MANAGEMENT SCIENCE
Vol. 55, No. 10, October 2009, pp. 1605–1622
ISSN 0025-1909 | EISSN 1526-5501 | 09 | 5510 | 1605
DOI 10.1287/mnsc.1090.1054

An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets

Anindya Ghose
Leonard N. Stern School of Business, New York University, New York, New York 10012,
aghose@stern.nyu.edu

Sha Yang
Leonard N. Stern School of Business, New York University, New York, New York 10012; and
School of Economics and Management, Southwest Jiaotong University, 610031 Chenfgdu, China,
shayang@stern.nyu.edu

The phenomenon of sponsored search advertising—where advertisers pay a fee to Internet search engines to be displayed alongside organic (nonsponsored) Web search results—is gaining ground as the largest source of revenues for search engines. Using a unique six-month panel data set of several hundred keywords collected from a large nationwide retailer that advertises on Google, we empirically model the relationship between different sponsored search metrics such as click-through rates, conversion rates, cost per click, and ranking of advertisements. Our paper proposes a novel framework to better understand the factors that drive differences in these metrics. We use a hierarchical Bayesian modeling framework and estimate the model using Markov Chain Monte Carlo methods. Using a simultaneous equations model, we quantify the relationship between various keyword characteristics, position of the advertisement, and the landing page quality score on consumer search and purchase behavior as well as on advertiser's cost per click and the search engine's ranking decision. Specifically, we find that the monetary value of a click is not uniform across all positions because conversion rates are highest at the top and decrease with rank as one goes down the search engine results page. Though search engines take into account the current period's bid as well as prior click-through rates before deciding the final rank of an advertisement in the current period, the current bid has a larger effect than prior click-through rates. We also find that an increase in landing page quality scores is associated with an increase in conversion rates and a decrease in advertiser's cost per click. Furthermore, our analysis shows 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. Besides providing managerial insights into search engine advertising, these results shed light on some key assumptions made in the theoretical modeling literature in sponsored search.

Key words: online advertising; search engines; hierarchical Bayesian modeling; paid search; click-through rates; conversion rates; keyword ranking; cost per click; electronic commerce; Internet monetization
History: Received August 30, 2007; accepted April 14, 2009, by Jagmohan S. Raju marketing. Published online in *Articles in Advance* July 27, 2009.

1. Introduction

The Internet has brought about a fundamental change in the way users generate and obtain information, thereby facilitating a paradigm shift in consumer search and purchase patterns. In this regard, search engines are able to leverage their value as information location tools by selling advertising linked to user-generated search queries. Indeed, the phenomenon of sponsored search advertising—where advertisers pay a fee to Internet search engines to be displayed alongside organic (nonsponsored) Web search results—is gaining ground as the largest source of revenues for search engines. The global paid search advertising market is predicted to have a 37% compound annual growth rate, to more than $33 billion in 2010, and has become a critical component of firms' marketing campaigns.

Search engines such as Google, Yahoo, and MSN have discovered that as intermediaries between users and firms, they are in a unique position to sell new forms of advertisements without annoying consumers. In particular, sponsored search advertising has gradually evolved to satisfy consumers' penchant for relevant search results and advertisers' desire for inviting high-quality traffic to their websites. These advertisements are based on customers' own queries and are hence considered far less intrusive than online banner ads or pop-up ads. The specific "keywords" in response to which the ads are displayed are often chosen based on user-generated content in online product reviews, social networks, and blogs where users have posted their opinions about firms' products, often highlighting the specific product features they value the most (Dhar and Ghose 2009). In many ways, the increased ability of users to interact with firms in the online world has enabled a shift from "mass" advertising to more "targeted" advertising.

How does this mechanism work? In sponsored search, firms that wish to advertise their products or services on the Internet submit their product information in the form of specific keyword listings to search engines. Bid values are assigned to each individual ad to determine the position of each competing listing on the search engine results page when a user performs a search. When a consumer searches for a term on the search engine, the advertisers' webpage appears as a sponsored link next to the organic search results that would otherwise be returned using the neutral criteria employed by the search engine. By allotting a specific value to each keyword, advertisers only pay the assigned price for the users who actually click on their listing to visit their websites in the most prevalent payment mechanism, known as cost per click (CPC). Because listings appear only when a user generates a keyword query, an advertiser can reach a more targeted audience on a relatively lower budget through search engine advertising.

Despite the growth of search advertising, we have little understanding of how consumers respond to contextual and sponsored search advertising on the Internet. In this paper, we focus on previously unexplored issues: How does sponsored search advertising affect consumer search and purchasing behavior on the Internet? More specifically, what kinds of sponsored keyword advertisement most contribute to variation in advertiser value in terms of consumer click-through rates and conversions? What is the relationship between different kinds of keywords and the advertiser's actual CPC and the search engine's keyword ranking decision? An emerging stream of theoretical literature in sponsored search has looked at issues such as mechanism design in auctions, but no prior work has empirically analyzed these kinds of questions. Given the shift in advertising from traditional banner advertising to search engine advertising, an understanding of the determinants of conversion rates and click-through rates in search advertising is essential for both traditional and Internet retailers.

Using a unique panel data set of several hundred keywords collected from a nationwide retailer that advertises on Google, we examine the relationship between various keyword characteristics, position of the keyword advertisement on the search engine results page, and the landing page quality score on consumer and firm behavior. In particular, we propose a hierarchical Bayesian modeling framework in which we build a simultaneous model to jointly estimate the impact of various keyword attributes on consumer click-through and purchase propensities, on the advertiser's CPC, and on the search engine ad ranking decision.

Our empirical analyses provide several descriptive insights. The presence of retailer-specific information in the keyword is associated with an increase in click-through and conversion rates, by 14.72% and 50.6%, respectively; the presence of brand-specific information in the keyword is associated with a decrease in click-through and conversion rates, by 56.6% and 44.2%, respectively; and the length of the keyword is associated with a decrease in click-through rates by 13.9%. Keyword rank is negatively associated with the click-through rates and conversion rates such that both these metrics decrease with ad position as one goes down the search engine results page. Furthermore, this relationship is increasing at a decreasing rate for both metrics. An increase in the landing page quality score of the advertiser by 1 unit is associated with an increase in conversion rates by as much as 22.5%. CPC is negatively associated with the landing page quality. Finally, our data suggest that profits are not necessarily monotonic with rank such that keywords that have more prominent positions on the search engine results page and thus experience higher click-through rates as well as higher conversion rates are not necessarily the most profitable ones. In fact, we find that profits are often higher for keywords that are ranked in the middle positions than for those in the very top on the search engine's results page.

Our key contributions are summarized as follows. First, our paper is the first empirical study that simultaneously models and documents the impact of search engine advertising on all three entities involved in the process—consumers, advertisers, and search engines. The proposed simultaneous model provides a natural way to account for endogenous relationships between decision variables, leading to a robust identification strategy and precise estimates. The model can be applied to similar data from other industries. Moreover, unlike previous work, we jointly study consumer click-through behavior and conversion behavior conditional on a click-through in studying consumer search behavior. Ignoring consumer click-through behavior can lead to selectivity bias if the error terms in the click-through probability and in the conditional conversion probability are correlated (Maddala 1983), and this is an additional contribution. The proposed Bayesian estimation algorithm provides a convenient way to estimate such a model by using data augmentation. The empirical estimates provide descriptive insights about what kinds of keyword advertisements contribute to variation in consumer behavior and advertiser value. In particular, our study examines the relationship between branded/retailer/generic and shorter/longer keywords and demand-side variables like click-through rates and conversion rates—a question of increasing interest to many firms.

