Marketing Science

Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?

Sha Yang, Anindya Ghose,

Please scroll down for article—it is on subsequent pages

Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?

Sha Yang, Anindya Ghose
Stern School of Business, New York University, New York, New York 10012
{shayang@stern.nyu.edu, aghose@stern.nyu.edu}

The phenomenon of paid search advertising has now become the most predominant form of online advertising in the marketing world. However, we have little understanding of the impact of search engine advertising on consumers' responses in the presence of organic listings of the same firms. In this paper, we model and estimate the interrelationship between organic search listings and paid search advertisements. We use a unique panel data set based on aggregate consumer response to several hundred keywords over a three-month period collected from a major nationwide retailer store chain that advertises on Google. In particular, we focus on understanding whether the presence of organic listings on a search engine is associated with a positive, a negative, or no effect on the click-through rates of paid search advertisements, and vice versa for a given firm. We first build an integrated model to estimate the relationship between different metrics such as search volume, click-through rates, conversion rates, cost per click, and keyword ranks. A hierarchical Bayesian modeling framework is used and the model is estimated using Markov chain Monte Carlo methods. Our empirical findings suggest that click-throughs on organic listings have a positive interdependence with click-throughs on paid listings, and vice versa. We also find that this positive interdependence is asymmetric such that the impact of organic clicks on increases in utility from paid clicks is 3.5 times stronger than the impact of paid clicks on increases in utility from organic clicks. Using counterfactual experiments, we show that on an average this positive interdependence leads to an increase in expected profits for the firm ranging from 4.2% to 6.15% when compared to profits in the absence of this interdependence. To further validate our empirical results, we also conduct and present the results from a controlled field experiment. This experiment shows that total click-through rates, conversions rates, and revenues in the presence of both paid and organic search listings are significantly higher than those in the absence of paid search advertisements. The results predicted by the econometric model are also corroborated in this field experiment, which suggests a causal interpretation to the positive interdependence between paid and organic search listings. Given the increased spending on search engine-based advertising, our analysis provides critical insights to managers in both traditional and Internet firms.

Key words: paid search advertising; organic search listings; search engines; click-through rates; conversion rates; electronic commerce; Internet markets; monetization of user-generated content; hierarchical Bayesian modeling

History: Received: February 10, 2008; accepted: October 21, 2009; processed by Alan Montgomery. Published online in *Articles in Advance* April 8, 2010.

1. Introduction

Over the past few years, search engines like Google, Yahoo!, and MSN have discovered that as intermediaries between consumers and firms, they are in a unique position to sell new forms of advertisements. This has led to the proliferation of sponsored search advertising—where advertisers pay a fee to Internet search engines to be displayed alongside nonsponsored or organic Web search results. Because search engine-based advertising is directly related to users' search queries, it is considered by users as being far less intrusive relative to other forms of online advertising. From the firm side, this kind of advertising leads to more qualified prospects because the ads are displayed in response to user-originated search behavior. From the consumer perspective, this is a much more relevant form of advertising because the keywords and the ad message are typically matched with user-generated queries. These features have lead to a widespread adoption of this form of Web 2.0 media by firms, with the global paid search market expected to reach almost $10 billion by the end of 2009.[1]

How does this mechanism work? In sponsored search, firms who wish to advertise their product or services on the Internet create text-based ads and submit that information in the form of "keyword" listings to search engines. A keyword is a combination of words or terms that best describes the product, brand, or retailer being advertised. Bid values are

---

[1] See eMarketer (2006).

assigned to each individual keyword and then search engines pit advertisers against each other in second-price auction-style bidding for the highest positions on search engine result pages. When users search for that keyword on a search engine, the relevant ad along with the advertisers' Web page appears as a sponsored link on the top and the right side of the organic search results. When users click on the sponsored ad, they are taken to the advertiser's website.

An important determinant of the effectiveness of sponsored search advertising for a given advertiser is the likelihood of the same advertiser appearing in the natural or organic listings of the search engine, and its position on the organic listings for a given keyword. Organic rankings of advertisers' websites are based on a complex and proprietary indexing algorithm devised by the search engine involving the quality of the website and the website's "relative importance" with respect to other links. Thus, consumers often face two competing lists of results that may both be relevant to their search query: (i) the sponsored search listing and (ii) the organic search listing.

Advertisers have been grappling with the trade-offs in each of these two forms of referrals. On the one hand, because a firm can control the message of paid search ads, one would expect higher conversions from them. On the other hand, because people value the perceived "editorial integrity" of organic listings, one would expect higher conversions from them. Some anecdotal evidence suggests that there is a potential disconnect between the perception of sponsored listings by business and users, with consumers having a positive bias towards organic search listings. For example, Hotchkiss et al. (2005) find that more than 77% of participants favored organic links more than the sponsored links as offering sources of trusted, unbiased information. A similar set of findings are reported in Greenspan (2004). On the other hand, Jansen (2007) finds that sponsored links are more relevant than organic links in the context of e-commerce search queries, and this finding is robust to usage across Google, Yahoo!, and MSN search engines. Moreover, there is also some anecdotal evidence suggesting that paid search may lead to higher conversions than organic search.[2] This study looked at search engine visits on Google, Yahoo!, and MSN and showed a median order conversion rate of 3.4% for paid search compared to a conversion rate of 3.13% for organic search results during the same time frame.

These mixed findings then motivate the question regarding to what extent should firms invest in sponsored search advertisements when they also appear in the organic listings for a given search query in that

search engine. After all, firms incur costs while engaging in paid search advertising. Despite the growth of search advertising, we have little understanding of the impact of sponsored search advertising on consumers' overall reaction in the presence of organic search listings for a given firm. Our key objective in this paper is to compare and analyze the interdependence between paid ad listings with organic listings for a given advertiser. Hence, we focus on the following questions: How do different keyword-level attributes impact performance metrics such as average click-through rates and conversion rates in paid search advertising as compared with those in organic search listings? What is the nature of the relationship between average click-through rates on organic and paid listings for a given set of keywords? Are the combined click-through rates, conversion rates, and revenues from sponsored and organic listings higher or lower than those from organic listings only (i.e., when paid search advertising is paused)?

Although an emerging stream of theoretical literature in sponsored search has looked at issues such as mechanism design in auctions, no prior work has empirically analyzed these kinds of questions. By modeling the association between paid and organic clicks, we aim to examine if there is a positive, negative, or zero association between them. Towards this goal, we use a unique panel data set of aggregate consumer response to several hundred keywords over three months collected from a large nationwide retailer that advertises on Google. To be clear, we only have aggregate keyword-level data, not disaggregate user-level data. We propose a hierarchical Bayesian modeling framework in which we model consumers' aggregate responses jointly with the advertiser's and search engine's decisions. Our paper is the first academic study that estimates the effect of sponsored search advertising on consumer search, click, and conversion behavior in the presence of organic listings of the same firm for the same set of keywords at the same time. We aim to make the following four contributions:

First, we build an integrative simultaneous equation-based model to empirically estimate the impact of ad rank and other keyword-level attributes on consumer aggregate responses, and advertiser and search engine decisions in the presence of both paid and organic search listings. Specifically, we examine the relationship of these attributes with average search, click-through, and purchase propensities as well as with the advertiser's cost per click (*CPC*) and the search engine's decision with respect to allocating keyword ranks in accordance with institutional practices. We find that on average the presence of retailer-specific information and brand-specific information is a significant predictor of conversion rates

---

[2] See http://www.searchnewz.com/blog/talk/sn-6-20060925Organic VersusPaidSearchResults.html.

in paid and organic listings, respectively. A simultaneous equations model as proposed in our study makes it possible for us to describe current phenomena and prescribe some recommendations to advertisers. It also provides two key benefits. On one hand, we are able to account for the potential endogeneity of keyword rank and *CPC*. On the other hand, this allows us to do some simple policy simulations to infer the optimal *CPC* for different keywords and thereby impute the impact of sponsoring ads in the presence of organic listings on firms' profits.

Second, we investigate the value to firms from participating in such sponsored search advertising by examining the nature of interdependence between organic and paid listings using an aggregate keyword-level data set of clicks and conversions on paid and organic links. We also explore the asymmetric nature of the relationship between these two forms of search engine listings. Based on our model, we find that average click-throughs on organic search listings have a positive interdependence with average click-throughs on paid listings, and vice versa. This positive interdependence is also asymmetric such that the impact of organic clicks on increase in utility from paid clicks for the same firm is 3.5 times stronger than the impact of paid clicks on increases in utility from organic clicks. We provide details of alternative econometric models that we build and estimate to explore this interdependence. All models provide qualitatively similar results.

Third, we conduct some counterfactual experiments using policy simulations to highlight the magnitude of the positive interdependence between these two forms of advertising. We find that on an average, this positive interdependence leads to an increase in expected profits for the firm ranging from 4.2% to 6.15% when compared to profits in the absence of either of these. Furthermore, the positive interdependence is the strongest in the case of the "least competitive" keywords (such as retailer-specific keywords) and weakest in the case of the "most competitive" keywords (such as brand-specific and generic keywords). Therefore, the proposed parsimonious modeling framework can help advertisers make optimal decisions and investigate the value from participating in such sponsored search advertising in the presence of natural or organic listings in search engines.

Finally, we describe a simple field experiment that sheds further light on the predictions from our empirical model and analysis. It shows that given a set of keywords, combined click-through rates, conversion rates, and total revenues accruing to the firm in the presence of both paid and organic listings is higher than those in the absence of paid search advertisements. We find that although the presence of paid search advertisements takes some traffic away from organic listings for some keywords, for a vast majority

of keywords in our sample, the average click-through rate of organic listings when paid search was on was higher than the average click-through rate of organic listings when paid search was inactive. This result shows that the positive interdependence effect predicted by the econometric model is also corroborated in controlled experiments and potentially suggests a causal interpretation to the results obtained from the model.

To evaluate the consistency of results between the field experiment and the estimated model, we conduct analysis using the specific sample of keywords for the duration for which the experiment was run with paid links. The results from these analyses are consistent with those from the integrated model and show that on an average, paid click-through rates are positively associated with organic click-through rates, and vice versa. Furthermore, the positive interdependence is also asymmetric in nature. These results remain qualitatively the same even when we estimate additional model specifications such as the autologistic model and the simultaneous-move game structural model.

The remainder of this paper is organized as follows. Section 2 gives an overview of the different streams of literature related to our paper. Section 3 presents a simultaneous model of consumer search, click, and purchase; the advertiser's keyword pricing decision; and the search engine's decision of assigning keyword ranks. Section 4 presents an empirical application of the proposed model along with a description of the data that are used. We also present a discussion of a number of robustness checks we have conducted using alternative model specifications such as an autologistic model and a simultaneous-move game model. The details are in the electronic companion to this paper, available as part of the online version that can be found at http://mktsci.pubs.informs.org. We describe the counterfactual experiments conducted through some simple policy simulations in §5. Section 6 presents a controlled field experiment that also analyzes the effect of organic search listings on paid search performance. Section 7 presents some managerial implications and concludes the paper with a discussion of some limitations.

2. Literature and Theoretical Foundation

Our paper is related to several streams of research. First, it relates to recent research in online advertising. A number of approaches have modeled the effects of advertising based on aggregate data (Tellis 2004). However, much of the existing academic research (e.g., Gallagher et al. 2001, Drèze and Hussherr 2003) on advertising in the online world has focused on measuring changes in brand awareness, brand

attitudes, and purchase intentions as a function of exposure. Other studies measure (individual) exposure to advertising via aggregate advertising dollars (e.g., Mela et al. 1998, Ilfeld and Winer 2002). Because banner ads have been perceived by many consumers as being annoying, traditionally they have had a negative connotation associated with it. Moreover, it was argued that because there is considerable evidence that only a small proportion of visits translate into final purchase (Moe and Fader 2003, Chatterjee et al. 2003), click-through rates may be too imprecise for measuring the effectiveness of banners served to the mass market. Interestingly, however, Manchanda et al. (2006) found that banner advertising actually increases purchasing behavior, in contrast to conventional wisdom. These studies therefore highlight the importance of investigating the impact of other kinds of online advertising such as search keyword advertising on actual purchase behavior, because the success of keyword advertising is also based on consumer click-through rates. Our study is also related to other forms of paid placements available to retailers on the Internet such as sponsored listings on shopping bots. For example, Baye and Morgan (2001), Montgomery et al. (2004), and Baye et al. (2009) have studied the role of shopping bots as information gatekeepers and estimated the impact of retailers' rank during placement on click-through rates.

From a theoretical perspective, our paper has interesting parallels to the traditional product placement based advertising next to editorial content in the mass media. Shapiro et al. (1997) report a study of magazine advertising in which participants were required to read a magazine article delivered on a computer screen. The article was flanked by target ads designed to receive minimal reader attention. They found that users were more likely to include the product featured in the ad in their consideration set compared to the control participants who had not viewed the ads. Geerardyn and Fauconnier (2000) discuss the advent of "advertorials," which are printed advertising messages but have the look and content of an ordinary newspaper or magazine article. In creating the ad, the editorial form of the medium and article content next to which the ad is to be placed is taken into account. Research in consumer behavior has theorized about these effects. For example, the truth effect (the increased belief in an ad claim because of a previous exposure; see Hawkins and Hoch 1992, Law and Braun 2000) and the mere exposure effect (the formulation of a positive affect from exposure to a brief stimulus; see Janiszewski 1993) both illustrate changes in consumer behavior following a single exposure to a stimuli without awareness of the prior exposure.

