auction mechanism; i.e. whenever a user clicks on an ad in position k, the advertiser pays an amount equal to the minimum bid needed to secure that position (Lahaie and Pennock, 2007). After clicking on the ad, the consumer is redirected to the advertiser's website, and then chooses whether to purchase a product or register for a service (which we define as an order).

The search engines provide daily reports to advertisers on the status of their campaigns. These reports provide statistics on the number of impressions, clicks, and the average position for each keyword in the advertiser's portfolio. The continuous nature of the auction allows an advertiser to change the portfolio of keywords as well as the bids, ad copies, and landing pages for each keyword in real-time. The bid submitted by the advertiser implicitly determines the target position for the ad. These decisions ultimately drive the advertiser's Return on Ad Spend (RoAS), a key metric used to evaluate return on investments in advertising. In our study, we

focus on the impact of the organic results on the click and conversion performance of sponsored search results. The ad copies and landing pages associated with these keywords do not change over time for the advertiser under consideration.

= = Insert Figure 1 about here = =

LITERATURE REVIEW

Prior Work

The literature most relevant to our study includes past research on consumers' online search behavior, with a special emphasis on the impact 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. As the advertisers are ranked in a particular order, the advertiser position may influence its quality perception. 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. Search engines can also be viewed as tools that aid consumer decision-making. Haubl and Trifts (2000) find that the use of decision aids reduces the size of consumers' consideration set but improves the quality of their consideration sets and the quality of purchase decisions in the online shopping environment. This again suggests that consumers are likely to evaluate only a few sponsored search results as they might expect that the results are in the decreasing order of relevance. There is some evidence for this. 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. This would suggest that position of an advertiser plays an important role in forming a perception about the quality of the advertiser. An important question, would be whether the consumers pay attention to the relative importance of an advertiser above and beyond the position of the ad. 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. 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 influence by the quality perception of links appearing in sponsored search.

Click and conversion performance also 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. Broad pattern of search may reflect less sensitivity to the organic results. On the other hand, focused search on part of buying consumers may reflect more selectivity in clicking ads. Advertiser revenues depend not just on clicks but also on conversion probability. Some studies show that consumers tend to de-emphasize the prescreening information in their search process

(Diel, Kornish and Lynch, 2003; Chakravarti et. al 2006). This suggests that the criteria used for selecting an ad may not have an effect on the final order as compared to the information obtained after visiting the associated website. Thus, if a consumer discounts all pre-screening information and buys from the website that maximizes his utility, conversion rate may be independent of the sponsored links

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 Zhang (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 sponsored 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. However, Ghose and Yang (2009) find complementarity between clicks on organic results and clicks on sponsored results. A relevant question is what drives this complementarity.

Thus, the prior literature show that competing ads can influence the performance of ads. However, the impact of organic results, such co-existence of advertiser's link and the relative

search quality, on the performance of sponsored search is not known. Also, there is conflicting evidence regarding the influence of sponsored links on the performance of sponsored search.

FIELD EXPERIMENT & 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. The data were generated by submitting randomized bids for several keywords that had generated orders in the past, and measuring the consumer response in terms of clicks and orders for different positions of the ads corresponding to the keywords.[3] 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 along with the rank in the organic results for our online retailer. Following the literature, we use alexa rank obtained from alexa.com as a measure of the perceived quality of a firm appearing in organic and sponsored 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). In order to determine the quality perception we consider the difference in the log values of the alexa rank of our advertiser and the top competing link. We calculate the perceived quality separately for organic results and sponsored search results. This is to account for the differential

---

[3] We used Google's keyword tool to determine the range of bids for each keyword for each position.

impact of the two types of results. It is possible that for a keyword, there is high correlation in the competitive intensity in organic and sponsored search results. In order to reduced the impact of this correlation we mean center the perceived quality measures using keyword level mean values.