Second, our paper provides insights into assumptions made in the theoretical modeling literature on search engine advertising. By showing a direct negative relationship between conversion rates and rank, we show that the value per click to an advertiser is not uniform across slots. This finding refutes a commonly held assumption in prior work that the value of a click from a sponsored search campaign is independent of the position of the advertisement. Prior theoretical work (e.g., Aggarwal et al. 2006, Edelman et al. 2007, Varian 2007) also makes a common assumption of uniform value per click across all ranks and shows that under this condition, sponsored search auctions maximize social welfare. Our finding of nonuniformity in value per click paves the way for future theoretical models in this domain that could relax this assumption and design newer mechanisms with more robust equilibrium or welfare-maximizing properties. The recent work by Börgers et al. (2007) and Xu et al. (2009) that allows value per click to vary across positions in their theoretical models is a step in this direction.

Finally, we find that (i) whereas search engines take into account the current period's bid as well as prior click-through rates before deciding the final rank of an advertisement in the current period, the current bid has a larger effect than prior click-through rates; (ii) an increase in landing page quality scores is associated with an increase in conversion rates and a decrease in advertiser's CPC; and (iii) even though the more prominent positions on the search engine results page experience higher click-through or conversion rates, they may not be the most profitable ones—profits are often higher at the middle positions than at the top or the bottom positions. Our findings thus corroborate claims about institutional practice in this industry and shed new light on conventional wisdom about profitability associated with ad position.

The remainder of this paper is organized as follows. Section 2 gives an overview of the different streams of literature from marketing and computer science related to the topic of our paper. Section 3 describes the data and gives a brief background into some different aspects of sponsored search advertising that could be useful before we proceed to the empirical models and analyses. In §4, we present a model to study the click-through rate, conversion rate, and keyword ranking simultaneously and discuss our identification strategy. In §5, we discuss our empirical findings. In §6, we discuss some implications of our findings and conclude.

2. Literature and Theoretical Background

Our paper is related to several streams of research. A number of approaches have been build to model the effects of advertising based on aggregate data (Tellis 2004). Much of the existing literature (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. This is usually done via field surveys or laboratory experiments using individual (or cookie) level data. Sherman and Deighton (2001) and Ilfeld and Winer (2002) show that using aggregate data that increased online advertising leads to more site visits. In contrast to other studies that measure (individual) exposure to advertising via aggregate advertising dollars (e.g., Mela et al. 1998, Ilfeld and Winer 2002), we use data on individual search keyword advertising exposure. Manchanda et al. (2006) look at online banner advertising. Because banner ads have been perceived by many consumers as being annoying, traditionally they have had a negative connotation associated with them. Moreover, it was argued that because there is considerable evidence that only a small proportion of visits translate into a final purchase (Sherman and Deighton 2001, Moe and Fader 2003, Chatterjee et al. 2003), click-through rates may be too imprecise to measure 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 studies of paid placements available to retailers on the Internet in the form of sponsored listings on shopping bots (Baye and Morgan 2001, Baye et al. 2009).

There is also an emerging theoretical stream of literature exemplified by Aggarwal et al. (2006), Edelman et al. (2007), Feng et al. (2007), Varian (2007), and Liu et al. (2009) that analyzes mechanism design and equilibria in search engine auctions. Chen and He (2006) and Athey and Ellison (2008) build models that integrate consumer behavior with advertiser decisions, and the latter paper theoretically analyzes several possible scenarios in the design of sponsored keyword auctions. Katona and Sarvary (2007) build a model of competition in sponsored search and find that the interaction between search listings and paid links determines equilibrium bidding behavior.

Gerstmeier et al. (2009) discuss some interesting bidding heuristics and highlight which of these leads to higher profits for the advertiser.

Despite the emerging theory work, little empirical work exists in online search advertising. This is primarily because of the difficulty researchers have in obtaining such advertiser-level data. Existing work has so far focused on search engine performance (Telang et al. 2004, Bradlow and Schmittlein 2000). Moreover, the handful of studies that exist in search engine marketing has typically analyzed publicly available data from search engines. Animesh et al. (2009) look at the presence of quality uncertainty and adverse selection in paid search advertising on search engines. Goldfarb and Tucker (2007) examine the factors that drive variation in prices for advertising legal services on Google.

A more closely related stream of work is the one that uses advertiser-level data in sponsored search. Ghose and Yang (2008) 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. Rutz and Bucklin (2008) showed that there are spillovers between search advertising on branded and search advertising on generic keywords, as some customers may start with a generic search to gather information but later use a branded search to complete their transaction. Yang and Ghose (2008) examine the interdependence between paid search and organic search listings and find a positive interdependence between the two forms of listings with regard to their impact on click-through rates. Agarwal et al. (2008) provide quantitative insights into the profitability of advertisements associated with differences in keyword position and show that profits may not be monotonic with rank. In an interesting paper related to our work, Rutz and Bucklin (2007) studied hotel marketing keywords to analyze the profitability of different campaign management strategies. However, our paper differs from theirs and extends their work in several important ways. Rutz and Bucklin (2007) only model the conversion probability conditional on a positive number of click throughs. However, our paper models click-through and conversion rates simultaneously to alleviate potential selectivity biases. In addition, we also model the search engine's ranking decision and the advertiser's decision on CPC, both of which are absent in the other paper. Our analysis reveals that it is important to model the advertiser and the search engine's decisions simultaneously with clicks and conversion since both CPC and Rank have been found to be endogenous.

To summarize, our research is distinct from extant online advertising research because it has largely been limited to the influence of banner advertisements on attitudes and behavior. We extend the literature by empirically comparing the relationship of different keyword characteristics with various performance metrics in search engine advertising toward understanding the larger question of analyzing how keyword characteristics are associated with variation in consumers' search and purchase behavior, as well as advertisers' CPC and search engines' ranking decisions.

3. Data

The data-generation process for paid keyword advertisement differs on many dimensions from traditional offline advertisement. Advertisers bid on keywords during the auction process. (A keyword may consist of one or more "words.") Once the advertiser gets a rank allotted for its keyword ad, these sponsored ads are displayed on the top left and right of the computer screen in response to a query that a consumer types on the search engine. The match between a user query and the advertisement could be based on a broad, exact, or phrase match. The ad typically consists of headline, a word or a limited number of words describing the product or service, and a hyperlink that refers the consumer to the advertiser's website after a click. The serving of the 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 usually pay on a per click basis. 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, which advertises on Google.[1] The data span all keyword advertisements by the company during a period of six months in 2007, specifically for the 24 calendar weeks from January to June. Each keyword in our data has a unique advertisement ID. The data are for a given keyword for a given week and are based on an "exact match" between the user query and sponsored ad. It consists of the number of impressions, number of clicks, average CPC, rank of the keyword, number of conversions, and the total revenues from a conversion. An impression often leads to a click, but it may not lead to an actual purchase (defined as a conversion). Based on these data, we compute the *Click-Through Rate* (clicks/impressions) and *Conversion Rate* (conversions/clicks) variables.

[1] The firm is a large Fortune 500 retail store chain with several hundred retail stores in the United States. Because of the nature of the data-sharing agreement between the firm and us, we are unable to reveal the name of the firm.

The product of *CPC* and number of clicks gives the total costs to the firm for sponsoring a particular advertisement. Based on the contribution margin and the revenues from each conversion through a paid search advertisement, we are able to compute the gross profit per keyword from a 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 is the *Profit* variable. We use this variable primarily in our robustness tests (see the online appendix, provided in the e-companion).[2]

Finally, although we have data on the URLs of the landing page corresponding to a given keyword, we do not have data on landing page quality scores or content, because the exact algorithm used by Google to impute the landing page quality is not disclosed to the public.[3] Hence, we use a semiautomated approach with content analysis to impute the landing page quality based on the three known metrics that Google uses. Google uses a weighted average of relevancy, transparency, and navigability to impute the landing page quality of a given weblink. We hired two independent annotators to rate each landing page based on each of these metrics and then computed the weighted average of the scores. The interrater reliability score was 0.73, indicating a very high level of reliability.