There is also an emerging theoretical stream of literature exemplified by Edelman et al. (2007), Feng et al. (2007), Varian (2007), and Liu et al. (2010) that study mechanism design in sponsored keyword auctions. Athey and Ellison (2008) build a model that integrates consumer behavior with advertiser decisions. Wilbur and Zhu (2008) examine the incentives of search engines to prevent click fraud. Katona and Sarvary (2010) build a model of competition in sponsored search and find that the interaction between search listings and paid links determines equilibrium bidding behavior. Xu et al. (2009) find that whereas organic listing may hurt search engine revenue, it could induce higher social welfare and sales diversity. Despite the emerging theory work, very little empirical work exists in online sponsored search advertising. Existing work has so far focused on search engine performance (Bradlow and Schmittlein 2000, Telang et al. 2004) and examined issues related to adverse selection (Animesh et al. 2010) and pricing differences based on ad context (Goldfarb and Tucker 2007).

Our paper is closely related to an emerging stream of work that uses firm-level data from search engine advertisers. Rutz and Bucklin (2007) study conversion probability for hotel marketing keywords in Los Angeles. Rutz and Bucklin (2008) also show that there are spillovers between search advertising on branded and generic keywords; some customers may start with a generic search to gather information, but they later use a branded search to complete their transaction. Ghose and Yang (2008a) compare paid search advertising to organic listings with respect to predicting order values and profits. Ghose and Yang (2009) also quantify the impact of keyword attributes on consumer search and purchase behavior as well as on advertiser's cost per click and the search engine's ranking decision for different ads. They find that keyword rank influences not only ad click-through rates but also the final conversion rates from the advertiser's website, thus implying that the value per click is not uniform across slots on the search engine's results page. They also show that keywords that have more prominent positions on the search engine results page, and thus experience higher click-through or conversion rates, are not necessarily the most profitable ones—profits are often higher at the middle positions than at the top or the bottom ones. Agarwal et al. (2008) provide quantitative insights into the profitability of advertisements associated with differences in keyword position. Ghose and Yang (2008b) build a model to map consumers' search–purchase relationship in sponsored search advertising. They provide evidence of horizontal spillover effects from search advertising resulting in purchases across other product categories. Yao and Mela (2009) build a dynamic structural model to explore how the interaction of consumers, search engines, and advertisers affects consumer welfare and firm profits. Gerstmeier

et al. (2009) discuss some interesting bidding heuristics and highlight which of these leads to higher profits for the advertiser. Jerath et al. (2009) find that a superior firm may obtain a position below the inferior firm but still obtain more clicks than the inferior firm. Thus, a firm may not always want to be in the topmost position, a finding consistent with the rank-profitability relationship in Ghose and Yang (2009).

3. An Integrative Model of Consumer Response and Firms' Decisions

We next present a model that integrates consumer search and purchase behavior with firms' decision-making behavior such as price of a keyword and the rank of a keyword ad after the auction. This model considers the simultaneous presence of both paid and organic search listings. A hierarchical Bayesian modeling framework is used and the models are estimated using Markov chain Monte Carlo (MCMC) methods (Rossi and Allenby 2003).

Let us denote N_{it} as the total number of searches for keyword i in week t. We model the total number of searches over time as a log-normal regression specified as follows:[3]

$$\ln(N_{it}) = \psi_i + \delta_1\,Retailer_i + \delta_2\,Brand_i$$
$$+ \delta_3\,Length_i + \delta_4\,Time_{it} + \kappa_{it}, \qquad (1a)$$

$$\psi_i \sim N(\bar{\psi}, \sigma_\psi^2) \qquad (1b)$$

The covariates, *Brand*, *Retailer*, and *Length*, in the above equation are described below. Prior work (Broder 2002) has analyzed the goals for users' Web searches and classified user queries in search engines into three categories of searches: *navigational* (for example, a search query consisting of a specific firm or retailer), *transactional* (for example, a search query consisting of a specific product), or *informational* (for example, a search query consisting of longer words). Search engines not only sell nonbranded or generic keywords as advertisements but also well-known product or manufacturer brand names, as well as keywords indicating the specific advertiser in order for the firm to attract consumers to its website.[4] Moreover, advertisers also have the option of making the keyword generic or specific by altering the number

of words contained in the keyword. Hence, we focus on the three important keyword-specific characteristics for a firm (the advertiser) when it advertises on a search engine. This includes whether the keyword has (i) retailer-specific information (for example, "Wal-Mart," "walmart.com") or not, (ii) brand-specific information (for example, "Sealy mattress") or not, and (iii) the length (in words) of the keyword (which determines how narrow or broad the consumer search is). Based on these factors, we construct three keyword-specific characteristics denoted by *Brand*, *Retailer*, and *Length*. The first two variables are coded as dummy variables. The industry dynamic effects are controlled for by adding a time trend in all equations.[5] This is consistent with prior work in this area (see, for example, Ghose and Yang 2009).

For a given keyword, although some searches do not lead to any clicks at all, some searches lead to clicks on either organic or paid listings, and some searches lead to clicks on both listings. This dual-click search behavior is also shown by Jansen et al. (2007), and hence it is important to incorporate such dual-click search behavior in the model and analysis. Let $N_{it}^{1,0}$ denote the number of click-throughs on the same keyword only in the organic listing, let $N_{it}^{0,1}$ denote the number of click-throughs on the same keyword only in the paid listing, and let $N_{it}^{1,1}$ denote the number of click-throughs on the same keyword in both the organic and paid listings. The remaining $N_{it}^{0,0} = N_{it} - N_{it}^{1,0} - N_{it}^{0,1} - N_{it}^{1,1}$ searches lead to zero click-throughs. Based on this, the likelihood function is multinomial:

$$f(N_{it}^{1,0}, N_{it}^{0,1}, N_{it}^{1,1}, N_{it}^{0,0})$$
$$\propto (p_{it}^{1,0})^{N_{it}^{1,0}}(p_{it}^{0,1})^{N_{it}^{0,1}}(p_{it}^{1,1})^{N_{it}^{1,1}}(p_{it}^{0,0})^{N_{it}^{0,0}}. \qquad (2)$$

Let us further assume the probability of a click-through takes the logit form $\exp(\pi_{it1})/[1+\exp(\pi_{it1})]$ for organic listings, and $\exp(\pi_{it2})/[1+\exp(\pi_{it2})]$ for paid listings. Here, π_{it1} denotes the latent utility of an average click on organic listings, and π_{it2} denotes the latent utility of an average click on paid listings. Then by assuming the conditional independence of a click-through on organic listings and a click-through on paid listings, we can write down the four probabilities specified in Equation (2) as follows:

$$p_{it}^{1,0} = \int_{\pi_{it2}} \int_{\pi_{it1}} \frac{\exp(\pi_{it1})}{1+\exp(\pi_{it1})}$$
$$\cdot \frac{1}{1+\exp(\pi_{it2})} p(\pi_{it1}, \pi_{it2})\,d\pi_{it1}\,d\pi_{it2}, \quad (3a)$$

---

[3] We use a log-normal regression on the total search volume because of the existence of outliers, which do not make the normal distribution a good one to apply.

[4] For example, a consumer seeking to purchase a digital camera is as likely to search for a popular manufacturer brand name such as Nikon, Canon, or Kodak on a search engine as searching for the generic phrase "digital camera." Similarly, the same consumer may also search for a retailer such as Best Buy or Circuit City to buy the digital camera directly from the retailer.

[5] In addition to adding a time trend, we have also explored the specification of correlated error terms over time. However, the autoregressive coefficients did not come out significant for any of the seven equations. Therefore, we present the results with time trend effect.

$$p_{it}^{0,1} = \int_{\pi_{it2}} \int_{\pi_{it1}} \frac{1}{1 + \exp(\pi_{it1})}$$
$$\cdot \frac{\exp(\pi_{it2})}{1 + \exp(\pi_{it2})} p(\pi_{it1}, \pi_{it2}) \, d\pi_{it1} \, d\pi_{it2}, \quad (3b)$$

$$p_{it}^{1,1} = \int_{\pi_{it2}} \int_{\pi_{it1}} \frac{\exp(\pi_{it1})}{1 + \exp(\pi_{it1})}$$
$$\cdot \frac{\exp(\pi_{it2})}{1 + \exp(\pi_{it2})} p(\pi_{it1}, \pi_{it2}) \, d\pi_{it1} \, d\pi_{it2}, \quad (3c)$$

$$p_{it}^{0,0} = \int_{\pi_{it2}} \int_{\pi_{it1}} \frac{1}{1 + \exp(\pi_{it1})}$$
$$\cdot \frac{1}{1 + \exp(\pi_{it2})} p(\pi_{it1}, \pi_{it2}) \, d\pi_{it1} \, d\pi_{it2}. \quad (3d)$$

We further model the interdependent relationship of the two latent utilities associated with the two types of click-throughs for the same keyword as follows:

$$\pi_{it1} = \chi_{it1} + \theta_i^{12} \pi_{it2} + \eta_{it1}, \quad (4a)$$
$$\pi_{it2} = \chi_{it2} + \theta_i^{21} \pi_{it1} + \eta_{it2}. \quad (4b)$$

The above two equations incorporate the notion that the latent utility of an organic search and a paid search is dependent on both its intrinsic utility (χ_{it}) and the extrinsic utility from each other ($\theta_i^{12} \pi_{it2}$ and $\theta_i^{21} \pi_{it1}$). With respect to the intrinsic utility, we model it to be dependent on the kind of keyword that is displayed in response to a search query. This is modeled as

$$\chi_{its} = \beta_{is1} + \beta_{is2} \, Rank_{its} + \alpha_{s1} \, Retailer_i + \alpha_{s2} \, Brand_i$$
$$+ \alpha_{s3} \, Length_i + \alpha_{s4} \, Time_{it}, \quad s = 1, 2. \quad (5)$$

With respect to the extrinsic utility, θ_i^{12} and θ_i^{21} indicate the effect that maps the interdependence between paid listings and organic listings for keyword i. A positive sign on θ_i^{12} and θ_i^{21} suggest a positive interdependency or complementary relationship between the click-throughs via organic and paid listings. That is, the click-through on the organic (paid) listing tends to increase the utility of a click-through on the paid (organic) listing. Similarly, a negative sign on θ_i^{12} and θ_i^{21} suggests a negative interdependency or substitutive relationship between the click-throughs via organic and paid listings. Finally, a zero value of θ_i^{12} and θ_i^{21} suggests independence between the click-through via the organic and paid listings. Finally, we model the unobserved heterogeneity across keywords as

$$[\beta_{i11}, \beta_{i12}, \theta_i^{12}]' \sim \text{MVN}(\bar{\beta}_1, \Sigma_1^{\beta\theta}), \quad (6a)$$
$$[\beta_{i21}, \beta_{i22}, \theta_i^{21}]' \sim \text{MVN}(\bar{\beta}_2, \Sigma_2^{\beta\theta}). \quad (6b)$$

So far, we have modeled the click-through rates. Next, we model the conversion behavior conditional on the click-through. Denote M_{it}^1 as the total number of conversions for keyword i in week t from organic searches, M_{it}^2 as the total number of conversions for keyword i in week t from paid searches, and q_{it1} (q_{it2}) as the conversion probability for organic (paid) searches conditional on a click-through. Then assuming the conversions are independent events, we can write down the likelihood of M_{it}^1 and M_{it}^2 as follows:

$$f(M_{it}^1 \mid N_{it}^{1,0}, N_{it}^{0,1}, N_{it}^{1,1}, N_{it}^{0,0})$$
$$= (q_{it1})^{M_{it}^1} (1 - q_{it1})^{N_{it}^{1,0} + N_{it}^{1,1} - M_{it}^1}, \quad (7a)$$
$$f(M_{it}^2 \mid N_{it}^{1,0}, N_{it}^{0,1}, N_{it}^{1,1}, N_{it}^{0,0})$$
$$= (q_{it2})^{M_{it}^2} (1 - q_{it2})^{N_{it}^{0,1} + N_{it}^{1,1} - M_{it}^2}. \quad (7b)$$

Prior work (Brooks 2005) has shown that there is an intrinsic trust value associated with the rank of a listing on a search engine, which leads to the conversion rate dropping significantly with an increase in the rank (i.e., with a lower position on the screen). Another factor that can influence conversion rates is the quality of the landing page of the advertiser's website. Anecdotal evidence suggests that if online consumers use a search engine to direct them to a product but do not see it addressed adequately on the landing page, they are likely to abandon that site. Different keywords lead to different kinds of landing pages. In keeping with the institutional practices of search engines, we use the click-through rates (CTRs) (standardized in our empirical analysis) to control for the landing page quality score,[6] where click-through rate is defined as the number of clicks over the number of searches. Furthermore, different keywords are associated with different products. It is possible that some product-specific characteristics influence consumer conversion rates, and thus, it is important to control for the unobserved product characteristics that may influence conversion rates once the consumer is on the website of the advertiser. Hence, we include the three keyword characteristics to proxy for the unobserved keyword heterogeneity stemming from the different products sold by the advertiser. These factors lead us to model the conversion probabilities

---

[6] Google computes a quality score for each keyword as a function of the relevancy, transparency, and navigability of information on the landing page and the past click-through rate of that keyword (Ghose and Yang 2009). The key idea is to provide a higher user experience after a click-through to the advertiser's site from their search engine. However, because we do not have information on the landing page quality scores, we use the click-through rate as the proxy for landing page quality scores. Further information on this metric is available at http://www.adwords.google.com.