The data set consists of 1468 observations of daily impressions, clicks, and orders for 39 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. In addition, the position reported for any keyword is the average position on a given day. The position of an ad can vary within a day because the set of advertisers may be different for different queries associated with a keyword. For example, the ad for keyword "red dress" may be shown if the consumer types "red party dress" or "red dress". The competitors and their bids may be different for these two queries causing the position to vary. By focusing on keywords with "exact match", we eliminate a major source of such intra-day variation in position. An ad for a keyword with "exact match" is shown only if the query is exactly the same as that keyword and hence the competitors are fixed for such keywords. Another reason for the position to vary is that competitors may change their bids multiple times within a day. While firms change bid periodically, typically weekly and sometimes even daily, we do not find significant intra-day variation in ad position for keywords with exact match.[4]

= = Insert Table 1 about here = =

---

[4] We have separately verified this for our sample keywords by monitoring the relative ad positions across multiple queries in a day. For the keywords with a phrase match we have used a large set of queries that have been associated with these keywords in the past several months.

SIMULTANEOUS MODEL OF CLICKS, CONVERSIONS, AND AD 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 expected advertiser quality. Competing links can appear in both organic results and sponsored search results. Quality perception can be influenced by the relative quality perception in the sponsored results as well as organic results. In order to delineate these effects we assume that overall quality perception is additively separable into the relative search quality and relative ad quality. Each of these variables can be positive or negative depending on whether the advertiser is perceived to be of higher quality or lower quality. Advertisers' own link can also appear in organic results. Thus, apart from the quality perception, the presence of advertiser's link can also influence consumer's clicking decision.

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 nearness of its landing page to a product. Advertisers organize their websites in a hierarchical fashion to accommodate the search intent of the users and reduce their search cost. Various levels in the hierarchy represent the product categories, sub-categories and products. When consumers are routed through a search engine, the landing page coincides with a level in the website hierarchy that is chosen based on the search intent of the consumer as reflected in the keyword. We define specificity as the level in the product hierarchy of the advertiser. For example, a top level such as 'men's clothing' would have the specificity value of 0, second level such as 'shirts' would have the specificity value of 1, and so on.

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 et. al (2010), 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).

For a keyword k at time t, latent utility of clicking can be expressed as

$$U_{kt}^{CTR} = Y_{kt}^{CTR} + \varepsilon_{kt}^{\theta}$$

(1)

$$Y_{kt}^{CTR} = \theta_0^k + \theta_1^k Pos_{kt} + \theta_2 \delta_{kt}^{Organic} + \theta_2 SearchQuality_{kt} + \theta_2 AdQuality_{kt} + \theta_t Time_{kt}$$

$$\theta^k = \Delta^\theta z_k + u_k^\theta \qquad u_k^\theta \sim N(0, V^\theta) \text{ where } \theta^k = [\theta_0^k, \theta_1^k]$$

Pos represents the position of the ad in sponsored search results

$\delta_{kt}^{Organic}$ is the dummy for the advertiser's organic link for keyword k and time t and is 1 when the link is present, 0 otherwise

$SearchQuality$ is the perceived quality of the advertiser among organic links

$AdQuality$ is the perceived quality of advertiser among sponsored results

z_k represents keyword specific characteristics: brand and specificity. Δ^θ is a matrix which captures the relationship between the keyword characteristics and the mean values of coefficients.

u_k^β represents the unobservable heterogeneity for each keyword, which we assume is normally distributed with a mean 0 and covariance matrix V^β.

We also control for the time dynamics of the auction using a time variable $Time_{kt}$.

We use a logit model to represent the click probability for a keyword k at time t as follows

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

(2)

Conversion Rate per Click (CONV)

Consumer choice of purchasing from an ad website after clicking the ad can be modeled in terms of the latent utility of conversion. This may depend on the position of the ad. For a keyword k at time t, this latent utility can be expressed as

$$U_{kt}^{CONV} = Y_{kt}^{CONV} + \varepsilon_{kt}^\beta$$

(3)

where $Y_{kt}^{CONV} = \beta_0^k + \beta_1^k Pos_{kt} + \beta_2 \delta_{kt}^{Organic} + \beta_3 SearchQuality_{kt} + \beta_4 AdQuality_{kt} +$

$$\beta_t Time_{kt}$$

$\beta^k = \Delta^\beta z_k + u_k^\beta \qquad u_k^\beta \sim N(0, V^\beta)$ where $\beta^k = [\beta_0^k, \beta_1^k]$

Similar to the CTR model, we have controls for time and a constant term. The conversion

probability can be expressed as follows

$$\Lambda_{kt}^{CONV} = \frac{exp(Y_{kt}^{CONV})}{1 + exp(Y_{kt}^{CONV})}$$

(4)

Ghose and Yang (2009) and Yang and Ghose (2009) have used similar models to represent

aggregate demand for clicks and orders.