Our final data set includes 9,664 observations from a total of 1,878 unique keywords. Note that our main interest in this empirical investigation is to examine various keyword-level factors that induce differences in click-throughs and conversions. Hence, we analyze click-through rates, conversion rates, CPC, and rank by jointly modeling the consumers' search and purchase behavior, the advertiser's decision on CPC, and the search engine's keyword rank-allocating behavior. Table 1 reports the summary statistics of our data set. As shown, the average weekly number of impressions is 411 for one keyword, among which around 46 lead to a click-through and 0.85 lead to a purchase. Our data suggest the average CPC for a given keyword is about 25 cents, and the average rank (position) of these keywords is about 6.92. Finally, we have information on three important keyword characteristics.

[2] An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

[3] Google computes a quality score for each landing page as a function of the site's navigability as well as the relevance and transparency of information on that page to provide higher user experience after a click-through to the site. Besides these relevancy factors, the quality score is also based on click-through rates. However, the exact algorithm for computing this score is not publicly available. The quality score is then used in determining the minimum bid price, which in turn affects the rank of the ad, given the typical advertiser budget constraints. Further information on these aspects is available at http://www.adwords.google.com.

Table 1 Summary Statistics of the Paid Search Data ($N = 9{,}664$)

Variable	Mean	Std. dev.	Min	Max
Impressions	411.694	2,441.488	1	97,424
Clicks	46.266	716.812	0	38,465
Orders	0.860	11.891	0	644
Click-Through Rate (*CTR*)	0.156	0.262	0	1
Conversion Rate	0.023	0.132	0	1
Cost per Click (*CPC*)	0.245	0.181	0.001	1.46
Lag_Rank	6.473	9.139	1	131
log(*Profit*)	0.036	1.771	−5.210	11.282
log(*Lag_Profit*)	0.026	1.726	−5.210	11.282
Rank	6.926	10.027	1	131
Lag_CTR	0.154	0.250	0	1
Retailer	0.076	0.265	0	1
Brand	0.427	0.494	0	1
Length	2.632	0.755	1	6
LandingPageQuality	8.556	1.434	4	10
Competitor Price	1.514	1.811	0.18	45.42

As Table 1 shows, there is a substantial amount of variation in clicks, conversion, rank, and CPC of each keyword over time.

We enhanced the data set by introducing keyword-specific characteristics such as *Brand*, *Retailer*, and *Length*. For each keyword, we constructed two dummy variables, based on whether they were (i) branded keywords or not (for example, "Sealy mattress," "Nautica bedsheets") and (ii) retailer-specific advertisements (for example, "Walmart," "walmart.com") or not. To be precise, for creating the variable in (i), we looked for the presence of a brand name (either a product-specific or a company-specific use) in the keyword and labeled the dummy as 1 or 0, with 1 indicating the presence of a brand name. For (ii), we looked for the presence of the specific advertiser's (retailer) name in the keyword and then labeled the dummy as 1 or 0, with 1 indicating the presence of the retailer's name. *Length* is defined as the number of words contained in the keyword.

4. A Simultaneous Model of Click-Through, Conversion, CPC, and Rank

We cast our model in a hierarchical Bayesian framework and estimate it using Markov Chain Monte Carlo (MCMC) methods (see Rossi and Allenby 2003 for a detailed review of such models). We postulate that the decision of whether to click and purchase in a given week will be affected by the probability of advertising exposure (for example, through the rank of the keyword) and individual keyword-level differences (both observed and unobserved). We simultaneously model consumers' click-through and conversion behavior, the advertiser's keyword pricing behavior, and the search engine's keyword rank-allocating behavior.

4.1. Theoretical Setup

Assume for search keyword i at week j that there are n_{ij} click-throughs among N_{ij} impressions (the number of times an advertisement is displayed by the retailer), where $n_{ij} \leq N_{ij}$ and $N_{ij} > 0$. Suppose that among the n_{ij} click-throughs, there are m_{ij} that lead to purchases, where $m_{ij} \leq n_{ij}$. Let us further assume that the probability of having a click-through is p_{ij} and the probability of having a purchase conditional on a click-through is q_{ij}. In our model, a consumer faces decisions at two levels—one, when she sees a keyword advertisement, she makes a decision whether to click it; two, if she clicks on the advertisement, she can either make a purchase or not make a purchase.

Thus, there are three types of observations. First, a person clicked through and made a purchase. The probability of such an event is $p_{ij}q_{ij}$. Second, a person clicked through but did not make a purchase. The probability of such an event is $p_{ij}(1-q_{ij})$. Third, an impression did not lead to a click-through or purchase. The probability of such an event is $1-p_{ij}$. Then the probability of observing (n_{ij}, m_{ij}) is given by

$$f(n_{ij}, m_{ij}, p_{ij}, q_{ij}) = \frac{N_{ij}!}{m_{ij}!(n_{ij}-m_{ij})!(N_{ij}-n_{ij})!}(p_{ij}q_{ij})^{m_{ij}} \cdot [p_{ij}(1-q_{ij})]^{n_{ij}-m_{ij}}(1-p_{ij})^{N_{ij}-n_{ij}}. \quad (1)$$

4.2. Modeling the Consumer's Decision: Click-Through

Prior work (Broder 2002, Jansen and Spink 2007) has analyzed the goals for users' Web searches and classified user queries in search engines into three categories of searches: *navigational* (e.g., 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). Being cognizant of such user behavior, search engines sell not only nonbranded or generic keywords as advertisements, but also well-known product or manufacturer brand names as well as keywords indicating the specific advertiser so the firm can attract consumers to its website.[4] Moreover, advertisers also have the option of making the keyword advertisement either generic or specific by altering the number of words contained in the keyword. Finally, the length of the keyword is also an important determinant of search and purchase behavior, but anecdotal evidence on this varies across trade press reports. Some studies have shown that the percentage of searchers who use a combination of keywords is 1.6 times the percentage of those who use single keyword queries (Kilpatrick 2003). In contrast, another study found that single keywords have on average the highest number of unique visitors (Oneupweb 2005). In our data, the average length of a keyword is about 2.6 words. In sum, the number of advertisers placing a bid, which can affect the number of clicks received by a given ad, will vary based on the kind of keyword that is advertised. Hence, we focus on the three important keyword-specific characteristics for the firm when it advertises on a search engine: *Brand*, *Retailer*, and *Length*. The click-through probability is likely to be influenced by the position of the ad (*Rank*), how specific or broad the keyword is (*Length*), and whether it contains any retailer-specific (*Retailer*) or brand-specific information (*Brand*). Hence, in Equation (1), p_{ij} the click-through probability is modeled as

$$\begin{aligned} p_{ij} = {} & [\exp(\beta_{i0} + \beta_{i1} Rank_{ij} + \alpha_1 Retailer_i + \alpha_2 Brand_i \\ & + \alpha_3 Length_i + \alpha_4 Time_{ij} + \varepsilon_{ij})] \\ & \cdot [1 + \exp(\beta_{i0} + \beta_{i1} Rank_{ij} + \alpha_1 Retailer_i + \alpha_2 Brand_i \\ & + \alpha_3 Length_i + \alpha_4 Time_{ij} + \varepsilon_{ij})]^{-1}. \end{aligned} \quad (2)$$

We capture the unobserved heterogeneity with a random coefficient on the intercept by allowing β_{i0} to vary along its population mean $\bar{\beta}_0$ as follows:

$$\beta_{i0} = \bar{\beta}_0 + \varsigma_{i0}^{\beta}. \quad (3)$$

We also allow the *Rank* coefficient of the ith keyword to vary along the population mean $\bar{\beta}_1$ and the keywords' characteristics as follows:

$$\beta_{i1} = \bar{\beta}_1 + \gamma_1 Retailer_i + \gamma_2 Brand_i + \gamma_3 Length_i + \varsigma_{i1}^{\beta}, \quad (4)$$

$$\begin{bmatrix} \varsigma_{i0}^{\beta} \\ \varsigma_{i1}^{\beta} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{11}^{\beta} & \Sigma_{12}^{\beta} \\ \Sigma_{21}^{\beta} & \Sigma_{22}^{\beta} \end{bmatrix} \right). \quad (5)$$

4.3. Modeling the Consumer's Decision: Conversion

Next, we model the conversion rates. Prior work (Brooks 2004) has shown that there is an intrinsic trust value associated with the rank of a firm's listing on a search engine, which could lead to the conversion rate dropping significantly with an increase in the rank (i.e., with a lower position on the screen). Hence, we include rank as a covariate. 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 don't see it addressed adequately on the landing page, they are likely to abandon their search and purchase process.