on organic listings ($s = 1$) and on paid listings ($s = 2$) as follows:

$$
\begin{aligned}
q_{its} = \big(&\exp(c_{is1} + c_{is2} Rank_{its} + \gamma_{s1} CTR_{its} + \gamma_{s2} Retailer_i \\
&+ \gamma_{s3} Brand_i + \gamma_{s4} Length_i + \gamma_{s5} Time_{it} + \mu_{its})\big) \\
\cdot \big(1 + &\exp(c_{is1} + c_{is2} Rank_{its} + \gamma_{s1} CTR_{its} + \gamma_{s2} Retailer_i \\
&+ \gamma_{s3} Brand_i + \gamma_{s4} Length_i + \gamma_{s5} Time_{it} + \mu_{its})\big)^{-1}, \quad (8a)
\end{aligned}
$$

$$
[c_{is1}, c_{is2}]' \sim \text{MVN}(\bar{c}_s, \Sigma_s^c). \quad (8b)
$$

Next, we model the cost per click (*CPC*) of the ad keywords posted in the sponsored search list.[7] Because different keyword attributes determine whether it is a generic or branded keyword, the advertiser's *CPC* for a given keyword will depend on these attributes. The advertiser decides on its bid price by tracking the performance of a keyword over time such that the current bid price is dependent on past performance of that keyword. It does this in two ways.[8] First, the keyword's current bid price is a function of the rank of the same keyword in the previous period, in both the paid and organic listing. Second, the keyword's bid price is also based on the extent of profits from conversions through paid and organic listings, respectively, in the previous period. Here, profit is defined as the revenue from advertising net of the variable cost of the product being sold through that keyword minus the costs of placing that advertisement for the firm (the advertisement cost is equal to the total number of clicks times cost per click). Hence, we include both these sets of covariates.[9] Finally, we also control for possible competitive effects of other advertisers by including the maximum bid price for a given keyword (denoted by *Competitor_Price*). This leads to the following equation:

$$
\begin{aligned}
\ln(CPC_{it}) = \omega_i &+ \lambda_1 Rank_{i, t-1, 1} + \lambda_2 Rank_{i, t-1, 2} \\
&+ \lambda_3 \ln(Profit_{i, t-1, 1}) + \lambda_4 \ln(Profit_{i, t-1, 2}) \\
&+ \lambda_5 Retailer_i + \lambda_6 Brand_i + \lambda_7 Length_i \\
&+ \lambda_8 Competitor\_Price_i + \lambda_9 Time_{it} + \varsigma_{it}, \quad (9a)
\end{aligned}
$$

$$
\omega_i \sim N(\bar{\omega}, \sigma_\omega^2). \quad (9b)
$$

[7] Because we do not have data on actual bids, we use the actual cost per click as a proxy for the bid price. These two are highly correlated.

[8] This information about current bid prices being based on the two metrics of past performance (lagged profit and lagged rank) was given to us by the advertiser. Our results are robust to the use of either one of these two metrics as well as to their exclusion. Gerstmeier et al. (2009) also discuss that current period bid can be a function of past profits from that keyword.

[9] Our results are robust to the use of gross profits in which we consider only the advertisement revenues and advertisement-related costs as well net profits in which we consider the variable costs of the products.

Finally, we model the search engine's decision on assigning ranks for the sponsored keyword. Search engines like Google, MSN, and Yahoo! decide on the ranks during the keyword auction by taking into account both the current bid price and the quality score. Because the quality score is most affected by the prior click-through rates, and more recent CTR is given higher weightage by the search engine in computing this score, we use the one period lagged value of CTR (standardized in our empirical analysis) as a control variable. As before, we use the three keyword attributes to proxy for the different unobserved characteristics of the landing page. We control for possible competitive effects of other advertisers by including the max bid price for a given keyword, *Competitor_Price*. This leads to the following equation for the rank of a keyword in sponsored search:[10]

$$
\begin{aligned}
\ln(Rank_{it2}) = \phi_i &+ \tau_1 \ln(CPC_{it}) + \tau_2 CTR_{i, t-1, 2} \\
&+ \tau_3 Retailer_i + \tau_4 Brand_i \\
&+ \tau_5 Length_i + \tau_6 Competitor\_Price_i \\
&+ \tau_7 Time_{it} + v_{it}, \quad (10a)
\end{aligned}
$$

$$
\phi_i \sim N(\bar{\phi}, \sigma_\phi^2). \quad (10b)
$$

We allow the error terms in the five levels of decisions to be correlated. That is,

$$
(\kappa_{it}, \eta_{it1}, \eta_{it2}, \mu_{it1}, \mu_{it2}, \varsigma_{it}, v_{it})' \sim \text{MVN}(0, \Omega). \quad (11)
$$

3.1. Econometric Issues and Identification

The specification of the covariance of the error terms is important to help control the potential endogeneity in the keyword ranks for paid searches and conversions. To show this endogeneity issue and the identification of the proposed system of simultaneous equation model, we provide a sketch of the model below. The proposed seven equations, in essence, can be written as follows:

$$
Rank_2 = f_1(CPC, X_1, \varepsilon_1), \quad (12a)
$$

$$
CPC = f_2(X_2, \varepsilon_2), \quad (12b)
$$

$$
q_1 = f_3(Rank_1, X_3, \varepsilon_3), \quad (12c)
$$

$$
q_2 = f_4(Rank_2, X_4, \varepsilon_4), \quad (12d)
$$

$$
\pi_1 = f_5(Rank_1, X_5, \varepsilon_5), \quad (12e)
$$

$$
\pi_2 = f_6(Rank_2, X_6, \varepsilon_6), \quad (12f)
$$

$$
N = f_7(X_7, \varepsilon_7). \quad (12g)
$$

[10] There are three reasons for using the log (rank). First, rank is measured as weekly average rank for a keyword, and therefore it is a continuous variable rather than an integer. Second, the log transformation makes its distribution closely mimic a normal distribution, and this mitigates the effect of outliers. Third, we have also tried an alternative linear model and found that the log-transformed model performs slightly better than the linear model for both the in-sample fit and out-sample fit.

In the above simultaneous equations system, X_1–X_7 are the exogenous covariates associated with the seven equations, respectively. Note that X_5 includes one covariate that is the instrument of π_2, that is, $\hat{\pi}_2$. Similarly, X_6 includes one covariate that is the instrument of π_1; that is, $\hat{\pi}_1$. The terms $\varepsilon_1, \ldots, \varepsilon_7$ are the error terms associated with the seven equations, respectively. These error terms are mainly capturing information that is observed by the decision makers (consumer, advertiser, and search engine) but not observed by the researcher. The proposed system of simultaneous equations presented in Equations (12a), (12b), (12d), and (12f) closely resembles the triangular system in standard econometric textbooks (Greene 1999, p. 679). This model is identified based on the following argument. First, Equation (12b) is a classical regression model whose parameters can be naturally identified. To see this more clearly, *CPC* is modeled as exogenously determined (modeled as the advertiser's decision and a function of the advertiser's past performance with the same keyword in both paid and organic listings, and other keyword related characteristics as specified in Equation (12b)). *CPC*, in turn, affects the search engine's ranking decision for paid ads, $Rank_2$, and finally $Rank_2$ affects both click-through and the conversion probabilities. Thus, Equations (12b)–(12g) can be identified accordingly. In fact, if the correlation between ε_1 and each of $\varepsilon_2, \ldots, \varepsilon_7$ is equal to zero, then we can estimate the seven equations separately. Now, if the correlation between ε_1 and any of $\varepsilon_2, \ldots, \varepsilon_7$ is not equal to zero, then $Rank_2$ will be endogenous, and estimating these equations separately will lead to inconsistent estimates. We give a simple example in the following. Suppose that ε_1 and ε_2 have a nonzero correlation. Then in Equation (12b), $Rank_2$ will be correlated with ε_6 because $Rank_2$ is correlated with ε_1 and ε_1 is correlated with ε_6. The way to account for this endogeneity problem is to simultaneously estimate Equations (12a)–(12g). Because we are not able to predict the correlation structure in the proposed simultaneous equations model, i.e., do not know which correlation is zero and which correlation is not zero, we estimate the full covariance matrix and let the data inform us.

As shown in Lahiri and Schmidt (1978) and discussed in Greene (1999), a triangular system of simultaneous equations can be identified without identification constraints such as nonlinearity or correlation restriction. In particular, the identification of such a triangular system comes from the likelihood function. This is also noted by Hausman (1975), who observes that in a triangular system, the Jacobian term in the likelihood function vanishes so that the likelihood function is the same as for the usual seemingly unrelated regressions problem (Hausman 1975). Hence, a GLS (generalized least squares)- or SUR (seemingly unrelated regression)-based estimation leads to uniquely identified estimates in a triangular system with a full covariance on error terms as shown by Lahiri and Schmidt (1978).

Furthermore, we did a simulation analysis and found that our estimation procedure accurately recovers the true parameter values. This suggests that identification is not a problem.

4. An Empirical Application

In this section, we present an empirical application of the proposed model using a unique panel data set of aggregate keyword-level data on clicks and conversions collected from a Fortune 500 firm that advertises on Google. We first describe the data-generating process and the data used in the estimation, then discuss our empirical findings, and finally conduct policy simulations based on the proposed model and parameter estimates.

4.1. Data

When an Internet user enters a search query into a search engine, he gets back a page with results containing both the organic links most relevant to the query and the sponsored links, i.e., paid advertisements that are ranked sequentially by the search engine. The serving of a text ad in response to a query for a certain keyword is denoted as an impression. If the consumer clicks on the ad, he is led to the landing page of the advertiser's website. This is recorded as a click, and advertisers pay the search engine on a per-click basis. This is known as the cost per click. In the event that the consumer ends up purchasing a product from the advertiser, this is recorded as a conversion.

Our data contain weekly information on paid search advertising from a large nationwide retail chain that advertises on Google and is similar to the data used in Ghose and Yang (2009).[11] The data span all keyword advertisements by the company during a period of three months in the first quarter of 2007, specifically for the 13 calendar weeks from January 1 to March 31. The data are based on an "exact match" between the user query and sponsored ad. Note that search engines only provide aggregate-level daily or weekly data to advertisers. The use of exact match (instead of "broad" or "phrase" match) prevents any concern from possible aggregation biases arising as a result of the absence of data from every single auction that occurred in a given week or in a given day. Moreover, the firm providing us the data for this study

---

[11] The firm is a Fortune 500 firm with a strong national and international presence, but because of the nature of the data-sharing agreement between the firm and us, we are unable to reveal the name of the firm.

had confirmed that for an overwhelming majority of the keywords in our sample, there was very little variation in the number of competitors for a given keyword across the time period of our data. This feature also minimizes the impact of competition on the extent of variation in the rank for a given keyword, further alleviating any concern of biases from data aggregation.

The data consist of keyword ads from all six categories of products that this nationwide chain retailer sells (bedding, bath, dining, kitchen, electronics, and home décor). These keyword ads encompass all the 40 departments subsumed within these six product categories. Between them, the keywords represent several hundred unique stock-keeping units. We have 106 unique brands represented by these keywords, and for the same advertiser we have several different combinations of its name, each represented by a unique keyword.

Each keyword in our data has a unique advertisement ID. For any given keyword search query, the retailer is linked through to both organic and paid search results on the search engine. In a given search session, whereas the majority lead to clicks on either organic or paid (or no clicks at all), some searches lead to clicks on both organic and paid listings during the same search session. We obtained the data on the total number of searches for a given keyword in a given week from the Google Keyword Metrics Tool.[12] The data provided had information on the number of clicks on "paid listings only," number of clicks on "organic listings only," and number of clicks on "both organic and paid listings" for any given keyword on a daily basis that was aggregated by the advertiser to a weekly basis. The number of searches that do not lead to any clicks is calculated by subtracting the sum of the aforementioned three types of clicks from the total number of searches. Similar to data on conversions through paid search, we have data on conversions through organic listings for any given keyword in the same week. Note that these are all aggregate keyword-level data, not disaggregate user-level data.

Data on organic search rankings for the same set of keywords were obtained based a Web crawler that gathered information on where the advertiser's link would appear on Google's organic listings. The Web crawler was constructed in PERL. To get a more precise estimate of the rank of the organic listing of the advertiser, we retrieved these data once every week over a six-week period from Google.[13] Because there was very little change in the rank of the organic listing (the standard deviation across ranks for a given keyword was very low), we used the weekly average rank of a keyword in the organic listing for the purpose of our analysis. Finally, to control for competitive bid prices in our estimations, we collected data from Google's Keyword Pricing Tool, which gives estimates of advertisers' maximum cost per click for any given keyword. Google's keyword estimator tools give two key pieces of information: the estimated upper range and lower range for the cost per click of that keyword (roughly corresponding to the price of appearing ranked first and third on the sponsored links related to that keyword). We take the average of these two values to construct the *Competitor_Price* variable.

Then for both paid and organic listings, we have information on the number of clicks, number of conversions, the total revenues from a conversion for a given keyword for a given week, the average *CPC* in paid search, the maximum *CPC* for a given keyword across competitors, and the rank of the keyword. Although a search can lead to an impression, and often to a click, it may not lead to an actual purchase (which we define as a conversion). The product of *CPC* and *number of clicks* gives the total costs to the firm for sponsoring a particular advertisement. We have data on the contribution margin of each keyword based on the subproduct category (department) that it represents. Based on the contribution margin and the revenues from each conversion through a paid search advertisement, we compute the gross profit per keyword from a paid search conversion. The difference between gross profits and keyword advertising costs (the number of clicks times the cost per click) gives the net profits accruing to the retailer from a sponsored keyword conversion. This gives us the *Paid_Profit* variable. Similarly, the *Organic_Profit* variable is computed based on the contribution margin of the keyword and the revenues from each conversion through an organic search listing.[14]

Our sample only includes those keywords for which we had access to data on total weekly searches for those keywords on Google. This resulted in a data set with 426 unique keywords and a total of 1,400

---

[12] This tool is available at http://www.technobloggie.com/keyword-tool/index.php and retrieves both daily and monthly searches for a given keyword on Google. We computed the weekly number based on the daily data. As a robustness check, we also computed the weekly numbers based on the monthly data and find that it yields similar results. We used the total number of searches to compute the click-through rates because we do not have data on the total number of impressions on natural listings for a given keyword.