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 the competing ads. This

relative quality measure is called quality score and is available to the advertisers as listed quality

score.[5] The dependence of ad position on bid and past performance introduces two sources of

endogeneity related to the advertiser's decision and 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 and

---

[5] http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=6111
https://adwords.google.com/support/aw/bin/answer.py?answer=100305
https://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=115967

this takes away the 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 get exposed to consumers over a wide range of positions.

Because search engines use ad performance data to compute an ad's position, an ad's position is still endogenous. In order to explicitly account for this, we also model the search engine's decision. The ad position for a keyword k at time t it can be expressed as

$$pos_{kt} \propto \delta_0^k \left(bid_{k,t} \right)^{\delta_1^k} \left(LQscore_{\ k,t} \right)^{\delta_2}$$

(5)

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 rank 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 the Google's keyword tool. [6]Substituting, taking the log, and using controls for time we get

$$ln(pos_{kt}) = \alpha_0^k + \alpha_1^k ln\left(bid_{k,t} \right) + \alpha_2 ln\left(LQscore_{k,t} \right) + \alpha_3 CompBid_{kt} + \alpha_t Time_{kt} + \varepsilon_{kt}^\alpha$$

(6)

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

Finally, as the position of the ad depends on the search engine's decision and is endogenous, the error terms for the equations representing consumer decisions will be correlated with the error term for the equation representing the search engine decision. In order to account for the correlation between the error terms for clickthrough rate, conversion rate, and position equations we use the following distribution

---

[6] https://adwords.google.com/select/KeywordToolExternal

$$\begin{bmatrix} \varepsilon_{kt}^{\theta} \\ \varepsilon_{kt}^{\beta} \\ \varepsilon_{kt}^{\alpha} \end{bmatrix} \sim N(0, \Omega) \text{ where } \Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix}$$

(7)

Identification

Our identification of the effects of organic results comes from the variation in the search results for all but one keyword in our sample. We treat these changes as exogenous to the system. It is possible that advertisers can be changing their websites which results in the changes in the search results. However, given the short period of our experiment it is not possible for advertisers 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, the alexa rank is based on the consumer traffic to the entire site. Traffic due to sponsored search would be a fraction of this traffic. 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.

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

$$U_{kt}^{CTR} = f(Position, X1, \varepsilon_{kt}^{\theta})$$

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

$$Position = f(X3, \varepsilon_{kt}^{\alpha})$$

Position is endogenous while variables X1-X3 are exogenous. The identification comes from the fact that rank 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. Rank 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 Δ^θ matrix and other parameters from equation 1. The coefficient for *pos* is negative and significant indicating that the click performance decays with position. The coefficient for organic link dummy is positive and significant suggesting that the the overall clickthrough rate improves due to the presence of advertiser's link in organic results. This is similar to the finding by Yang and Ghose (2010). The coefficients for AdQuality and SearchQuality are not significant. This suggests that overall consumers do not pay attention to the familiarity of the link while clicking during the sequential evaluation. Table 3 reports the values for the covariance matrix V^θ and highlights the importance of accounting for heterogeneity.

$$= = \text{Insert Table 2 \& 3 here} = =$$

Conversion Rate (CONV)

Table 4 provides the mean values for the posterior distribution of the Δ^β matrix and other parametersfrom equation 3. The coefficient for pos is +ve and significant indicating that on an average conversion rate increases with position. This suggests that the serious buyers are actually visiting the lower positions more than the information seekers and are buying from these positions. The coefficient for organic link is –ve and significant. This suggests that the buying consumers are less likely to convert using the sponsored search link of the advertiser when it also appears in the organic results. A possible explanation is that these consumers are visiting the organic link instead of clicking on the advertiser's link. The coefficient for AdQuality is positive and significant. This suggests that the buying consumers are paying attention to relative quality perception of the advertiser as compared to the competitors appearing in organic search results.

Table 5 reports the values for the covariance matrix. The values are significant, which again underscores the importance of accounting for heterogeneity in this environment.