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

Different keywords from a given advertiser lead to different kinds of landing pages. Hence, it is important to incorporate the landing page quality as a covariate in the model. Furthermore, different keywords are associated with different products. It is possible that product-specific characteristics influence consumer conversion rates, so 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. Thus, the conversion probability is likely to be influenced by the position of the ad on the screen, the three keyword specific characteristics, and the landing page quality score. These factors lead us to model the conversion probabilities as follows:

$$
\begin{aligned}
q_{ij} = \big[&\exp(\theta_{i0} + \theta_{i1} Rank_{ij} + \delta_1 Retailer_i + \delta_2 Brand_i \\
&+ \delta_3 Length_i + \delta_4 LandingPageQuality_i \\
&+ \delta_5 Time_{ij} + \eta_{ij})\big] \\
\cdot \big[&1 + \exp(\theta_{i0} + \theta_{i1} Rank_{ij} + \delta_1 Retailer_i + \delta_2 Brand_i \\
&+ \delta_3 Length_i + \delta_4 LandingPageQuality_i \\
&+ \delta_5 Time_{ij} + \eta_{ij})\big]^{-1}. \qquad (6)
\end{aligned}
$$

As before, we capture the unobserved heterogeneity with a random coefficient specified on both the intercept and the *Rank* coefficient, as follows:

$$\theta_{i0} = \bar{\theta}_0 + \varsigma_{i0}^{\theta}, \qquad (7)$$

$$
\begin{aligned}
\theta_{i1} = \bar{\theta}_1 + \kappa_1 Retailer_i + \kappa_2 Brand_i + \kappa_3 Length_i \\
+ \kappa_4 LandingPageQuality_i + \varsigma_{i1}^{\beta}, \qquad (8)
\end{aligned}
$$

$$\begin{bmatrix} \varsigma_{i0}^{\theta} \\ \varsigma_{i1}^{\theta} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{11}^{\theta} & \Sigma_{12}^{\theta} \\ \Sigma_{21}^{\theta} & \Sigma_{22}^{\theta} \end{bmatrix} \right). \qquad (9)$$

Thus, Equations (1)–(9) model the demand for a keyword, i.e., the consumer's decision.

4.4. Modeling the Advertiser's Decision: CPC

Next, we model the advertiser's (i.e., the firm's) strategic behavior. The advertiser has to decide how much to bid for each keyword *i* in week *j* and thus the CPC that it is willing to incur.[5] The advertiser decides on its CPC by tracking the performance of a keyword over time such that the current CPC is dependent on past performance of that keyword.[6] Specifically, the keyword's current CPC is a function of the rank of the same keyword in the previous period. In keeping with the institutional practices of Google, which decides the minimum bid price of any given keyword ad as a function of landing page quality associated with that keyword, we control for the landing page quality in the advertiser's CPC decision. Different keyword attributes determine the extent of competitiveness in the bidding process for that keyword, as can be seen in the number of advertisers that, place a bid. For example, a "retailer" keyword is likely to be far less competitive, because the specific advertiser is usually the only firm that will bid on such a keyword. In contrast, "branded" keywords are likely to be much more competitive because there are several advertisers (retailers that sell that brand) that will bid on that keyword. Similarly, smaller keywords typically tend to indicate more generic ads and are likely to be much more competitive, whereas longer keywords typically tend to indicate more specific ads and are likely to be less competitive. Hence, the advertiser's CPC for a given keyword also depends on the three keyword attributes. Thus, the CPC will be influenced by the rank of the ad in the previous time period, the three keyword-specific characteristics, and the landing page quality. This leads to the following equation for the CPC of an advertiser:

$$
\begin{aligned}
&\ln(CPC_{ij}) \\
&= \omega_{i0} + \omega_{i1} Rank_{i,j-1} + \lambda_1 Retailer_i + \lambda_2 Brand_i + \lambda_3 Length_i \\
&\quad + \lambda_4 LandingPageQuality_i + \lambda_5 Time_{ij} + \mu_{ij}, \qquad (10)
\end{aligned}
$$

$$\omega_{i0} = \bar{\omega}_0 + \varsigma_{i0}^{\omega}, \qquad (11)$$

$$
\begin{aligned}
\omega_{i1} = \bar{\omega}_1 + \rho_{11} Retailer_i + \rho_{12} Brand_i + \rho_{13} Length_i \\
+ \rho_{14} LandingPageQuality_i + \varsigma_{i1}^{\omega}, \qquad (12)
\end{aligned}
$$

The error terms in Equations (11) and (12) are distributed as follows:

$$\begin{bmatrix} \varsigma_{i0}^{\omega} \\ \varsigma_{i1}^{\omega} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{11}^{\omega} & \Sigma_{12}^{\omega} \\ \Sigma_{21}^{\omega} & \Sigma_{22}^{\omega} \end{bmatrix} \right). \qquad (13)$$

4.5. Modeling the Search Engine's Decision: Keyword Rank

Finally, we model the search engine's decision on assigning ranks for a sponsored keyword advertisement. During the auction, search engines such as Google, MSN, and Yahoo decide on the keyword rank

[5] Because we do not have data on actual bids, we use the actual CPC as a proxy for the bid price. According to the firm whose data we use, they are very strongly correlated, and hence it is a very reasonable proxy.

[6] This information about current bids being based on past performance (lagged Rank) was given to us by the advertiser. The qualitative nature of all our results is robust to the use of both one-period lagged Rank and one-period lagged Profit, from a given keyword ad, which is another heuristic used by some advertisers to decide how much to bid for a given keyword in a given period.

by taking into account both the current CPC and a "quality score" that is determined by the prior click-through rate (CTR) of that keyword (Varian 2007, Athey and Ellison 2008) among other factors. Because more recent click-through rate is given more weight by the search engine in computing this score, we use the one-period lagged value of *CTR*. The three keyword attributes are used to control for unobserved characteristics such as the extent of competition in the auction bidding process as before in the CPC decision. Hence, the rank is modeled as being dependent on these three keyword attributes. This leads to the following equation for the rank of a keyword in sponsored search:

$$\ln(Rank_{ij}) = \phi_{i0} + \phi_{i1} CPC_{i,j} + \bar{\phi}_2 CTR_{i,j-1} + \tau_1 Retailer_i + \tau_2 Brand_i + \tau_3 Length_i + \tau_4 Time_{ij} + v_{ij}, \quad (14)$$

$$\phi_{i0} = \bar{\phi}_0 + \varsigma^{\phi}_{i0}, \quad (15)$$

$$\phi_{i1} = \bar{\phi}_1 + \pi_1 Retailer_i + \pi_2 Brand_i + \pi_3 Length_i + \varsigma^{\pi}_{i1}. \quad (16)$$