[13] The Web crawler performed searches with the exact keyword that we have in our data to search for the page number and rank of the organic link of the advertiser that is retrieved by the search engine.

[14] The *CPC* for clicks on organic listing is always zero. We also ran all the estimations without factoring in the contribution margin of the different keyword. All our results are robust to the use of gross margins only.

Table 1   Summary Statistics of the Data ($N = 1{,}400$)

Variable	Mean	Std. dev.	Min	Max
Total searches	5,400.035	19,250.420	7	165,361
Paid clicks	98.509	991.143	0	33,330
Organic clicks	7.017	50.353	1	1,355
Both clicks	12.708	125.022	0	4,200
No clicks	5,281.802	19,185.500	0	165,265
Paid conversions	1.574	15.680	0	527
Organic conversions	0.095	0.536	0	12
Cost per click (*CPC*)	1.211	0.210	1.003	2.72
Competitor_Price	1.165	0.893	0.165	7.04
Lag_Cost per click	1.224	0.20	1.001	2.72
Paid rank	4.301	7.208	1	63
Organic rank	10.860	15.988	1	100
Lag_Paid rank	4.574	7.038	1	63
Lag_Organic rank	10.860	15.988	1	100
Log(*Paid profit*)	0.762	2.385	−4.859	10.711
Log(*Lag_Paid profit*)	0.652	2.163	−4.859	8.945
Log(*Organic profit*)	0.251	1.061	0	7.392
Log(*Lag_Organic profit*)	0.265	1.022	0	7.392
Paid click-through rate	0.066	0.149	0	0.988
Organic click-through rate	0.028	0.047	0.001	0.571
Lag_Paid click-through rate	0.061	0.116	0	0.945
Lag_Organic click-through rate	0.025	0.042	0.001	0.571
Length	2.363	0.870	1	5
Retailer	0.188	0.391	0	1
Brand	0.656	0.475	0	1

observations. Table 1 reports the summary statistics. Not surprisingly, a majority of the keywords have brand-specific information (65%), whereas only 18.8% of the keywords have retailer-specific information. Interestingly, we note that the mean click-through rate was 6.6% and 2.77%, respectively, from paid and organic searches. The mean conversion rate was 5.71% and 1.67%, respectively, from paid and organic searches. Finally, note that the mean profit from paid search advertisements was 3.1 times higher than that from organic search listings.

4.2.   Empirical Findings

First note from Table 2 that all three keyword-specific characteristics (*Retailer*, *Brand*, and *Length*) significantly predict the search volume. Specifically, retailer-specific keywords are associated with a higher volume of searches, whereas keywords containing information on brands (manufacturer or product) are associated with a lower volume of searches. Moreover, the volume of searches decreases with an increase in the length of the keyword, i.e., an increase in the specificity of the search.

From Tables 3(a) and 3(b) we see that the average magnitude of interdependence (the parameter *θ*) between paid clicks and organic clicks is positive and statistically very significant. Therefore, a higher probability (number) of clicks on organic listings is correlated with a higher probability (number) of clicks

Table 2   Results on Search Volume

	Response estimates	Heterogeneity estimate
Intercept	**5.74**	**3.71**
	(0.12)	**(0.25)**
Retailer	**1.45**	
	(0.26)	
Brand	**−0.23**	
	(0.10)	
Length	**−0.55**	
	(0.00)	
Time	0.001	
	(0.002)	

Note. Posterior means and posterior standard deviations (in the parentheses) are reported, and estimates that are significant at 95% are in bold in Tables 2–6.

Table 3(a)   Results on Paid Click-Throughs

Intercept	Rank	Retailer	Brand	Length	Time	Utility_Organic
−0.25	**−0.02**	−0.13	0.11	−0.10	−0.02	**0.98**
(0.45)	**(0.01)**	(0.32)	(0.24)	(0.13)	(0.03)	**(0.09)**

$\Sigma_2^{\beta\theta}$	$\beta_{Intercept}$	β_{Rank}	θ^{21}			
$\beta_{Intercept}$	**1.17**	0.01	**0.21**			
	(0.43)	(0.05)	**(0.09)**			
β_{Rank}		**0.11**	0.01			
		(0.02)	(0.01)			
θ^{21}			**0.13**			
			(0.03)			

Table 3(b)   Results on Organic Click-Throughs

Intercept	Rank	Retailer	Brand	Length	Time	Utility_Paid
−3.58	**−0.06**	**−0.58**	**0.29**	**0.17**	0.01	**0.28**
(0.26)	**(0.02)**	**(0.21)**	**(0.11)**	**(0.05)**	(0.02)	**(0.05)**

Σ^{β}	$\beta_{Intercept}$	β_{Rank}	θ^{12}			
$\beta_{Intercept}$	**0.69**	**0.02**	**0.11**			
	(0.28)	**(0.01)**	**(0.05)**			
β_{Rank}		**0.04**	−0.01			
		(0.01)	(0.01)			
θ^{12}			**0.08**			
			(0.01)			

on paid search ads, and vice versa.[15] This highlights that the presence of organic search listings has a positive association with the average click-through rates in paid search advertisements, and vice versa. We also find that the magnitude of this positive interdependence between paid search and organic search is asymmetric. On average, the average impact of

---

[15] The three characteristics of a keyword (*Retailer*, *Brand*, and *Length*) are all mean centered, and hence the intercept can be viewed as the mean effect.

organic clicks on increases in the utilities of a paid click (0.98) is 3.5 times stronger than the average impact of paid clicks on increases in organic clicks (0.28). In addition, from Equations (3a)–(3d), we also computed the individual probabilities $p_{it}^{1,0}$, $p_{it}^{0,1}$, and $p_{it}^{1,1}$ and the lift in probabilities from a one unit increase in utility based on the empirical estimates and the summary statistics. We found that one unit increase in paid utility increases organic click-through probability by 1.25 times more than vice versa.

As predicted, *Rank* has an overall negative relationship with *CTR* in both paid and organic listings as seen in Table 2. The position of the advertisement link on the search engine results page clearly plays an important role in influencing click-through rates. This kind of primacy effect has also been seen in other empirical studies of the online world (Ansari and Mela 2003, Brynjolfsson et al. 2004, Rutz and Bucklin 2007, Baye et al. 2009, Ghose and Yang 2009). Interestingly, the magnitude of the effect of rank on click-through rates is different for paid searches from organic searches. The magnitude of the *Rank* coefficient is smaller (−0.02) for paid searches than for organic searches (−0.06), suggesting that keyword position on the screen plays a relatively more important role in influencing clicks in organic search compared to clicks in paid search.

We find that neither the presence of a brand name nor the retailer's own name in the search keyword has a statistically significant effect on click-through rates in paid listings. However, in the case of organic listings we find that the coefficient of *Brand* is positive and significant (0.29), whereas that for *Retailer* is negative and significant (−0.58). The coefficient of *Length* is significantly positive in organic search (0.17), suggesting that longer keywords that typically represent more goal-oriented searches for specific products tend to experience higher click-through rates on organic listings. As shown in Table 2, many of the estimated variances (unobserved heterogeneity) of the intercept, the interdependence effect, and the *Rank* coefficient are significant in both organic and paid search click-through probabilities. This suggests that the baseline click-through rates and the way that keyword ranking predicts the click-through rates are different across keywords, driven by factors beyond the three observed keyword characteristics.

Next, consider Tables 4(a) and 4(b) with findings on conversion rates. We find that *Rank* has a negative relationship with conversion rates from both paid and organic searches but is statistically significant only for organic listings. This implies that the lower the *Rank* (i.e., higher the position of the sponsored listing on the screen), the higher is the *Conversion Rate* in organic search. On the other hand, CTR has a positive effect on the *Conversion Rate* in paid search but a

Table 4(a)   Results on Paid Conversions

Intercept	Rank	CTR	Retailer	Brand	Length	Time
−5.35 **(0.33)**	−0.04 (0.05)	**0.11** **(0.04)**	**0.76** **(0.33)**	−0.16 (0.42)	−0.05 (0.11)	−0.02 (0.04)
Σ^β	$\beta_{\text{Intercept}}$	β_{Rank}				
$\beta_{\text{Intercept}}$	**3.87** **(0.91)**	**1.19** **(0.33)**				
β_{Rank}		**0.52** **(0.12)**				

Table 4(b)   Results on Organic Conversions

Intercept	Rank	CTR	Retailer	Brand	Length	Time
−8.49 **(0.72)**	**−0.19** **(0.05)**	−0.14 (0.11)	−0.83 (0.81)	**1.03** **(0.35)**	−0.57 (0.38)	−0.06 (0.08)
Σ^β	$\beta_{\text{Intercept}}$	β_{Rank}				
$\beta_{\text{Intercept}}$	**1.36** **(0.39)**	0.04 (0.07)				
β_{Rank}		**0.14** **(0.02)**				

statistically insignificant effect in organic search conversion rates. Overall, this suggests that there is an indirect effect of *Rank* on the conversion probability for paid search (through its effect on click-through rate) but a direct effect for organic search. Our analysis also reveals the coefficient of *Retailer* is positive and significant for conversions in paid search but statistically insignificant for organic search. As shown in Tables 4(a) and 4(b), many of the estimated variances (unobserved heterogeneity) for the intercept and the *Rank* coefficient are significant in both organic and paid search conversion rates. This suggests that the baseline conversion rates and the way that keyword ranking predicts the conversion rates are different across keywords, driven by unobserved factors beyond the observed characteristics.

The analysis of *CPC* in Table 5 reveals that *Brand* has a statistically significant and negative effect on keyword cost per click. The coefficient of *Lag_Paid Rank* is negative and statistically significant whereas that for *Lag_Organic Rank* is negative but statistically insignificant. Similarly, the coefficient of *Lag Profit* is statistically significant only in the case of organic search listings. These findings suggest that whereas past performance metrics (lagged paid rank and lagged organic profit) are being incorporated by the firm prior to bidding for keywords, it may not be bidding in an optimal manner. This is further evident in the negative correlation of the *Brand* dummy with *CPC*. Given that brand-specific keywords are more competitive, one would expect this association to be the other way around.

Table 5   Estimates of the Equation of ln(CPC) for Paid Searches

	Response estimates	Heterogeneity estimate
Intercept	**0.162**	**0.047**
	(0.009)	**(0.021)**
Lag_Paid rank	**−0.071**	
	(0.022)	
Lag_Organic rank	−0.028	
	(0.019)	
Log(Lag_Paid profit)	0.002	
	(0.002)	
Log(Lag_Organic profit)	**−0.003**	
	(0.001)	
Retailer	0.001	
	(0.001)	
Brand	**−0.017**	
	(0.003)	
Length	−0.004	
	(0.006)	
Competitor_Price	0.005	
	(0.016)	
Time	−0.001	
	(0.001)	

Table 6   Estimates of the Equation of Log(Rank) for Paid Searches

	Response estimates	Heterogeneity estimate
Intercept	**1.51**	**0.84**
	(0.06)	**(0.11)**
CPC	**−4.19**	
	(0.27)	
Lag_Paid click-through rate	**−0.18**	
	(0.05)	
Retailer	**−1.01**	
	(0.15)	
Brand	**−0.47**	
	(0.13)	
Length	−0.13	
	(0.09)	
Competitor_Price	**0.33**	
	(0.08)	
Time	**0.07**	
	(0.01)	

On the analysis of keyword rank in Table 6, we find that *Brand* and *Retailer* have a statistically significant and negative relationship with *Rank*, suggesting that keywords that have brand-specific information or retailer-specific information generally tend to be listed higher up on the screen. Both *CPC* and *Lag CTR* are statistically significant and inversely related to *Rank* as expected (see, for example, Ghose and Yang 2009). Furthermore, the effect of maximum bid price across all competing firms on *Rank* is also positive and statistically significant, thereby confirming that the extent of competition from other advertisers plays a key role in the sponsored search auctions' outcomes.

Finally, it is worth noting in Table 7 that the unobserved covariance between several variables is statistically significant. This suggests that keyword ranking is endogenous and a firm's bids for a given keyword are likely to be based on the same keyword's past performance.

4.3. Robustness Checks: Alternate Model Specifications

To demonstrate the robustness of our main results using the integrated model, we also explore two alternative model specifications. In particular, we build two nonnested models. First, we build and estimate an autologistic model to examine the sign of the interdependence between paid and organic click-throughs. Thereafter, we explore the nature of this interdependence to examine if there is some kind of asymmetry in the relationship using a simultaneous-move game structural model. For brevity, we only provide a high-level description of the models below. Detailed

descriptions along with the estimates are given in the electronic companion.

4.3.1.   Autologistic Model. The autologistic model that is used to model the relationship between consumer click-throughs on paid and organic links has been adopted in the marketing literature (for example, Moon and Russell 2008). The main idea behind the autologistic model is to allow the click-through action to be interdependent on each other instead of the latent utility as in our main model. More specifically, the autologistic model starts with a specification on the conditional distribution of one event such as a click-through on a paid or organic listing. Then based on Besag's (1974) theorem (also known as Brook's lemma), this conditional specification leads to a proper and well-defined joint distribution of the click-through probabilities of paid and organic search, given that the interdependence effect is symmetric. The intrinsic utility functions for paid and organic clicks are a function of the different keyword-level covariates and other factors that determine potential benefits to the consumer from a click.