= = Insert Tables 4 & 5 about here = =

Ad Position

Tables 6 and 7 provide the mean values for the posterior distribution of the Δ^α matrix and V^α from equation 6. 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 rank and higher values of these should move the ad higher.

= = Insert Tables 6 & 7 about here = =

Finally, Table 8 shows covariance between unobservables for CTR, CONV, and ad positions from equation 8. Covariance between the unobservables for CONV and CTR is significant. This indicates that the unknown factors influencing consumer clicks also influence the subsequent conversion behavior. The covariance between the unobservables for CONV and position is statistically significant. Similarly, covariance between the unobservables for CTR and position is statistically significant. This correlation between the error terms for CONV and CTR with the error term for ad position shows that the position is endogenous and the proposed simultaneous equation model helps to capture the effect of this endogeneity.

= = Insert Table 8 about here = =

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 the keyword level. The error values are reported in Table 9 and indicate that the model prediction accuracy is similar for both the estimation and holdout samples. This suggests that our model estimates are robust.

= = Insert Table 9 about here = =

Alternate Measure of Quality: We have verified our results with alternate measure of quality perception of our advertiser which is based on comparing the no of links across different competing web sites both in sponsored search results and organic results. No of links is a key input used by the search engines to evaluate the relative ranking of websites. This measure of quality has also been used by previous research (Animesh et al, 2010). Higher number of links would be an indication of higher popularity of a website. In order to determine the relative quality perception we consider the difference in the log values of the no of links of the top competing link and our advertiser. We calculate the perceived quality separately for organic results and sponsored search results. It is possible that for a keyword, there is high correlation in the competitive intensity in organic and sponsored search results. In order to reduced the impact of this correlation we mean center the perceived quality measures using keyword level mean values. Results are shown in Table 10-13. The qualitative results are similar to our main analysis. Note that higher value of quality indicates that our advertiser less popular than the top competing link.

DISCUSSION AND CONCLUSION

In this paper, we analyze the impact organic results on the performance of sponsored search advertisements that appear alongside regular organic search results in search engines. Sponsored search is viewed as an effective ad mechanism to reach out to the consumers.

We analyze the impact of advertiser's own link as well competing link 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 advertiser's link has different effect on the click through rate and the conversion rate. While the click through rate increases due to the presence of advertiser's link in organic results, the conversion rate decreases. Our results also show that the relative quality perception of the advertiser in the organic search results is important for the conversion rate.

These findings are important to the industry as advertisers are not clear about the value proposition of sponsored search when they can easily show up in the organic results. Our results suggest that the advertisers can improve their branding by the collocation of the links as the clicking propensity increases. However, they may not get the transactional benefits as the rate of conversion for sponsored search decreases due to the collocation of links. Our results also suggest that advertisers should pay attention to the competition from the organic links. If they are perceived to be of lower quality then they should aim for higher positions in sponsored search in order to increase their click through rate. Also they should improve their website content which would improve their organic ranking and lead to higher quality perception..

Our result suggests that there may be inefficiencies in the current mechanisms used by search engines. A common perception is that sponsored search is a mechanism for the small advertisers to reach out to consumers. However, our results show that organic results impose externality on sponsored search results. Familiar brands are likely to appear in organic results and be perceived as higher quality advertisers. As a consequence, relatively unknown advertisers are at a

disadvantage as their ROI can be significantly lower as compared to familiar brands even if they appear at higher positions in sponsored search results.

Finally, our study sheds light on consumer behavior in sponsored search environments. Conversion rate depends on the relative quality perception of the advertiser in organic search results. It also decreases if the advertiser's link is present in organic results. This suggests consumer behavior is driven by their type. 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. We only consider the relative quality of the advertiser determined by the popularity of the advertiser's website. However, the clicking and conversion behavior can also depend on the content of the links as well as the landing page. While we do not consider these factors, these are not likely to change in our panel duration. The differences at the keyword level are accounted by the random effects in our model. Future research should investigate the impact of content as well. For example, Animesh et al (2009) have shown that the unique selling proposition of the ad content influences its click through rate. One possibility is to extend this analysis to evaluate the conversion rate as well as the impact of the textual content of organic results on the sponsored search performance. 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. Our results show that the conversion rate decreases when the advertiser link appears in organic result. However, if this in turn increases the conversion rate for organic results then the investment in sponsored search can still be justified. Yang and Ghose (2009) show that the clicking propensity for sponsored results increases the clicking propensity for organic results. However, the magnitude of this effect is lower as compared to the effect of organic results on sponsored search. 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

Agarwal, Ashish, Kartik Hosanagar, Michael Smith. 2010. Location, location and location: An analysis of profitability of position in online advertising markets. Working Paper, SSRN.