The error terms in Equations (15) and (16) are distributed as follows:

$$\begin{bmatrix} \varsigma^{\phi}_{i0} \\ \varsigma^{\phi}_{i1} \end{bmatrix} \sim \text{MVN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma^{\phi}_{11} & \Sigma^{\phi}_{12} \\ \Sigma^{\phi}_{21} & \Sigma^{\phi}_{22} \end{bmatrix} \right). \quad (17)$$

Finally, to model the unobserved covariation among click-through, conversions, CPC, and the keyword ranking, we let the four error terms be correlated in the following manner:

$$\begin{bmatrix} \varepsilon_{ij} \\ \eta_{ij} \\ \mu_{ij} \\ \nu_{ij} \end{bmatrix} \sim \text{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} & \Omega_{14} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} & \Omega_{24} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} & \Omega_{34} \\ \Omega_{41} & \Omega_{42} & \Omega_{43} & \Omega_{44} \end{bmatrix} \right). \quad (18)$$

A couple of clarifications are useful to note here. First, the three characteristics of a keyword (*Retailer*, *Brand*, and *Length*) are all mean centered. This means that $\bar{\beta}_1$ is the average effect of $\bar{\beta}_{i1}$ in Equation (4). A similar interpretation applies to the parameters θ_{i1}, ω_{i1}, ω_{i2}, and ϕ_{i1}. Second, in Equations (2), (6), (10), and (14), we have controlled for the temporal effects by estimating time-period effects that capture unobserved industry dynamics.

4.6. Identification

To ensure that the model is fully identified even with sparse data (data in which a large proportion of observations are zero), we conduct the following simulation. We picked a set of parameter values and generated the number of click-throughs, the number of purchases, CPC, and ranking for each keyword, which mimicked their actual observed values in the data according to the model and the actual independent variables observed in our data. We then estimated the proposed model with the simulated data set and found that we were able to recover the true parameter values. This relieves a potential concern on empirical identification of the model due to the sparseness of the data.

To show any endogeneity issues and the identification of the proposed system of simultaneous equation model, we provide a sketch of the model below. Note that our proposed model boils down to the following simultaneous equations:

$$p = f_1(Rank, X_1, \varepsilon_1), \quad (19)$$

$$q = f_2(Rank, X_2, \varepsilon_2) \quad \text{conditional on the number of click-throughs} > 0, \quad (20)$$

$$CPC = f_3(X_3, \varepsilon_3), \quad (21)$$

$$Rank = f_4(CPC, X_4, \varepsilon_4). \quad (22)$$

Here p is the click-through probability, q is the conversion as probability conditional on click-through, *CPC* is cost per click, and *Rank* is the position of a keyword in the listing. X_1–X_4 are the exogenous covariates corresponding to the four equations. ε_1–ε_4 are the error terms associated with the four equations, respectively. These error terms are mainly capturing information that is observed by the decision makers (consumer, advertiser, and search engine) but not by the researcher. Further, if ε_1 or ε_2 is correlated with ε_4, Rank will be endogenous. If ε_3 is correlated with ε_4, CPC will be endogenous.

Our proposed simultaneous model closely resembles the triangular system in standard econometric textbooks (Lahiri and Schmidt 1978, Greene 1999). 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 and other keyword-related characteristics). CPC, in turn, affects the search engine's ranking decision, and finally Rank affects both click-through and the conversion probabilities. As shown in Lahiri and Schmidt (1978) and discussed in Greene (1999), a triangular system of simultaneous equations can be identified without any further identification constraint 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 (SUR) problem (Hausman 1975). Hence, a generalized least squares-based estimation (GLS) leads to uniquely identified estimates in a triangular system with a full covariance on error terms (Lahiri and Schmidt 1978).[7]

We also provide the parameters produced by the estimation of this system under the assumption of diagonality (restricting covariance elements to be zero) to be able to compare them to the generalized results. These are given in the tables in the online appendix. These estimates show that it is important to control for endogeneity because the parameter estimates are attenuated when we restrict the covariance elements to be zero, and thus biased. For example, in the case of estimating *CTR* and *Conversion Rates*, the parameter estimates on *Rank* are much closer to zero under the assumption of diagonality than otherwise. Similarly, in the case of estimating *Rank*, the parameter estimates on *Lag_CTR* and *CPC* are significantly closer to zero under the assumption of diagonality than otherwise.

Note that the conversion probability q is only defined when the number of click-throughs is greater than zero. In this case, if ε_1 and ε_2 are correlated, as in our data, the conditional mean of ε_2 conditional on a positive click-through probability is not going to be zero. Then, a model in which one only looks at the conversion conditional on positive number of click-throughs (i.e., does not model the click-through behavior simultaneously) is going to suffer from the selection bias. By jointly modeling click-through and conversion behavior, our proposed model accounts for such selectivity issues. The proposed Bayesian estimation approach also offers a computationally convenient way to deal with the selectivity problem by augmenting the unobserved click-through intention when there are no clicks.

5. Empirical Analysis

5.1. Results

5.1.1. Click-Through Rate. The coefficient of *Retailer* in Table 2(a) is positive and significant, indicating that keyword advertisements that contain retailer-specific information are associated with a significant increase in click-through rates. Specifically, this corresponds to a 14.72% increase in click-through rates with the presence of retailer information. Further, the coefficient of *Brand* in Table 2(a)

Table 2(a) Coefficient Estimates on Click-Through Rate

	Intercept	*Retailer*	*Brand*	*Length*
Intercept	$\bar{\beta}_0$	α_1	α_2	α_3
	−1.654	**1.290**	**−0.299**	**−0.106**
	(0.063)	**(0.124)**	**(0.065)**	**(0.045)**
Rank	$\bar{\beta}_1$	γ_1	γ_2	γ_3
	−0.264	**−0.205**	**−0.049**	−0.004
	(0.017)	**(0.031)**	**(0.018)**	(0.010)
Time	α_4			
	0.051			
	(0.003)			

Table 2(b) Unobserved Heterogeneity Estimates in the Click-Through Model (Σ^{β})

	β_{i0} (*Intercept*)	β_{i1} (*Rank*)
β_{i0} (*Intercept*)	**1.053**	**−0.095**
	(0.078)	**(0.014)**
β_{i1} (*Rank*)	**−0.095**	**0.035**
	(0.014)	**(0.004)**

Note. Posterior means and posterior standard deviations (in parentheses) are reported, and estimates that are significant at 95% are bolded in Tables 2–7.

is negative and significant, indicating that keyword advertisements that contain brand-specific information are associated with a 56.6% decrease in click-through rates. These results imply that keyword advertisements that explicitly contain information identifying the advertiser are associated with higher click-through rates, whereas those that explicitly contain information identifying the brand are associated with lower click-through rates than keywords which lack such information. In contrast, the coefficient of *Length* in Table 2(a) is negative, suggesting that longer keywords typically tend be associated with lower click-through rates. Specifically, we find that all else being equal, an increase in the length of the keyword by one word is associated with a decrease in the click-through rates of 13.9%.