We find that the average magnitude of interdependence (the interdependence parameter θ) between paid clicks and organic clicks is significantly positive. Therefore, a higher probability (number) of clicks on organic listings is correlated with a higher probability (number) of clicks on paid search ads, and vice versa. The results of the click-through estimates are given in Tables A.1(a)–A.1(c) of the electronic companion.

4.3.2. Simultaneous-Move Game Structural Model. Given the positive interdependence we find between paid and organic clicks in our data using the autologistic specification, we further explore the presence of an asymmetric effect in this relationship by adopting a structural approach (Bresnahan and Reiss 1991).

Table 7 **Cross-Equation Covariance Matrix Estimate**

	Search volume	Paid_ Clicks	Organic_ Clicks	Paid_ Conversions	Organic_ Conversions	CPC	Paid_Rank
Search volume	**3.84** (**0.31**)	−1.09 (0.93)	**0.52** (**0.11**)	0.21 (0.36)	**1.03** (**0.35**)	−0.35 (0.39)	0.63 (0.42)
Paid_Clicks		**3.87** (**0.24**)	**1.33** (**0.08**)	**0.67** (**0.10**)	−0.46 (0.28)	**0.42** (**0.08**)	**0.65** (**0.19**)
Organic_Clicks			**0.79** (**0.04**)	−0.19 (0.17)	**0.98** (**0.21**)	−0.23 (0.27)	0.11 (0.14)
Paid_Conversions				**0.41** (**0.09**)	**0.31** (**0.16**)	−**0.28** (**0.04**)	0.39 (0.28)
Organic_Conversions					**4.03** (**1.05**)	0.12 (0.25)	−0.27 (0.19)
CPC						**0.026** (**0.001**)	−**0.16** (**0.04**)
Paid_Rank							**1.16** (**0.07**)

As shown in Bresnahan and Reiss, we need to know the signs of the interaction effects θs to solve for the equilibrium in this structural modeling approach. For example, θs are assumed to be negative in a discrete entry game by assuming that competitors' entry tends to lower the company's profit at entry; θs are assumed to be positive in studying social interactions or peer effects among consumers. As shown in Bresnahan and Reiss, a violation of this assumption can lead to the absence of an equilibrium. Given the positive interdependence we find between paid and organic clicks in our data using the autologistic specification, we can explore the existence of an asymmetric interdependence by constraining the interaction effects to be positive. We use the hierarchical Bayesian method to estimate this model. As before, we find evidence of a strong asymmetric effect. On average, the effect of organic search click-throughs on paid ad clicks is 3.1 times the effect of paid ad click-throughs on organic search clicks. The results of the click-through estimates are given in Tables A.2(a)–A.2(c) of the electronic companion.

We have also run multivariate regressions with panel data methods such as fixed effect models, random effect models, and Tobit regressions. Those analyses also show a strong positive interdependence between paid clicks and organic clicks and are thus very consistent with our current model. The regression model is specified in the appendix, and the results are in Table A.3 of the electronic companion.[16]

4.3.3. Both Interdependence and Independence. As a robustness check, we also explore an alternative specification that incorporates both interdependence ($\theta \neq 0$) and independence ($\theta = 0$) in modeling the paid and organic click-throughs. This is consistent with modeling structural heterogeneity in the marketing literature. In such a model the probability (likelihood) of one observation y_k can be written as

$$L_k = pL_k^1(y_k \mid \theta \neq 0) + (1-p)L_k^2(y_k \mid \theta = 0). \quad (13)$$

In this mixture specification, p is the point mass or probability of following the interdependence model and $1 - p$ is the point mass or probability of following the independence model. We estimate this model, using the Bayesian approach developed in Yang and Allenby (2000) for dealing with the structural heterogeneity. The qualitative nature of all our estimates remains the same. More importantly, our results showed that there is a point mass of 0.972 on the interdependence model ($L_k^1(y_k \mid \theta \neq 0)$). This provides strong evidence for confirming the validity of our original model that incorporates the interdependence specification only.

4.3.4. Out-of-Sample Prediction. To demonstrate the fit of our model, we conduct an out-of-sample prediction. Based on the mean absolute deviations (MAD), our results suggest that the proposed simultaneous equation model predicts better than the same model estimated equation by equation, suggesting the importance of accounting for the simultaneity. In addition, we find that our proposed model predicts substantially better than a naïve non-model-based forecasting approach (i.e., predicting with sample average), which emphasizes the need for a model.

[16] We ran further robustness checks such as adding (i) a dummy variable indicating if the organic listing was on the first page when the paid listing was on the first page and (ii) another specification if the organic listing of this advertiser appeared more than once in the first 10 pages of Google's search engine results. We also ran additional robustness tests such as whether the paid listing for a given keyword appeared on the first page of the search engine results

page when the organic listing was on the first page. The qualitative nature of our main result of positive and asymmetric interdependence between the two forms of listings remains the same in all these analyses.

5. Policy Simulations and Managerial Implications

We aim to assess how much the advertiser benefits from the simultaneous presence of both paid and organic search listings. To this end, we first need to infer what the maximum profits for the advertiser are from both paid and organic search, which in turn requires that we infer the optimal cost per click for each keyword based on the empirical estimates from §4.2. The advertiser can determine the optimal *CPC* for each keyword to maximize the expected profit (Π):

$$\Pi_{it} = (p_{it}^{1,0})(q_{it1}m_{it1}) + (p_{it}^{0,1} + p_{it}^{1,1})(q_{it2}m_{it2} - CPC_{it}). \quad (14)$$

In Equation (17), p_{it} is the expected CTR for keyword i at week t and the superscripts indicate click-through rates on organic, paid, or both, consistent with Equation (2). q_{it} is the expected conversion rate conditional on a click-through, and m_{it} is the expected gross profit from a conversion that is observed from our data. Subscripts 1 and 2 in q_{it} and m_{it} indicate organic and paid search, respectively. CPC_{it} is the actual cost per click paid by the advertiser to the search engine for a given keyword. p_{it}, q_{it}, and $Rank_{it}$ are predicted based on Equations (3a)–(3c), (8a)–(8b), and (10a)–(10b), respectively, using the estimates obtained from the proposed model. Note that this kind of analysis cannot be done by eyeballing the summary statistics of the data because it requires us to find the optimal profits of the advertiser based on imputing the optimal *CPC*.

We conduct the optimization routine to maximize the expected profit from each consumer impression of the advertisement for each keyword in each week, using the grid search method. Our simulation results highlight that there is a considerable difference in the optimal *CPC* and the actual *CPC* incurred by the firm for a given keyword. Furthermore, there is also a difference between optimal expected profits and actual profits accruing to the firm from the current *CPC*. These results show that the firm was not bidding optimally during the time period of our data. For the majority of the keywords, we saw evidence of overbidding by the firm. Note that a similar finding regarding the suboptimal bidding behavior was discussed in Ghose and Yang (2009), who used a similar data set from the same firm.

We conduct a counterfactual experiment to infer the magnitude of the positive interdependence between organic search and paid search, and vice versa. To this end, we run the policy simulation in the absence of the cross-advertising effect (θ) parameter. That is, we set the cross-advertising effect parameter to zero and then calculate the optimal *CPC* and the expected profit, given the optimal *CPC*. We find that the expected profit in the presence of the θ parameter is 4.25% higher than that in its absence.

To check for the robustness of these results, we also did a few simulations that are more specific. First, we reran the simulation using a sample consisting of only retailer-specific keywords. This is the set of keywords where we expect the advertiser to face the least amount of competition because only the advertiser is likely to bid on such keyword ads that prominently display its name.[17] Second, we reran the simulation using a sample consisting of branded and generic keywords only. This is the set of keywords where we expect the advertiser to face the most competition because all firms selling a similar portfolio of products are likely to bid for such ads. Indeed, our data suggest that the *CPC* for branded and generic keywords was 2.1 times that of retailer-specific keywords. Furthermore, the *CPC* of generic keywords was 1.2 times that of branded keywords.

Our findings are consistent with the main result. Using the "retailer-specific keywords" only, we find that the expected profit in the presence of the interdependence effect parameter is 6.15% higher than that in its absence. Using the "generic and branded keywords," we find that the expected profit in the presence of the cross-advertising effect parameter is 4.2% higher than that in its absence. Thus, the additional implication we can derive from these experiments is that the positive interdependence is the strongest in the case of the "least competitive" keywords (retailer-specific keywords) and weakest in the case of the "most competitive" keywords (brand-specific and generic keywords).

We also conducted simulations using different subsamples as robustness checks. Specifically, we looked at different ways to consider the cases where the paid link was not displayed on the first page in the event that this increase in profits was mostly for such links. Note that Google's sponsored search auction has no fixed format for the number of links that appear on the first page for any given search query. Anecdotal evidence suggests that the number of sponsored links on the first page could vary anywhere between 3 and 10 for a given keyword. Thus in our data we looked at cases where the rank of the keyword in the paid search auction was (i) more than 3, (ii) more than 5, and (iii) more than 10, and we ran the policy simulations separately on these subsamples. We found that even if the paid link did not appear on the first page, there was an increase in profits from the positive interdependence between paid and organic ranging from 2.6% to 4.2%.

---

[17] We verified this information from the Search Analytics tool of http://www.compete.com and found that on an average the total click-through share of the retailer-specific keywords belonging to this advertiser's direct competitors is only 1.8%. In other words, almost all referrals to the advertiser's website stemming from retailer-specific keywords originate from the firm's own sponsored advertisement.

6. A Field Experiment

Our empirical model and analysis in §4 predicts that the combined CTR from paid and organic search listings when both these options are available to users is going to be greater than or equal to the CTR from the scenario where there are only organic search listings available. To put it in terms of the model parameters, recall from Equations (3a)–(3d) that the combined CTR in the presence of the interdependence effect parameter θ is given by $p_{it}^{1,0} + p_{it}^{0,1} + p_{it}^{1,1}$. This expression is greater than or equal to $p_{it}^{1,0}$ in the absence of paid search, depending on the magnitude of searchers belonging to the $p_{it}^{0,1}$ and $p_{it}^{1,1}$ categories who shift to clicking on organic search links. Furthermore, the model predicts a positive interdependence between clicks on paid and organic listings. In this section, we describe a simple field experiment that was conducted to shed further light on these predictions from the empirical model.

Let Λ be the fraction of users who move from clicking on paid search to natural links when the firm pauses sponsoring ads. Intuitively, this shift occurs because users who would have clicked on paid links would only have the option of clicking on organic links when the firm does not engage in paid search. Based on the model it is not obvious if all users of firm's paid search ads would migrate to click on its organic links ($\Lambda = 1$) or only a fraction of them would do so ($\Lambda < 1$). If it turned out that $\Lambda = 1$, that would imply that the combined CTR was the same as the CTR when only organic listings are available, and one would question the value to the firm from engaging in sponsored searches given the costs involved in paid search advertising. Perhaps even more importantly, if the total revenues in the presence of both paid and organic search listings are not significantly higher than the revenues when only organic search listings are present, the potential value from sponsored searches would be rather murky.

A second related question would be to analyze how the click-through rate in organic listings specifically would change if the advertiser moved from sponsoring paid searches to turning the ads off. In other words, we are interested in comparing the click-through rate of organic searches in the "both organic and paid" scenario versus the "only organic" scenario to examine the effect of paid ads on the CTR of organic listings per se. Our model suggests that this would be a function of the respective CTR probabilities and the fraction of people who migrate from clicking on paid ads to clicking on organic listings for a given keyword.

To be precise, we derive this cutoff in the following way. Recall from Equations (8a)–(8d) that in the presence of a paid ad, the total probability of clicking on organic listings is $p_{it}^{1,0} + p_{it}^{1,1}$. In the absence of paid ads, the total probability of clicking on organic listings is $p_{it}^{1,0} + \Lambda(p_{it}^{0,1} + p_{it}^{1,1})$. This implies that the organic CTR will be higher or lower in the absence of paid search depending on a critical value of Λ given by $p_{it}^{1,1}/(p_{it}^{0,1} + p_{it}^{1,1})$. Whether this organic CTR would be higher or lower is therefore an empirical question.

To investigate these two issues, a field experiment was designed that sheds further light on the impact of the simultaneous presence of paid search and organic listings on the combined performance from these two listings. This experiment was conducted over an eight-week period from mid-March to mid-May in 2007 during which the firm pulsated between periodically sponsoring some keywords and then halting the process. Specifically, a sample of 90 keywords was randomly selected by the firm to conduct this experiment. Then the firm sponsored these keyword ads for a two-week period on Google and tracked the results from the organic and paid search advertisements. Then the firm paused the sponsored advertisements for the next two weeks and tracked only the results from the organic listings. Then it resumed sponsoring the same set of keywords again for the next two weeks and then paused the process for the remaining two weeks. During these pulsing periods, the firm measured both sponsored search and organic performance using different metrics such as click-through rates, conversion rates, and revenues.

Based on the analysis of this field experimental data, we found that when paid search advertising was active implying that both sponsored and organic listings were available to consumers, the combined CTR from both of these listings was 5.1% higher than when paid search advertising was inactive, and only the organic listings were present (see Figure 1(a)). A two sample *t*-test reveals that the difference is statistically significant at the 1% level. It is worth noting that for a vast majority of the keywords, we found that the CTR of organic listings in the both organic and paid scenario was higher than its CTR when paid search was inactive.