Animesh, A., V. Ramachandran, and S. Viswanathan. (2010). Quality Uncertainty and Adverse Selection in Online Sponsored Search Markets. Information Systems Research, 21, 190-201.

Animesh, Viswanathan, and Agarwal (2009): Competing "Creatively" in Sponsored Search Markets Information Systems Research, Articles in Advance, pp. 1–20.

Athey, Susan, Glenn Ellison. 2008. Position auctions with consumer search, Working Paper, Harvard University.

Ansari, A., C. Mela. 2003. E-customization. *Journal of Marketing Research*, 40(2), 131–145.

Arbatskaya, M. (2007). Ordered Search. *Rand Journal of Economics*, 38(1), 119-126.

Brooks, N. (2004). The atlas rank report II: How search engine rank impacts conversions. Atlas Institute, New York.

Brunel, F. F., Nelson, M. R. (2003). Message Order Effects and Gender Differences in Advertising Persuasion. *Journal of Advertising Research*, 43(3), 330-341.

Brynjolfsson, E., A.A. Dick, M.D. Smith. (2009). A Nearly Perfect Market? Differentiation vs. Price in Consumer Choice. *Quantitative Marketing and Economics*, **8**(1), 1-33.

Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A., Mathieu, C., and Schwarz, M. (2007). Greedy Bidding Strategies for Keyword Auctions. *Proceedings of the 8th ACM Conference on Electronic Commerce*, San Diego,CA.

Chakravarti, A., Janiszewski, C., and Ulkumen, G. (2006). The Neglect of Prescreening Information. *Journal of Marketing Research*, 63(11), 642-653.

Demers, E., and Lev, B. "A Rude Awakening: Internet Shakeout in 2000," Review of Accounting Studies (6:2/3) 2001, pp. 331-359.

Diehl, K., Kornish, L. J., and Lynch J, G. (2003). Smart Agents: When Lower Search Cost for Quality Information Increases Price Sensitivity. *Journal of Consumer Research*, 30 (6), 56–71.

Edelman, B., Ostrovsky, M. (2007). Strategic Bidder Behavior in Sponsored Search Auctions. *Decision Support Systems*, 43, 192-198.

Edelman, B., Ostrovsky, M. & Schwarz, M. (2007). Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords, *American Economic Review*, forthcoming March 2007.

Elberse, A., Eliashberg, J. (2003) Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. *Marketing Science*, 22(3), 329-354.

EMarketer (2009). US Online Ad Spend Turns the Corner. *www.emarketer.com*, December.

Feldman, J., Muthukrishnan, S., Pal, M., and Stein, C. (2007). Budget optimization in search-based advertising auctions. *Proceedings of the 8th ACM Conference on Electronic Commerce*, San Diego,CA.

Feng, F., Bhargava, H., and Pennock, D. 2007. Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms. *Informs Journal on Computing*, 19(1),137-148.

Ganchev, K., Kulesza, A., Tan, J., Gabbard, R., Liu, Q., Kearns, M. (2007). Empirical Price Modeling for Sponsored Search. Working Paper.

Godes, D., and Mayzlin, D. (2004) Using Online Conversations to Study Word of Mouth Communication. *Marketing Science*, 23 (4), 545-560.

Granka, Laura A., Thorsten Joachims, and Geri Gay (2004), \Eye-Tracking Analysis of User Behavior in WWW Search," Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, July 25-29, Sheffield, UK.

Ghose, A. and Yang, S. (2009). An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets. *Management Science,* 55(10), 1605-1622.

Greene, W. (1999). Econometric Analysis, 4th ed. Prentice Hall, Upper Saddle River, NJ.

Häubl, G. and Trifts, V. (2000). Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids. *Marketing Science*, 19 (Winter), 4–21.

Hausman, J. (1975). An instrumental variable approach to full information estimators for linear and certain nonlinear econometric models. *Econometrica*, 43, 727–738.