Intuitively, this result has an interesting implication if one were to tie this result with that in the literature on consideration sets in marketing. A longer keyword typically tends to suggest a more directed or specific search, whereas a shorter keyword typically suggests a more generic search. The shorter the keyword is, the less information it likely carries and the larger the context that should be supplied to focus the search (Finkelstein et al. 2002). This implies that the consideration set for the consumer is likely to shrink as the search term becomes "narrower" in scope. Danaher and Mullarkey (2003) show that user involvement during search (whether the use is in a purchasing or surfing mode) plays a crucial role in the effectiveness of online banner ads. One plausible explanation

[7] Ruud (2000) demonstrates that if Γ is restricted so that its determinant is a known constant, then $\log|\det\Gamma|$ plays no role in the in the maximization of the log likelihood, which leaves behind a seemingly unrelated regressions model that can be fit by generalized least squares. In the recursive case, in which the diagonal elements of the triangular matrix Γ are all equal to 1, $\det\Gamma = 1$ and the simultaneous equations log-likelihood function has the functional form of the seemingly unrelated regressions log-likelihood function. As a result, the full information maximum likelihood estimation function simplifies to generalized least squares.

is related to the extent of user involvement. Because consumers get to view ads of competing retailers, the probability of a goal-directed consumer clicking on the retailer's advertisement decreases unless the retailer carries the specific product that the consumer is searching for. In contrast, a consumer who does not have a goal-directed search (has a wider consideration set) and is in the surfing mode is likely to click on several advertising links before she finds a product that induces a purchase. Another plausible reason is that the quality of results retrieved by search engines for longer, more specific keywords is poorer than that for shorter, more generic ones. This can lead to reduced click-through rates for longer keywords.

Some additional substantive results are as expected. *Rank* has an overall negative relationship with *CTR* in Table 2(a). This implies that lower the rank of the advertisement (i.e., the higher the location of the sponsored ad on the computer screen) is, the higher the click-through rate is. The position of the advertisement link on the search engine page clearly plays an important role in influencing click-through rates. This kind of primacy effect is consistent with other empirical studies of the online world. Ansari and Mela (2003) suggested a positive relationship between the serial position of a link in an e-mail and recipients' clicks on that link. Similarly, Drèze and Zufryden (2004) implied a positive relationship between a link's serial position and site visibility. Brooks (2004) showed that the higher the link's placement in the results listing, the more likely a searcher is to select it. In the context of shopping search engines, Baye at al. (2009) find that there is a 17.5% drop in click-through rates when a retailer is moved down one position on the screen. Brynjolfsson et al. (2004) find similar evidence of the primacy effect in their study on shopbots. Thus, website designers, and online advertising managers would place their most desirable links toward the top of a webpage or e-mail and their least desirable links toward the bottom. A robustness test wherein we include a quadratic term for *Rank* highlights that the negative relationship between *CTR* and *Rank* increases at a decreasing rate. This finding has useful implications for managers interested in quantifying the impact of *Rank* on *CTR*.

When we consider the interaction effect of these variables on the relationship of *Rank* with click-through rates, we find that keywords that contain retailer- or brand-specific information are associated with an increase in the negative relationship between *Rank* and *CTR*. That is, for keywords that contain retailer- or brand-specific information, a lower rank (better placement) is associated with even higher click-through rates. However, we find that the coefficient of *Length* is statistically insignificant, suggesting that longer keywords do not seem to affect the negative relationship between click-through rates and ranks. As shown in Table 2(b), the estimated unobserved heterogeneity covariance is significant, including all of its elements. 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 unobserved factors beyond the three observed keyword characteristics.

5.1.2. Conversion Rate. Next consider Tables 3(a) and 3(b) with findings on conversion rates. Our analysis reveals that the coefficient of *Brand*, δ_2, is negative and significant, indicating that keywords that contain information specific to a brand (either product specific or manufacturer specific) experience lower conversion rates on an average. Specifically, the presence of brand information in the keyword decreases conversion rates by 44.2%. Similarly, the presence of retailer information in the keyword increases conversion rates by 50.6%. In contrast, *Length* is not statistically significant in its overall effect on conversion rates.

We find a significant relationship between *Rank* and *Conversion Rates*, such that the lower the *Rank* is (i.e., the higher the position of the keyword on the screen), the higher the *Conversion Rate* is. A decrease in the rank from the maximum possible position or worst case scenario (which is 131 in our data) to the minimum position or best-case scenario (which is 1 in our data) increases conversion rates by 92.5%. It is useful to discuss what this result suggests. Note that a prominent (or top) position on the search engine results page can be associated with at least two countervailing effects. First, a prominent position can be

Table 3(a) **Coefficient Estimates on Conversion Rate**

	Intercept	*Retailer*	*Brand*	*Length*	*LandingPageQuality*
Intercept	$\bar{\theta}_0$	δ_1	δ_2	δ_3	δ_4
	−4.457	**1.123**	**−0.879**	−0.041	**0.152**
	(0.097)	**(0.234)**	**(0.136)**	(0.110)	**(0.066)**
Rank	$\bar{\theta}_1$	κ_1	κ_2	κ_3	κ_4
	−0.282	−0.032	0.014	0.012	0.013
	(0.031)	(0.089)	(0.036)	(0.023)	(0.014)
Time	δ_5				
	0.067				
	(0.009)				

Table 3(b) **Unobserved Heterogeneity Estimates in the Conversion Model (Σ^θ)**

	θ_{i0} (*Intercept*)	θ_{i1} (*Rank*)
θ_{i0} (*Intercept*)	**1.436**	**−0.131**
	(0.285)	**(0.030)**
θ_{i1} (*Rank*)	**−0.131**	**0.058**
	(0.030)	**(0.007)**

associated with a high "quality" or "trust" perception in consumers' minds, where they associate higher positions with a higher product quality (Brooks 2004). This can induce higher conversion rates for the top positions. Yet Agarwal et al. (2008) point to the possibility in which nonserious buyers often click on the top slots but do not purchase and serious buyers buy from the middle positions because of a recency bias. This can induce lower conversion rates for the top positions. In our data, the first effect dominates the second effect, leading to higher conversion rates for the top positions. This finding of the relationship between rank and conversion rates can have an important implication for existing theoretical models in sponsored search advertising, which has typically assumed that the value per click to an advertiser is uniform across all ranks on the search engine results page. Our estimates suggest instead that the value per click is not uniform and thereby motivates future theoretical models that could modify this assumption to reexamine the social welfare-maximizing properties of generalized second price keyword auctions on the Internet.

As speculated in trade press reports, our analysis empirically confirms that *LandingPageQuality* has a positive relationship with conversation rates. To be precise, an increase in the landing page quality score from the lowest possible score (equal to 1) to the highest possible score (equal to 10) is associated with an increase in the conversion rates of 22.5%. These analyses suggest that in terms of magnitude, the rank of a keyword on the search engine has a larger impact on conversion rates than the quality of the landing pages does.

When we consider the effect of these keyword characteristics on the relationship of *Rank* with *Conversion Rates*, we find that none of the keyword attributes has a statistically significant effect on the relationship between rank and conversion rates. As shown in Table 3(b), the estimated unobserved heterogeneity covariance is significant, including all of its elements. 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.

5.1.3. Cost per Click. Next, we turn to advertiser bidding behavior, as shown by estimates in Tables 4(a) and 4(b). Interestingly, the analysis of CPC reveals that there is a negative relationship between *CPC* and *Retailer* but a positive relationship between *CPC* and *Brand*. This implies that the firm incurs a lower CPC for advertisements that contain retailer information and higher CPC for advertisements that contain brand information. This is consistent with theoretical predictions because *Retailer* keywords are

Table 4(a) Coefficient Estimates on CPC

	Intercept	*Retailer*	*Brand*	*Length*	*LandingPageQuality*
Intercept	$\bar{\omega}_0$	λ_1	λ_2	λ_3	λ_4
	−1.660	**−0.760**	**0.139**	−0.022	**−0.036**
	(0.024)	**(0.069)**	**(0.032)**	(0.023)	**(0.016)**
	$\bar{\omega}_1$	ρ_{11}	ρ_{12}	ρ_{13}	ρ_{14}
Lag_Rank	**−0.041**	**0.036**	−0.008	**0.018**	−0.001
	(0.005)	**(0.010)**	(0.008)	**(0.005)**	(0.004)
Time	λ_6				
	−0.020				
	(0.001)				

Table 4(b) Unobserved Heterogeneity Estimates in the CPC Model (Σ^{ω})

	ω_{i0} (*Intercept*)	ω_{i1} (*Lag_Rank*)
ω_{i0} (*Intercept*)	**0.555**	**−0.021**
	(0.030)	**(0.005)**
ω_{i1} (*Lag_Rank*)	**−0.021**	**0.011**
	(0.005)	**(0.001)**

far less competitive than *Brand* keywords, on average. Although *Length* does not have a direct statistically significant effect on *CPC*, it indirectly affects *CPC* through the interaction with *Rank*. There is a negative and statistically significant relationship between *CPC* and *LandingPageQuality*, implying that advertisers tend to place lower bids on keywords that are linked to landing pages with higher quality.