After verifying that paid search keywords provide additional visitors to the advertiser's website, it was important to monitor the quality of traffic driven by paid search keywords. We find that although paid keywords are running alongside the organic links, the combined conversion rate is higher than when the organic links stand alone on the search engine results page. When paid ads are active, the combined conversion rate from both paid and organic links was 1.53%. Thus, there is an 11.7% increase in the combined conversion rate when paid and organic links are present simultaneously relative to when only organic listings are present (Figure 1(b)). Furthermore, when paid ads are paused, the conversion rate from organic search

Figure 1(a)   Plot from the Field Experiment Showing Combined CTR from Paid and Organic Search in the Periods When Paid Search Advertising Was On and When Paid Search Advertising Was Paused

Total CTR from paid and organic search

	1	2
Total CTR	25.20%	20.10%
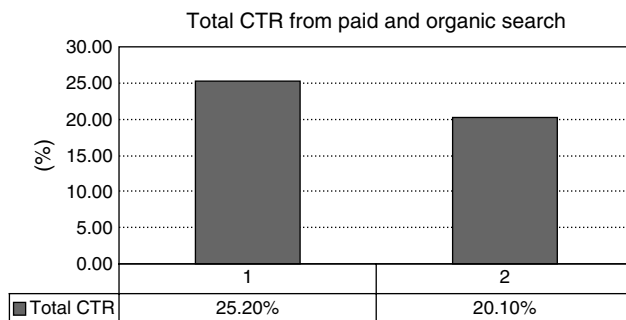
Figure 1(b)   Plot from the Field Experiment Showing Combined Conversions from Paid and Organic Search in the Periods When Paid Search Advertising Was On and When Paid Search Advertising Was Paused

Total conversion rates from paid and organic search

	1	2
Total conversion rates	1.53%	1.37%
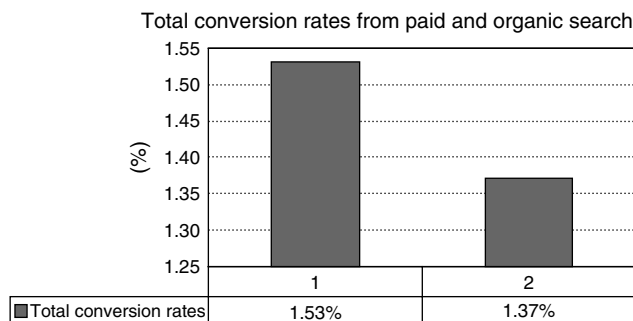
Figure 1(c)   Plot from the Field Experiment Showing Combined Revenues from Paid and Organic Search in the Periods When Paid Search Advertising Was On and When Paid Search Advertising Was Paused

Total revenues from paid and organic search

	1	2
Site revenue	$682,126.97	$435,301.77
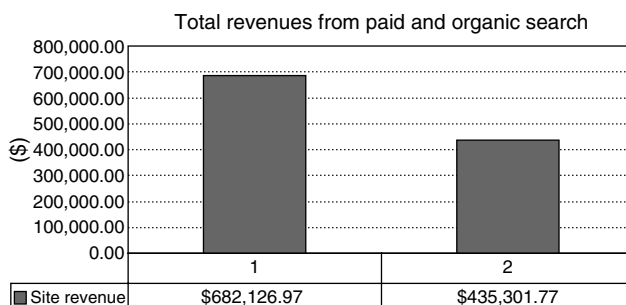
was 1.37%. It is worth noting that, for a vast majority of the keywords, the conversion rate of organic listings in the both organic and paid scenario was higher than its conversion rate when paid search was inactive.

With respect to revenues, we observed peaks and valleys in performance as and when the paid ads were pulsed. Because the firm tracked revenues from paid and organic searches separately, we were able to verify that although organic keywords made up a por-

tion of the lost revenue when paid search keywords were inactive, organic listings alone did not provide the full value of having both paid and organic search simultaneously available to users. Through the analysis of these data, we found that when paid search advertising is active, it drives an additional 54% incremental revenue lift from $435K to $682K in total revenue (Figure 1(c)). A *t*-test reveals that this difference increase is statistically significant. The substantially higher increase in combined revenues compared to the increase in combined click-through and conversion rates suggests that users tend to buy a greater number of products or buy higher-value items when paid search is active. This highlights the strong business potential of paid search.

To further evaluate the consistency of results between the field experiment and the estimated model, we conduct some panel data analysis using the specific sample of keywords for the duration for which the experiment was run with paid search advertisements on. We regress paid click-through rates against organic click-through rates after controlling for the *Rank* of the keyword in paid search, and vice versa. In particular, we estimate ordinary least-squares (OLS) regressions with keyword-level fixed effects. We also estimate the same regressions with keyword-level random effects and find that the estimates are very similar in magnitude and direction. These regressions predict that paid click-through has a statistically significant and positive relationship with organic click-through, and vice versa. In addition, the effect is also asymmetric. Based on the magnitude of the coefficients, we find that the effect of organic click-through on paid click-through is significantly higher than the effect of paid click-through on organic click-through. Note that this is consistent with our finding from the econometric model, including the alternate ones conducted for robustness purposes.

7.   Discussion, Implications, and Limitations

We build a model that integrates consumer searches and their reactions to the organic listings and sponsored ads associated with these searches with the advertiser's decision on cost per click and the search engine's decision on keyword ranks. Our goal is to estimate the interdependence between organic and paid ads. The model is general and can be directly applied to many Internet firms who sell their products on the Internet and advertise their products via search engines. Our data are also unique and carefully screened to prevent any possible concerns from the use of aggregate-level data.[18]

[18] As mentioned in §4, these features in the data include the use of "exact match" to display ads and the use of a sample with very

We show that the presence of organic listings is associated with a higher probability of click-throughs on paid ads, and vice versa. This suggests that firms, which tend to rank highly in organic search, are more likely to benefit from sponsored search advertising. In this regard, our finding that paid and organic listings have a positive interdependence on each other's click-through rates underscores that both search engine optimization (SEO) and search engine marketing (SEM)—along with other marketing channels—have a place in online customer acquisition campaigns. Thus, these results can have useful implications for a firm's marketing mix strategies. This empirical finding is consistent with claims in the trade press that more people will visit a website if it is listed in both paid and organic listings because there is a "second opinion effect." This happens because searchers are encouraged by the fact that a website is listed in both the organic and paid listings leading to higher click-through rates.[19] Another reason could be related to the quality of the link and the resultant consumer satisfaction from a click on each type of link. This is possible because the landing pages associated with a keyword might be different for organic and paid links, and the right landing page on the organic listing could be strategically decided by the search engine. In a game-theoretic model, Taylor (2009) shows that by adjusting the quality of organic links and thereby affecting the level of consumer satisfaction from clicking on a link, search engines can induce consumers to click on both organic and paid links, leading to a positive interdependence between organic and paid listings. White (2008) also explores the costs and benefits to a search engine of providing Internet users with high-quality organic search results in addition to showing them paid advertisements and discusses similar effects on user behavior.

Furthermore, we find that the relationship is asymmetric such that the impact of organic clicks on an increase in utility from paid clicks is on an average 3.5 times stronger than vice versa. From a search engine's perspective, this positive and asymmetric interdependence between paid and organic listings also implies that the top-ranking websites in organic search are likely to get a higher number of clicks in paid search as well. Because advertisers pay search engines on a per-click basis, this has implications for search engines' revenues. Indeed, there may be a moral hazard problem here as search engines may have an incentive to manipulate rankings in organic search and selectively present those firms on the

top in organic search that experience higher click-through rates in paid search. Search engines claim there is no direct linkage between sponsoring ads and organic ranking, but trade press reports speculate that there could be perverse incentives for search engines.[20] Google creates algorithms that generate organic search results based on indexing criteria such as relevance, PageRank, and the presence of user-generated content (UGC). Even though sponsored ads do not count towards the link popularity of an advertiser in the organic listings, there are other ways to tie together paid ads and organic listings. For example, Google has begun to serve organic search results based on user profiles in its recently developed personalized search results. Websites that users have already visited will usually rank higher on subsequent queries if users have that feature enabled. A potential implication is that a firm might pay to have a paid listing for the most generic of terms because a click on a paid ad helps it rank higher on the organic listings in subsequent searches as the user gets closer to the purchase. Thus, there could be deleterious effects of interactions between paid and organic listings. Another example of an intricate relationship between paid and organic listings is through UGC. Whereas Google ranks landing pages with UGC higher in its organic search, the presence of UGC can also increase the landing page quality scores. This highlights the importance of having UGC on websites for improving rankings in both paid and organic listings. On the other hand, selective presentation of paid links or organic links could also improve consumer utility by reducing the cognitive costs associated with evaluating different alternatives. This has been shown by prior research on information gatekeepers like shopbots (Montgomery et al. 2004) whose infrastructure for paid placement of retailers (pay per click) is very similar to that of search engines. This calls attention for the need for designing newer mechanisms that can preserve the integrity of organic search rankings while still increasing user welfare during search.

In most search-based advertising services, a company sets a daily budget, selects a set of keywords, determines a bid price for each keyword, and designates an ad associated with each selected keyword. With millions of available keywords and a highly uncertain click-through rate associated with

---

little variation in the number of competitors for a given keyword. These features reduce the variation in ranks within a given day and thereby allow us to gain precise estimates using aggregate data.

[19] See Hartzer (2005).

[20] Google's official claim is that "it is very important to note that there is absolutely no connection between being an AdWords advertiser, and having your site appear in the unpaid search results. One does not affect the other in any way. To put it another way, being an AdWords advertiser will neither help nor harm your chances of appearing on the 'organic' search engine" (Search Engine Roundtable 2007). White (2008) also highlights that despite the regulatory eyebrows Google has been raising, there is remarkable silence over the incentives to manipulate the organic listings.

the ad for each keyword, identifying the most profitable set of keywords given the daily budget constraint often becomes challenging for firms wishing to promote their goods and services via search-based advertising (Rusmevichientong and Williamson 2006). The analysis of keyword covariates on average click-through rates and conversion rates can provide some guidance to practitioners on the profitability of choosing different keywords. Indeed, such techniques could nicely follow the broader keyword selection techniques based on popularity and economic impact of occurrence of keywords in user-generated content sites such as product review forums and blogs (Archak et al. 2007, Dhar and Ghose 2009).

Our estimates also suggest that firms may not be bidding optimally based on the relationship between cost per click and different keyword attributes, consistent with the findings of Ghose and Yang (2009). This can provide additional managerial implications for firms engaging in paid search advertising. Based on an optimization algorithm that imputed the expected profits based on the optimal *CPC* for each keyword for the advertiser, we find that as a result of the positive interdependence, the firm's profits in the simultaneous presence of paid and organic search listings is 4.5% higher compared with the scenario when there are either only paid advertisements or organic search listings. Furthermore, we find that the positive interdependence is the strongest in the case of the "least competitive" keywords (retailer-specific keywords) and weakest in the case of the "most competitive" keywords (brand-specific and generic keywords). Therefore, the proposed parsimonious modeling framework can help advertisers make better decisions about investments in sponsored search in the presence of organic listings in search engines.

Finally, we describe a field experiment that shows that total revenues and combined conversion rates in the presence of both paid and organic listings are higher compared to when only organic listings are present. By examining the CTR, conversion rates, and total revenues, this experiment further corroborates the beneficial effect of the simultaneous presence of organic and paid listings to advertisers. For many keywords, the click-through rate in organic listings is higher when paid and organic listings are simultaneously available compared to when the firm does not sponsor keyword ads. Furthermore, the overall effect on combined click-through rates, conversion rates, and revenues is significantly positive. From a managerial standpoint, what makes matters a little more subtle is that the conversion rates are higher on the paid listings. This is true both in the empirical analyses as well as in the field experiment. It is possible that users are self-selecting: searchers who are more likely to convert are more likely to click on the paid

listing. It is somewhat intuitive that people who are less likely to convert (information seekers, consumers early in the purchase process, or those with other noncommercial goals) are going to lean more toward clicking on organic listings rather than paid listings. This would naturally lead to a higher conversion rate via paid ads. It is also possible that the sponsored ads are written better to grab more targeted traffic and send users to better landing pages than organic listings.

Our results have some implications on how advertisers should invest in SEO, in which firms try to improve their ranking in organic search by fine tuning their landing pages versus SEM, in which firms try to improve their performance in paid search auctions. This can be important because many advertisers engage in both kinds of activity. Our data reveal that the conversion rate is significantly higher in paid search than in organic listings. This underscores the importance of securing a higher rank and designing effective landing pages by advertisers. On the other hand, our analysis suggests that most of the keyword-level characteristics have a stronger impact on the performance of organic search than paid search. For a well-rounded and effective search marketing campaign that reaches the greatest number of searchers, marketers should blend both organic and paid listings, capitalizing on the positive interdependence in clicks between them. These results could shed light on understanding how firms should invest in search engine advertising campaigns relative to search engine optimization and the proportion of the advertisement budget allocated to search advertising.

This paper has many important limitations that suggest that these results might best be viewed as starting points for further research. Some of the limitations have to do with the lack of information in our data. For example, we do not have data on competition—that is, we do not have information on the competitors of the specific firm whose data we have used in this paper. Hence, we are not able to control for the impact of the number of ads displayed in response to a single search query while estimating the relationship between paid and organic links. Although we use the maximum of the competitors' bid prices as the proxy for the level of competition for a given keyword, it is possible that we are overestimating the extent of interdependence for highly competitive ads. That said, the firm providing us the data for this study had confirmed that there was very little variation in the number of competitors for a given keyword across the time period of our data for the majority of the keywords. Nevertheless, future research could use richer data sets to address this issue.

We do not know whether the same consumer clicked multiple times on a given listing (for example, by using the browser's back button to go back to the search engine results page from the advertiser's page) or whether they clicked only once on each of these listings within a given search session. Knowledge of this issue can motivate another empirical framework that incorporates an incidence and count model to describe this phenomenon.