Hoque, A.Y., and Lohse, G.L (1999) An Information Search Cost Perspective for Designing Interfaces for Electronic Commerce. *Journal of Marketing Research*, 36(3), 387-394.

Jansen, B. J. and Resnick, M. (2006). An Examination of Searchers' Perceptions of Nonsponsored and Sponsored Links During Ecommerce Web Searching J. Am. Soc. for Information Science and Technology, pp. 1949-1961.

Johnson, E. J., Moe, W. W., Fader, P. S., Bellman, S., Lohse, G.L. 2004. On the depth and dynamics of online search behavior. *Management Science*, 50(3), 299-308.

Kent, J. R. and Allen, Chris T (1994). Competitive Interference Effects in Consumer Memory for Advertising: The Role of Brand. The Journal of Marketing, Vol. 58, No. 3 (Jul., 1994), pp. 97-105.

Lahaie, S.; Pennock, D. M. (2007). Revenue Analysis of a Family of Ranking Rules for Keyword Auctions. *Proc. of the 8th ACM Conference on Electronic Commerce*, San Diego, CA

Lahiri, K. and   Schmidt, P. (1978). On the estimation of triangular structural systems. E*conometrica,* 46(5) 1217–1221.

Lizhen Xu, Jianqing Chen and Andrew Whinston. 2009. "Too Organic for Organic Listing? Interplay between Organic and Sponsored Listing in Search Advertising" http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1409450

Lohse, G. L. (1997). Consumer Eye Movement Patterns on Yellow Pages Advertising. *Journal of Advertising*, 26(1), 61-73.

Katona, Z., and M. Sarvary (2008), "The Race for Sponsored Links: Bidding Patterns for Search Advertising," Working paper, Haas School of Business, University of California, Berkeley.

Moe, W. W. 2003. Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. J. Consumer Psych. 13(1, 2) 29–40.

Moe, W. W. and Fader, P. S. (2004). Dynamic Conversion Behavior at e-Commerce Sites. *Management Science*, 50 (3), 326-335.

Montgomery, A. L., Hosanagar,K., Krishnan, R. and Clay. K. B. (2004). Designing a Better Shopbot. *Management Science*, 50(2), 189–206.

Montgomery, A. L., Li, S., Srinivasan, K., and Liechty, J. C. (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data", *Marketing Science*, 23(4), 579-595.

Rhodes, E.W., Teferman, N. B., Cook, E. and Schwartz, D. (1979). T-Scope tests of yellow pages advertising. *Journal of Advertising Research*, 19, 49-52.

Rossi, P. E. & Allenby, G. M.(2003). Bayesian Statistics and Marketing. *Marketing Science*, 22(3), 304-328.

Rossi, P. E. & Allenby, G. M. (2005). Bayesian Statistics and Marketing. *John Wiley and Sons.*

Rutz, O. & Bucklin, R (2010). From Generic to Branded: A Model of Spillover s in Paid Search Advertising. *Forthcoming in Journal of Marketing Research*.

Srinivasan, N., & Ratchford, B. T. (1991). An empirical test of a model of external search for automobiles. *Journal of Consumer Research*, 18, 233–242.

Trueman, B., Wong, M.H.F., and Zhang, X.-J. "Back to Basics: Forecasting the Revenues of Internet Firms," in: Review of Accounting Studies, Springer Science & Business Media B.V., 2001, pp. 305-329.

William B. Dodds, Kent B. Monroe, Dhruv Grewal, 1991: Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. Journal of Marketing Research, Vol. 28, No. 3 , pp. 307-319.

Varian, H.R (2006). Position Auctions. *International Journal of Industrial Organization*, 2006.

Weber, T.A. and Z. Zheng (2007). A Model of Search Intermediaries and Paid Referrals. *Information Systems Research*, 18 (4), 414-436.

Yang, S. and Ghose, A. (2009). Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence? *Marketing Science*, forthcoming.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57, 348-368.