Further, there is a negative relationship between *CPC* and *Lag_Rank*, such that a more prominent (lower rank) position on the search engine results screen is associated with a higher CPC and hence a higher actual bid. These results suggest that although there is some learning exhibited by the firm during the bidding process based on past performance metrics, it may not necessarily be bidding in the most profitable manner.

5.1.4. Rank. Finally, on the analysis of *Rank*, we find that all three covariates—*Retailer*, *Brand*, and *Length*—have a statistically significant and negative relationship with *Rank*, suggesting that the search keywords that have retailer-specific information or brand-specific information or are more specific in their scope generally tend to have lower ranks (i.e., they are listed higher on the search engine results screen).

How do search engines decide on the final rank? Anecdotal evidence and public disclosures by Google suggest that Google incorporates a performance criterion along with bid price when determining the ranking of the advertisers. The advertiser in the top position might be willing to pay a higher CPC than the advertiser in the second position, but there is no

Table 5(a) Coefficient Estimates on Keyword Rank

	Intercept	*Retailer*	*Brand*	*Length*
Intercept	$\bar{\phi}_0$	τ_1	τ_2	τ_3
	1.954	**−0.213**	**−0.279**	**−0.172**
	(0.031)	**(0.075)**	**(0.037)**	**(0.030)**
CPC	$\bar{\phi}_1$	π_1	π_2	π_3
	−2.028	0.361	0.185	−0.003
	(0.093)	(0.306)	(0.108)	(0.085)
Lag_CTR	$\bar{\phi}_2$			
	−1.289			
	(0.046)			
Time	τ_5			
	0.031			
	(0.001)			

Table 5(b) Unobserved Heterogeneity Estimates in the Keyword Rank Model (Σ^{θ})

	$\bar{\phi}_0$ (*Intercept*)	$\bar{\phi}_1$ (*CPC*)
$\bar{\phi}_0$ (*Intercept*)	**1.020**	**−1.677**
	(0.048)	**(0.108)**
$\bar{\phi}_1$ (*CPC*)	**−1.677**	**4.073**
	(0.108)	**(0.294)**

guarantee that its ad will be displayed in the first slot. This is because past performance, such as prior click-through rates, are factored in by Google before the final ranks are published. The coefficients of *CPC* and *Lag_CTR* are negative and statistically significant in our data. Thus, our results from the estimation of the *Rank* equation confirm that the search engine is indeed incorporating both the current CPC bid and the previous click-through rates in determining the final rank of a keyword. Note from Table 5(a) that the coefficient of *CPC* is almost twice the coefficient of *Lag_CTR*, suggesting that current bid price (CPC) has a larger role to play in determining the final rank than the "quality score"-related factors like prior click-through rates.

It is worth noting in Table 6 that the unobserved covariance between (i) click-through propensity and keyword rank, (ii) conversion propensity and keyword rank, and (iii) CPC and keyword rank all turn out to be statistically significant. This suggests the endogenous nature of CPC and Rank. Therefore, it is important to simultaneously model the consumer's click-through and purchase behavior and the advertiser's and search engine's decisions.

As mentioned before, we provide the parameter estimates produced by the estimation of this system under the assumption of diagonality (restricting covariance elements to be zero) to the generalized results. Refer to tables in the online appendix. These estimates further demonstrate that it is important to

Table 6 Estimated Covariance Across Click-Through, Conversion, CPC, and Rank (Ω)

	Click-Through	*Conversion*	*CPC*	*Rank*
Click-Through	**0.956**	**1.092**	**−0.082**	**0.472**
	(0.055)	**(0.086)**	**(0.009)**	**(0.022)**
Conversion	**1.092**	**2.429**	**−0.213**	**0.528**
	(0.086)	**(0.158)**	**(0.021)**	**(0.043)**
CPC	**−0.082**	**−0.213**	**0.220**	−0.003
	(0.009)	**(0.021)**	**(0.004)**	(0.005)
Rank	**0.472**	**0.528**	−0.003	**0.319**
	(0.022)	**(0.043)**	(0.005)	**(0.007)**

control for endogeneity, because the parameter estimates are attenuated when we restrict the covariance elements to be zero and thus are biased.

5.1.5. Profit. Finally, based on the above estimates and the summary statistics of the data, we find that profits are not necessarily the highest in the top slots—rather, profits are often higher in the middle positions than those in the most prominent (top) or least prominent (bottom) positions. We plotted profits with rank for several different keywords in our sample and consistently found that profits are higher in ranks 4–6 compared to ranks 1–3 or ranks 7–10. This was also true for an ad associated with a product priced at $22, which was the average price of an item in our sample. Despite conversion rates being the highest at the topmost slot (rank 1), this counterintuitive result occurs because the CPC for the middle (ranks 4–6) or bottom slots (ranks 7–10) on the search engine results page decays much faster that those for the most prominent slots (ranks 1–3).

5.2. Further Robustness Tests

We conducted a few additional robustness tests with additional data to examine whether the key results remain consistent. To control for other competitors' bid prices for a given keyword, we collected data from Google's keyword pricing tool, which gives estimates of advertisers' maximum CPC for any given keyword.[8] In addition, note that the final rank allocation for a given keyword is based on a second price auction, so a keyword's rank will be influenced by competitors' bid prices. Hence, we also control for the effect of competition on the search engine's ranking decision. Note also that the current period bid can also based on the extent of profits from that

[8] Google's keyword estimator tools in fact give two key pieces of information—the estimated upper and lower range for the CPC of that keyword (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 bid values to construct the *Competitor Price* variable and control for possible competitive effects of other advertisers.

keyword—a heuristic used by some firms in sponsored search advertising (Gerstmeier et al. 2009).[9] We find that the qualitative nature of our main results remains unchanged with the inclusion of these variables (see the online appendix).[10] Note that one needs to be cautious about these tests because competitor prices and lagged profits may be endogenous. For example, keyword profitability may be correlated across competitors, leading to correlation between this variable and the error term. Similarly, lagged profits may be correlated with the current profitability conditions. Hence, to account for the endogeneity of competitor price and lagged profits, we also estimate the entire system of equations by using two period lags of competitor price and profits as instruments. The qualitative nature of our key results remains the same.

As a robustness test, we also ran our empirical analyses using a quadratic term for "Rank" in addition to the linear term in both the conversion rate and click-through rate equation. This helped us examine the rate of change of clicks and conversions with position and controls for the fact that the relationship may not be linear. We found that the qualitative nature of our main results remained unchanged when we included a nonlinear term for Rank in both click-through and conversion rate equations. Furthermore, the inclusion of a quadratic term for *Rank* highlights that the negative relationship between *Conversion Rates* and *Rank* increases at a decreasing rate (see the online appendix). To our knowledge, this finding is relatively new in the literature on online advertising.[11]

6. Managerial Implications and Conclusion

The phenomenon of sponsored search advertising is gaining ground as the largest source of revenue for search engines. In this research, we focus on building a model that analyzes the relationship between different keyword-level covariates and different metrics of sponsored search advertisement performance taking consumer, advertiser, and search engine behavior into account. Our analyses reveal that there is a considerable amount of heterogeneity in terms of the effect of different kinds of keywords on the decision metrics of the various players—consumers, advertisers, and search engines.