Another key limitation of our field experiment is that we are unable to control the presence of organic listings for a given keyword. It would be interesting to observe what happens to clicks and conversions on paid ads when organic search listings are absent, although such an experiment is only possible with the explicit cooperation of the search engine or by carefully designed laboratory experiments with human subjects.

An important avenue for future research is to investigate the impact of consumer heterogeneity in search advertising by adopting methods from recent methodological advances in Bayesian modeling (Rossi and Allenby 2003, Chen and Yang 2007, Musalem et al. 2009). We did not model this because our proposed model based on aggregate data is already very complicated because of the simultaneous and nonrecursive nature of the model. Future research could use individual consumer-level data from multiple advertisers as opposed to one advertiser to model the impact of consumer heterogeneity.

Future research could also examine data on the textual content in the copy of the ad (ad creative) corresponding to the different keywords to examine how textual content affects the results identified in this paper. This can be done using recent advances in text mining methods for quantifying the economic impact of textual content (Archak et al. 2007), although some anecdotal evidence suggests that the presence of the keyword in the title of the ad is more important than that in the ad copy in influencing clicks (Marketing Experiments 2005). Notwithstanding these limitations, we hope that this study will generate further interest in exploring this important emerging area in marketing.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mktsci.pubs.informs.org/.

Appendix. The MCMC Algorithm

We ran the MCMC chain for 40,000 iterations and used the last 20,000 iterations to compute the mean and standard deviation of the posterior distribution of the model parameters.

1. Draw $\pi_{it} = (\pi_{it1}, \pi_{it2})'$.

As specified, the likelihood function of the number of clicks is

$$l\left(N_{it}^{1,0}, N_{it}^{0,1}, N_{it}^{1,1}, N_{it}^{0,0}\right)$$
$$\propto (p_{it}^{1,0})^{N_{it}^{1,0}} (p_{it}^{0,1})^{N_{it}^{0,1}} (p_{it}^{1,1})^{N_{it}^{1,1}} (p_{it}^{0,0})^{N_{it}^{0,0}}.$$

The probabilities of the four actions (based on whether to click on a paid listing, and whether to click on an organic listing) conditional on the latent utilities πs, are as follows:

$$p_{it}^{1,0} = \frac{\exp(\pi_{it1})}{1 + \exp(\pi_{it1})} \cdot \frac{1}{1 + \exp(\pi_{it2})},$$

$$p_{it}^{0,1} = \frac{1}{1 + \exp(\pi_{it1})} \cdot \frac{\exp(\pi_{it2})}{1 + \exp(\pi_{it2})},$$

$$p_{it}^{1,1} = \frac{\exp(\pi_{it1})}{1 + \exp(\pi_{it1})} \cdot \frac{\exp(\pi_{it2})}{1 + \exp(\pi_{it2})},$$

$$p_{it}^{0,0} = \frac{1}{1 + \exp(\pi_{it1})} \cdot \frac{1}{1 + \exp(\pi_{it2})},$$

$$\pi_{its} = m_{its} + \eta_{its}, \quad s = 1, 2,$$

$$m_{it1} = \beta_{i11} + \beta_{i12} Rank_{it1} + \alpha_{11} Retailer_i + \alpha_{12} Brand_i$$
$$+ \alpha_{13} Length_i + \alpha_{14} Time_{it} + \theta_i^{12} \hat{\pi}_{it2},$$

$$m_{it2} = \beta_{i21} + \beta_{i22} Rank_{it2} + \alpha_{21} Retailer_i + \alpha_{22} Brand_i$$
$$+ \alpha_{23} Length_i + \alpha_{24} Time_{it} + \theta_i^{21} \hat{\pi}_{it1},$$

where $\hat{\pi}_{it1}$ and $\hat{\pi}_{it2}$ are predicted utilities from the reduced-form model on click throughs, generated in a separate MCMC chain in parallel to this algorithm. More specifically, the reduced form of our click-through model can be written as

$$\pi_{it1} = f(x_{it}, \kappa_{i1}, \gamma_1) + \xi_{it1},$$
$$\pi_{it2} = f(x_{it}, \kappa_{i2}, \gamma_2) + \xi_{it2},$$

where

$$x_{it} = (intercept, Rank_{it1}, Rank_{it2}, Retailer_i, Brand_i,$$
$$Length_i, Time_{it}).$$

Then

$$\hat{\pi}_{it1} = f(x_{it}, \hat{\kappa}_{i1}, \hat{\gamma}_1),$$
$$\hat{\pi}_{it2} = f(x_{it}, \hat{\kappa}_{i2}, \hat{\gamma}_2).$$

This approach follows the econometric literature for estimating similar models with endogenous regressors (Nelson and Olsen 1978, Maddala 1983, Bajari et al. 2006). The standard identification condition applies, that is, the predictors of m_{it1} are not exactly the same as the predictors of m_{it2}. This condition for identification is met in our empirical context because rank on the paid listings is different from the rank on the organic listings for the same keyword in a given time period.

Let us denote D and E_{it} as the conditional covariance matrix and mean vector of $(\eta_{it1}, \eta_{it2})'$, respectively, conditioning on values of $(\kappa_{it}, \mu_{it1}, \mu_{it2}, s_{it}, \nu_{it})'$ and Ω. We use Metropolis–Hastings algorithm with a random walk chain to generate draws of $\pi_{it} = (\pi_{it1}, \pi_{it2})'$ (see Chib and Greenberg 1995, p. 330, method 1). Let $\pi_{it}^{(p)}$ denote the previous draw, and then the next draw $\pi_{it}^{(n)}$ is given by

$$\pi_{it}^{(n)} = \pi_{it}^{(p)} + \Delta,$$

with the accepting probability α given by

$$\min\left[\frac{\exp[-1/2(\pi_{it}^{(n)} - m_{it} - E_{it})'D^{-1}(\pi_{it}^{(n)} - m_{it} - E_{it})]l(\pi_{it}^{(n)})}{\exp[-1/2(\pi_{it}^{(p)} - m_{it} - E_{it})'D^{-1}(\pi_{it}^{(p)} - m_{it} - E_{it})]l(\pi_{it}^{(p)})}, 1\right].$$

Δ is a draw from multivariate Normal$(0, 0.05I)$, where I is the identity matrix.

2. Draw $\beta_{i1} = (\beta_{i11}, \beta_{i12}, \theta_i^{12})'$.

Let us denote d and e_{it} as the variance and mean of η_{it1}, conditional on the values of η_{it2} and D.

$$w_{it} = \pi_{it1} - (\alpha_{11} Retailer_i + \alpha_{12} Brand_i + \alpha_{13} Length_i + \alpha_{14} Time_{it}) - e_{it}$$
$$= \beta_{i1} x_{it} + \eta_{it1},$$

where

$$x_{it} = (1, Rank_{it1}, \hat{\pi}_{it2})$$

$$\beta_{i1} \sim MVN(A_i, B_i); \quad A_i = B_i(\Sigma_1^{\beta\theta-1}\bar{\beta}_1 + x_i'w_i/d)^{-1} \quad \text{and}$$
$$B_i = (\Sigma_1^{\beta\theta-1} + x_i'x_i/d)^{-1}.$$

3. Draw $\beta_{i2} = (\beta_{i21}, \beta_{i22}, \theta_i^{21})'$ similar to Step 2.

4. Draw $\Sigma_1^{\beta\theta}$.

$$\Sigma_1^{\beta\theta} \sim IW\left(\sum_i(\beta_{i1} - \bar{\beta}_1)(\beta_{i1} - \bar{\beta}_1)' + Q_0, N + q_0\right)$$

$Q_0 = 10I$ and $q_0 = 10$; $N = $ no. of keywords.

5. Draw $\Sigma_2^{\beta\theta}$ similar to Step 4.

6. Draw $\bar{\beta}_1$.

$$\bar{\beta}_1 \sim MVN(A, B); \quad A = \sum_i \beta_{i1}/N \quad \text{and} \quad B = \Sigma_1^{\beta\theta}/N.$$

7. Draw $\bar{\beta}_2$ similar to Step 6.

8. Draw $\alpha_1 = (\alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{14})'$.

Let us denote d and e_{it} as the variance and mean of η_{it1}, conditional on the values of η_{it2} and D.

Let $w_{it} = \pi_{it1} - \beta_{i11} - \beta_{i12} Rank_{it1} - \theta_i^{12}\hat{\pi}_{it2} - e_{it} = \alpha_1 x_{it} + \eta_{it1}$, where

$$x_{it} = (Retailer_i, Brand_i, Length_i, Time_{it})$$
$$\alpha_1 \sim MVN(A, B)', \quad A = B(\Sigma_0^{-1}\bar{\alpha}_1 + x'w/d)^{-1},$$
$$B = (\Sigma_0^{-1} + x'x/d)^{-1}, \quad \bar{\alpha}_1 = 0, \quad \text{and} \quad \Sigma_0 = 100I.$$

9. Draw $\alpha_2 = (\alpha_{21}, \alpha_{22}, \alpha_{23}, \alpha_{24})'$ similar to Step 8.

10. Draw u_{it1}.

$$u_{it1} = \bar{u}_{it1} + \mu_{it1},$$
$$\bar{u}_{it1} = c_{i11} + c_{i12} Rank_{it1} + \gamma_{11} CTR_{it1} + \gamma_{12} Retailer_i$$
$$+ \gamma_{13} Brand_i + \gamma_{14} Length_i + \gamma_{15} Time_{it}.$$

The likelihood function is

$$l(M_{it}^1 \mid N_{it}^{1,0}, N_{it}^{0,1}, N_{it}^{1,1}, N_{it}^{0,0}) \propto (q_{it1})^{M_{it}^1}(1 - q_{it1})^{N_{it}^{1,0}+N_{it}^{1,1}-M_{it}^1},$$
$$q_{it1} = \frac{\exp(u_{it1})}{1 + \exp(u_{it1})}.$$

We use the Metropolis–Hastings algorithm with a random walk chain to generate these draws (see Chib and Greenberg 1995, p. 330, method 1). Let $u_{it1}^{(p)}$ denote the previous draw, and then the next draw $u_{it1}^{(n)}$ is given by

$$u_{it1}^{(n)} = u_{it1}^{(p)} + \Delta,$$

with the accepting probability α given by

$$\min\left[\frac{\exp[-1/2(u_{it1}^{(n)} - \bar{u}_{it1})^2]l(u_{it1}^{(n)})}{\exp[-1/2(u_{it1}^{(p)} - \bar{u}_{it1})^2]l(u_{it1}^{(p)})}, 1\right].$$

Δ is a draw from Normal$(0, 0.0025)$.

11. Draw u_{it2} similar to Step 10.

12. Draw $c_{i1} = (c_{i11}, c_{i12})'$.

Let us denote d and e_{ij} as the variance and mean of μ_{it1}, conditional on the values of $(\eta_{it1}, \eta_{it2}, \mu_{it1}, s_{it}, \nu_{it})'$ and Ω.

$$w_{it} = u_{it1} - (\gamma_{11} CTR_{it1} + \gamma_{12} Retailer_i + \gamma_{13} Brand_i$$
$$+ \gamma_{14} Length_i + \gamma_{15} Time_{it}) - e_{it}$$
$$= c_{i1} x_{it} + \eta_{it1},$$

where

$$x_{it} = (1, Rank_{it1}),$$
$$c_{i1} \sim MVN(A_i, B_i), \quad A_i = B_i(\Sigma^{c_1-1}\bar{c}_1 + x_i'w_i/d)^{-1}, \quad \text{and}$$
$$B_i = (\Sigma^{c_1-1} + x_i'x_i/d)^{-1}.$$

13. Draw $c_{i2} = (c_{i21}, c_{i22})'$ similar to Step 12.

14. Draw Σ^{c_1}.

$$\Sigma^{c_1} \sim IW\left(\sum_i(c_{i1} - \bar{c}_1)(c_{i1} - \bar{c}_1)' + Q_0, N + q_0\right),$$

$Q_0 = 10I$ and $q_0 = 10$, $N = $ no. of keywords.