Table 1: Keyword Performance Summary Statistics

Variable	Mean	St. Dev.	Min	Max
Impressions	82.4	160	1	1666
Clicks	1.21	2.1	0	24
Orders	0.02	0.18	0	3
Pos	3.6	1.78	1	10
Organic Link	0.15	0.35	0	1
SearchQuality	-1.72	2.5	-11.9	10.7
AdQuality	-4.9	3.4	-12.7	4.3
LQScore	8.1	1.5	5	10
Brand	0.75	0.44	0	1
Specificity	0.81	0.68	0	2
Bid	0.74	0.35	0.08	2

Table 2: Estimates for the CTR

	Intercept	Brand	Specificity
Const	-2.7 (0.19)***	-0.32 (0.48)	0.12 (0.3)
Pos	-0.46 (0.08)***	-0.1 (0.19)	0.02 (0.13)
Organic Link	0.34 (0.15)**		
SearchQuality	0.01 (0.01)		
AdQuality	-0.01 (0.02)		
Time	-0.01 (0.0)***		

Table 3: Estimates for the covariance matrix V^{θ}

Variables	Const	pos
Const	0.89 (0.23)***	-0.12 (0.07)*
pos	-0.12 (0.07)*	0.17 (0.04)***

Table 4: Estimates for the CONV

	Intercept	Brand	Specificity
Const	-2.25 (0.38)***	0.29 (0.76)	-0.37 (0.48)
AdRank	0.26 (0.13)**	0.0 (0.25)	0.21 (0.16)
Organic Link	-0.78 (0.19)***		
SearchQuality	0.04 (0.01)**		
AdQuality	0.03 (0.02)*		
Time	-0.02 (0.0)***		

Table 5: Estimates for the covariance matrix V^{β}

Variables	Const	pos
Const	0.93 (0.2)***	-0.04 (0.06)
pos	-0.04 (0.06)	0.15 (0.03)***

Table 6: Estimates for the Ad Position

Variables	Intercept	Brand	Specificity
Const	1.53 (0.31)***	0.03 (0.26)	0.39 (0.19)**
bid	-0.51 (0.07)***	0.1 (0.18)	-0.03 (0.13)
LQScore	-0.32 (0.14)**		
CompBid	-0.05 (0.06)		
Time	0.0 (0.0)		

Table 7: Estimates for the Covariance Matrix V^{α}

Variables	Const	bid
Const	0.42 (0.1)***	-0.01 (0.04)

	bid	-0.01 (0.04)	0.18 (0.04)***

Table 8: Estimates for the Covariance Matrix Ω

.	CONV	CTR	Rank
CONV	0.5 (0.06)***	-0.17 (0.02)***	-0.04 (0.02)**
CTR	-0.17 (0.02)***	0.28 (0.03)***	0.02 (0.01)**
Rank	-0.04 (0.02)**	0.02 (0.01)**	0.09 (0.0)***

Table 9: Prediction Accuracy for Estimation & Holdout Samples

Models	CTR Fit (MAPE)		CONV Fit (MAPE)	
	Aggregate	Keyword	Aggregate	Keyword
Estimation Sample	0.44	0.43	0.30	0.30
Holdout Sample	0.46	0.46	0.33	0.33

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

Table 10: Parameter estimates for CTR using a different measure of quality

Variables	Intercept	Brand	Specificity
Const	-2.74 (0.2)***	-0.32 (0.47)	0.1 (0.3)
Pos	-0.46 (0.08)***	-0.1 (0.19)	0.02 (0.13)
Organic Rank	0.37 (0.15)**		
SearchQuality	-0.01 (0.04)		
AdQuality	0.02 (0.02)		
Time	-0.01 (0.0)***		

Table 11: Estimates for the covariance matrix V^θ using a different measure of quality

Variables	Const	pos
Const	0.88 (0.23)***	-0.11 (0.07)*
pos	-0.11 (0.07)*	0.17 (0.04)***

Table 12: Parameter estimates for CONV using a different measure of quality

	Intercept	Brand	Specificity
Const	-2.38 (0.38)***	0.12 (0.8)	-0.45 (0.5)
AdRank	0.25 (0.13)**	0.01 (0.25)	0.22 (0.16)
Organic Link	-0.78 (0.18)***		
SearchQuality	-0.14 (0.04)**		
AdQuality	0.03 (0.02)		
Time	-0.02 (0.0)		

Table 13: Estimates for the covariance matrix V^β using a different measure of quality

Variables	Const	pos
Const	2.54 (0.65)***	-0.45 (0.18)**
pos	-0.45 (0.18)**	0.3 (0.07)***

Figure 1: Search Results