15. Draw Σ^{c_2} similar to Step 14.
16. Draw $\bar{c}_1$.

$$\bar{c}_1 \sim \text{MVN}(A, B), \quad A = \sum_i c_{i1}/N, \quad \text{and} \quad B = \Sigma^{c_1}/N.$$

17. Draw $\bar{c}_2$ similar to Step 16.
18. Draw $\gamma_1 = (\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14}, \gamma_{15})'$.
Let $w_{it} = u_{it1} - c_{i11} - c_{i12} Rank_{it1} = \gamma_1 x_{it} + \mu_{it1}$, where $x_{ij} = (CTR_{ij1}, Retailer_i, Brand_i, Length_i, Time_{ij})$.

$$\gamma_1 \sim \text{MVN}(A, B)', \quad A = B(\Sigma_0^{-1} \bar{\alpha}_1 + x'w/d)^{-1},$$

$$B = (\Sigma_0^{-1} + x'x/d)^{-1}, \quad \bar{\alpha}_1 = 0, \quad \text{and} \quad \Sigma_0 = 100I.$$

19. Draw $\gamma_2 = (\gamma_{21}, \gamma_{22}, \gamma_{23}, \gamma_{24}, \gamma_{25})'$ similar to Step 18.
20. Draw ω_i.

$$w_{it} = \ln(CPC_{it})$$

$$- (\lambda_1 Rank_{i,t-1,1} + \lambda_2 Rank_{i,t-1,2} + \lambda_3 Profit_{i,t-1,1}$$

$$+ \lambda_4 Profit_{i,t-1,2} + \lambda_5 Retailer_i + \lambda_6 Brand_i + \lambda_7 Length_i$$

$$+ \lambda_8 Competitor\_Price_i + \lambda_9 Time_{it}) - e_{it}$$

$$= \omega_i x_{it} + \varsigma_{it},$$

where $x_{it} = 1$. Let us denote d and e_{ij} as the variance and mean of ς_{it}, conditional on the values of $(\eta_{it1}, \eta_{it2}, \mu_{it1}, \mu_{it2}, \nu_{it})'$ and Ω.

$$\omega_i \sim \text{MVN}(A_i, B_i), \quad A_i = B_i(\bar{\omega}/\sigma_\omega^2 + x_i'w_i/d)^{-1}, \quad \text{and}$$

$$B_i = (1/\sigma_\omega^2 + x_i'x_i/d)^{-1}.$$

21. Draw $\bar{\omega}$.

$$\bar{\omega} \sim N(A, B), \quad A = \sum_i \omega_i/N, \quad \text{and} \quad B = \sigma_\omega^2/N.$$

22. Draw σ_ω^2.

$$\sigma_\omega^2 \sim \text{Inverted Gamma}(A, B),$$

$$A = s_0 + N/2 \quad (s_0 = 5),$$

$$B = \frac{2}{\sum_{i=1}^{N}(\omega_i - \bar{\omega})^2 + 2/q_0} \quad (q_0 = 1).$$

23. Draw $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9)'$.

$$w_{it} = \ln(CPC_{it}) - \omega_i - e_{it} = \lambda x_{it} + \varsigma_{it},$$

$$x_{it} = (Rank_{i,t-1,1}, Rank_{i,t-1,2}, Profit_{i,t-1,1}, Profit_{i,t-1,2},$$

$$Retailer_i, Brand_i, Length_i, Competitor\_Price_i, Time_{it}).$$

Let us denote d and e_{it} as the variance and mean of ς_{it}, conditional on the values of $(\eta_{it1}, \eta_{it2}, \mu_{it1}, \mu_{it2}, \nu_{it})'$ and Ω.

$$\lambda \sim \text{MVN}(A, B), \quad A = B(\bar{\lambda}_0/\sigma_0^2 + x'w/d)^{-1},$$

$$B = (1/\sigma_0^2 + x'x/d)^{-1}, \quad \bar{\lambda}_0 = 0, \quad \text{and} \quad \sigma_0^2 = 100.$$

24. Draw ϕ_i similar to Step 20.
25. Draw $\bar{\phi}$ similar to Step 21.
26. Draw σ_ϕ^2 similar to Step 22.
27. Draw $\tau = (\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6)'$ similar to Step 23.
28. Draw $\underline{\psi}_i$ similar to Step 20.
29. Draw $\underline{\psi}$ similar to Step 21.

30. Draw σ_ψ^2 similar to Step 22.
31. Draw $\delta = (\delta_1, \delta_2, \delta_3, \delta_4)'$ similar to Step 23.
32. Draw Ω.
Let $f_{it} = (\kappa_{it}, \eta_{it1}, \eta_{it2}, \mu_{it1}, \mu_{it2}, \varsigma_{it}, \nu_{it})'$,

$$\Omega \sim \text{IW}\left(\sum_i \sum_t f_{it} f_{it}' + Q_0, N + q_0\right),$$

$$Q_0 = 10I \quad \text{and} \quad q_0 = 10, \quad K = \text{no. of observations.}$$

References

Agarwal, A., K. Hosanagar, M. D. Smith. 2008. Location, location, location: An analysis of profitability of position in online advertising markets. Working paper, http://ssrn.com/abstract=1151537.

Animesh, A., V. Ramachandran, S. Viswanathan. 2010. Quality uncertainty and the performance of online sponsored search markets: An empirical investigation. *Inform. Systems Res.* **21**(1) 190–201.

Ansari, A., C. F. Mela. 2003. E-customization. *J. Marketing Res.* **40**(2) 131–145.

Archak, N., A. Ghose, P. Ipeirotis. 2007. Deriving the pricing power of product features by mining consumer reviews. NET Institute Working Paper 07-36, http://ssrn.com/abstract=1024903.

Athey, S., G. Ellison. 2008. Position auctions with consumer search. Working paper, Harvard University, Cambridge, MA.

Bajari, P., H. Hong, J. Krainer, D. Nekipolov. 2006. Estimating static models of strategic interactions. Working paper, Duke University, Durham, NC.

Baye, M. R., J. Morgan. 2001. Information gatekeepers on the Internet and the competitiveness of homogeneous product markets. *Amer. Econom. Rev.* **91**(3) 454–474.

Baye, M. R., J. R. J. Gatti, P. Kattuman, J. Morgan. 2009. Clicks, discontinuities, and firm demand online. *J. Econom. Management Strategy* **18**(4) 935–975.

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36**(2) 192–236.

Bradlow, E. T., D. C. Schmittlein. 2000. The little engines that could: Modeling the performance of World Wide Web search engines. *Marketing Sci.* **19**(1) 43–62.

Bresnahan, T., P. Reiss. 1991. Entry and competition in concentrated markets. *J. Political Econom.* **99**(5) 977–1009.

Broder, A. 2002. A taxonomy of Web search. *ACM SIGIR Forum* **36**(2) 3–10.

Brooks, N. 2005. The Atlas Rank Report II: How search engine rank impacts conversions. Report, Atlas Institute, Seattle, http://www.atlassolutions.com/pdf/RankReportPart2.pdf.

Brynjolfsson, E., A. Dick, M. D. Smith. 2004. Search and product differentiation at an Internet shopbot. Working Paper 4441-03, Sloan School of Management, Massachusetts Institute of Technology, Cambridge.

Chatterjee, P., D. L. Hoffman, T. P. Novak. 2003. Modeling the clickstream: Implications for Web-based advertising efforts. *Marketing Sci.* **22**(4) 520–541.

Chen, Y., S. Yang. 2007. Estimating disaggregate models using aggregate data through augmentation of individual choice. *J. Marketing Res.* **44**(4) 613–621.

Chib, S., E. Greenberg. 1995. Understanding the Metropolis-Hastings algorithm. *Amer. Statistician* **49**(4) 327–335.

Dhar, V., A. Ghose. 2009. Market efficiency in sponsored search. Working paper, New York University, New York.

Drèze, X., F.-X. Hussherr. 2003. Internet advertising: Is anybody watching? *J. Interactive Marketing* **17**(4) 8–23.

Edelman, B., M. Ostrovsky, M. Schwarz. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* **97**(1) 242–259.

eMarketer. 2006. Paid search ad spend will hit $10 billion by 2009. (March 7), http://www.marketwire.com/press.release/Emarketer.688780.html.

Feng, J., H. K. Bhargava, D. M. Pennock. 2007. Implementing sponsored search in Web search engines: Computational evaluation of alternative mechanisms. *INFORMS J. Comput.* **19**(1) 137–148.

Gallagher, K., D. Foster, J. Parsons. 2001. The medium is not the message: Advertising effectiveness and content evaluation in print and on the Web. *J. Advertising Res.* **41**(4) 57–70.

Geerardyn, A., G. Fauconnier. 2000. Hybrid forms in marketing communication: The increasingly fuzzy boundaries between information and commerce. B. Pattyn, ed. *Media Ethics: Opening Social Dialogue.* Peeters Publishers, Leuven, Belgium, 329–355.

Gerstmeier, E., T. Stepanchuk, B. Skiera. 2009. An analysis of the profitability of different bidding heuristics in search engine marketing. Working paper, Goethe University, Frankfurt, Germany.

Ghose, A., S. Yang. 2008a. Comparing performance metrics in organic search with sponsored search advertising. *Proc. 2nd Annual Internat. Workshop Datamining Audience Intelligence Advertising (ADKDD 2008), Las Vegas, NV.*

Ghose, A., S. Yang. 2008b. Modeling cross-category purchases in sponsored search advertising. Working paper, http://ssrn.com/abstract=1312864.

Ghose, A., S. Yang. 2009. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Sci.* **55**(10) 1605–1622.

Goldfarb, A., C. Tucker. 2007. Search engine advertising: Pricing ads to context. NET Institute Working Paper 07-23, http://ssrn.com/abstract=1021451.

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

Greenspan, R. 2004. Searching for balance. ClickZ (April 30), http://www.clickz.com/3348071.

Hartzer, B. 2005. Balancing paid and organic search listings. *Search Engine Guide* (October 25), http://www.searchengineguide.com/bill-hartzer/balancing-paid.php.

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

Hawkins, S. A., S. J. Hoch. 1992. Low-involvement learning: Memory without evaluation. *J. Consumer Res.* **19**(2) 212–225.

Hotchkiss, G., M. Garrison, S. Jensen. 2005. Search engine usage in North America. White paper, Enquiro, Kelowna, BC, Canada.

Ilfeld, J., R. Winer. 2002. Generating website traffic. *J. Advertising Res.* **42**(5) 49–61.

Janiszewski, C. 1993. Preattentive mere exposure effects. *J. Consumer Res.* **20**(3) 376–392.

Jansen, B. J. 2007. The comparative effectiveness of sponsored and non-sponsored links for Web e-commerce queries. *ACM Trans. Web* **1**(1) Article 3.

Jansen, B. J., A. Brown, M. Resnick. 2007. Factors relating to the decision to click on a sponsored link. *Decision Support Systems* **44**(1) 46–59.

Jerath, K., L. Ma, Y.-H. Park, K. Srinivasan. 2009. A "position paradox" in sponsored search auctions. Johnson School Research Paper 36-09, Cornell University, Ithaca, NY.

Katona, Z., M. Sarvary. 2010. The race for sponsored links: Bidding patterns for search advertising. *Marketing Sci.* **29**(2) 199–215.

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

Law, S., K. A. Braun. 2000. I'll have what she's having: Gauging the impact of product placements on viewers. *Psych. Marketing* **17**(12) 1059–1075.

Liu, D., J. Chen, A. Whinston. 2010. Ex ante information and the design of keyword auctions. *Inform. Systems Res.* **21**(1) 133–153.

Maddala, G. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, Cambridge, UK.

Manchanda, P., J.-P. Dubé, K. Y. Goh, P. Chintagunta. 2006. The effect of banner advertising on Internet purchasing. *J. Marketing Res.* **43**(1) 98–108.

MarketingExperiments.com. 2005. PPC ad copy tested. (June 22), http://www.marketingexperiments.com/ppc-seo-optimization/ppc-ad-copy-tested.html.

Mela, C. F., S. Gupta, K. Jedidi. 1998. Assessing long-term promotional influences on market structure. *Internat. J. Res. Marketing* **15**(2) 89–107.

Moe, W. W., P. S. Fader. 2003. Dynamic conversion behavior at e-commerce sites. *Management Sci.* **50**(3) 326–335.

Montgomery, A. L., K. Hosanagar, R. Krishnan, K. B. Clay. 2004. Designing a better shopbot. *Management Sci.* **50**(2) 189–206.

Moon, S., G. J. Russell. 2008. Predicting product purchase from inferred customer similarity: An autologistic model approach. *Management Sci.* **54**(1) 71–82.

Musalem, A., E. T. Bradlow, J. S. Raju. 2009. Bayesian estimation of random-coefficients choice models using aggregate data. *J. Appl. Econometrics* **24**(3) 490–516.

Nelson, F. D., L. Olsen. 1978. Specification and estimation of a simultaneous-equation model with limited dependent variables. *Internat. Econom. Rev.* **19**(October) 695–709.

Rossi, P. E., G. M. Allenby. 2003. Bayesian statistics and marketing. *Marketing Sci.* **22**(3) 304–328.

Rusmevichientong, P., D. P. Williamson. 2006. An adaptive algorithm for selecting profitable keywords for search-based advertising services. *Proc. 7th ACM Conf. Electronic Commerce*, ACM, New York, 260–269.

Rutz, O. J., R. E. Bucklin. 2007. A model of individual keyword performance in paid search advertising. Working paper, http://ssrn.com/abstract=1024765.

Rutz, O. J., R. E. Bucklin. 2008. From generic to branded: A model of spillover dynamics in paid search advertising. Working paper, http://ssrn.com/abstract=1024766.

Search Engine Roundtable. 2007. Does Google pay per click influence organic search results? (May 31), http://www.seroundtable.com/archives/013662.html.

Shapiro, S., D. J. MacInnis, S. E. Heckler. 1997. The effects of incidental ad exposure on the formation of consideration sets. *J. Consumer Res.* **24**(1) 94–104.

Taylor, G. 2009. Competing search engines with utility maximising consumers and ordinary search results. Working paper, University of Southampton, Southampton, UK.

Telang, R., P. Boatwright, T. Mukhopadhyay. 2004. A mixture model for Internet search-engine visits. *J. Marketing Res.* **41**(2) 206–214.

Tellis, G. 2004. *Effective Advertising.* Sage, Thousand Oaks, CA.

Varian, H. 2007. Position auctions. *Internat. J. Indust. Organ.* **25**(6) 1163–1178.

White, A. 2008. Search engines: Left side quality versus right side profits. Working paper, Toulouse School of Economics, Toulouse, France.

Wilbur, K., Y. Zhu. 2008. Click fraud. *Marketing Sci.* **28**(2) 293–308.

Xu, L., J. Chen, A. B. Whinston. 2009. Too organic for organic listing? Interplay between organic and sponsored listing in search advertising. McCombs Research Paper IROM-13-09, McCombs School of Business, University of Texas at Austin, Austin.

Yang, S., G. M. Allenby. 2000. A model for observation, structural and household heterogeneity in panel data. *Marketing Lett.* **11**(2) 137–149.

Yao, S., C. F. Mela. 2009. A dynamic model of sponsored search advertising. Working paper, Duke University, Durham, NC.