Arguably, the mix of retailer-specific and brand-specific keywords in an online advertiser's portfolio has some analogies to other kinds of marketing mix decisions faced by firms in many markets. For instance, typically it is the retailer who engages in "retail store" advertising that has a relatively "monopolistic" market. In contrast, typically it is the manufacturer who engages in advertising national brands. From the retailer's perspective, these brand-specific advertisements are likely to be relatively more "competitive," because national brands are likely to be stocked by its competitors, too. Retailer name searches are navigational searches and are analogous to a user finding the retailer's or address in the White Pages. These searches are driven by brand awareness generated by catalog mailings, TV ads, etc., and are likely to have come from more "loyal" consumers. Even though the referral to the retailer's website came through a search engine, the search engine had very little to do with generating the demand in the first place. In contrast, searches on product or manufacturer-specific names are analogous to consumers going to the Yellow Pages—they know they need a branded product, but don't yet know where to buy it (Rimm-Kaufman 2007). These are likely to be "competitive" searches. If the advertiser wins the click and the order, that implies that it has taken market share away from a competitor. Thus, retailer-specific keywords are likely to be searched for and clicked by "loyal" consumers who are inclined to buying from that retailer, whereas brand-specific keywords are likely to be searched for and clicked by the shoppers who can easily switch to the competition. This would suggest that advertisers experience higher conversion rates on retailer-specific keywords and lower conversion rates on brand-specific keywords, a feature that we also observe in our data.

Our results provide descriptive insights into sponsoring such retail store keywords (retailer-specific keywords) with national-brand keywords (brand-specific keywords). Most firms that sponsor online keyword advertisements set a daily budget, select a set of keywords, determine a bid price for each keyword, and designate an ad associated with each selected keyword based on the kind of match type (broad, exact, or phrase). If the company's spending has exceeded its daily budget, however, its ads will not be displayed. With millions of available keywords and a highly uncertain click-through rate associated with each keyword, identifying the most profitable set of keywords given the daily budget constraint becomes challenging (Rusmevichientong and Williamson 2006). Analyzing click-through and conversion rates provide key insights into the cost per

[9] To normalize the distribution of this variable, we took the log(*Profit*).

[10] 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 based on data from the advertiser.

[11] As a further robustness test, we ran panel data models with keyword-level fixed effects to examine the relationship between rank and click-through rates as well as conversion rates. The qualitative nature of our main results remains unchanged; moreover, this holds both with and without the quadratic term.

conversion and the value per click of different keywords. Such empirical analyses could supplement the broader keyword selection techniques based on popularity and economic impact of occurrence of keywords in user-generated content sites such as reputation systems, product review forums, and blogs (Archak et al. 2008, Dhar and Ghose 2009).

We quantify the impact of landing page quality on product conversions and CPC of search advertising. A high landing page quality can also boost the organic or nonsponsored rankings of that retailer for a keyword. This is because organic rankings of advertisers' websites are based on a complex and proprietary algorithm devised by the search engine that involves the quality of the landing page and the website's "relative importance" with respect to other links. This can be important because of empirical evidence that more users will click on an advertiser's link if it is listed in both paid and organic listings, due to "second opinion" or "reinforcement effect" resulting in more conversions and higher revenues as well (Yang and Ghose 2008). An example of content that can increase the landing page quality score and thus lead to improvements in the ranking of an advertiser on paid and organic listings is the presence of user-generated content. Besides helping boost the ranking of organic listings, the presence of this content can also increase the landing page quality scores for paid search.

Our results have some similarities with the findings in the context of traditional media advertising in offline markets. Koschat and Putsis (2002) attempt to estimate the effect of unbundling in magazine advertising. They find that, in terms of the pricing of magazine advertising space, targeting specific reader segments is generally preferable to offering advertisers all the readers. This is consistent with our finding that advertisers incur a higher CPC for brand-specific keywords (that are relatively more targeted) than for generic keywords that do not highlight the manufacturer or product brand. Wilbur (2008) empirically examines the determinants of television advertising pricing to estimate viewer demand for programs and advertiser demand for audiences. His results suggest that advertiser preferences influence network choices more strongly than viewer preferences. This has an interesting parallel to our finding that search engines place a higher weight on advertisers' bid prices relative to consumer click-through rates in deciding their choice of rank for a given ad in a given auction. Using circulation data for U.S. daily newspapers, Chandra (2009) shows that newspapers facing more competition have lower circulation prices but higher advertising prices than similar newspapers facing little or no competition. This, however, differs from our finding that advertisers tend to incur a lower CPC on longer keywords (narrower searches), although this relationship is not statistically significant in our data. This suggests that the bidding behavior of this advertiser in sponsored search auctions is not optimal and is consistent with the findings of Gerstmeier et al. (2009), who show that a bidding heuristic based on "rank" only is not optimal. However, this is a topic that merits more detailed analysis, which is beyond the scope of this paper.

Another interesting observation that has similarities with the behavior of offline ads is the relationship between rank and profitability in search auctions. In the context of selling medical services, Tellis et al. (2001) find that effective TV ads that generate referrals may not necessarily be profitable, too. This is consistent with our data, which suggests that ads that have higher click-through rates may have lower profits (because of a higher CPC and lower revenues) than other ads. In fact, we see in our data that profit can be nonmonotonic with rank such that ads placed in the more prominent positions do not always maximize profit: profits are often higher in the middle positions than those in the top or bottom positions. This is similar to the findings of Agarwal et al. (2008), who also find that profits first increase and then decrease with ad position. However, the factors that drive this counterintuitive result are somewhat different in our paper than in theirs. In their paper, it is the combination of an increase in conversion rates and cost decay that result in lower profits at the more prominent positions. In our paper, the aggressive bidding behavior increases the total advertisement costs (given the high click-through rates) and results in lower profits despite the high conversion rates at the more prominent positions. Put simply, whereas revenues and costs both decrease for the less prominent slots than for the more prominent slots in our sample, CPC decays at a much faster rate, leading to lower profits from the topmost slots than from the middle or bottom slots.

Another aspect of sponsored search advertising that also seems important to keep in mind is that even if clicks on certain keywords do not lead to conversions in the same session, the mere act of repetitive exposure of a consumer to a stimulus can increase the user's familiarity with the brand name and lead to a brand preference (Tellis 2004). This in turn enhances the effectiveness of future advertising. Moreover, even in sponsored search, there is evidence that some keywords are often meant to perform the role of assisting in conversions through more specific keywords—a kind of spillover effect that has been documented in other studies (Rutz and Bucklin 2008, Ghose and Yang 2008). In other words, sponsored advertising can contribute to additional consumer purchases beyond the original intent and thus generate a longer-term business value for the advertiser.

Our estimates from the conversion rate equation show that an advertiser's relative value per click for each slot is not uniform. Instead, this decreases as one goes down the search engine results page, meaning that clicks from more prominent positions are more valuable than clicks from the lower slots. Prior work (e.g., Edelman et al. 2007, Varian 2007) showed that in a model where all clicks on an ad gain the advertiser the same value, generalized second price keyword auction maximizes the social welfare in equilibrium. Given that the probability that a click will convert to an actual action (e.g., a sale) for the advertiser depends on the position (rank) of the ad on the search engine results page, it is worth examining whether the equilibrium results of prior work necessarily hold. Our empirical results can thus pave the way for future theoretical models in this domain that could relax assumptions to design newer mechanisms with more robust equilibrium properties. A conclusion that can be made from our study is that in CPC models of online search advertising the value generated by clicks may vary for a number of reasons and that this should be taken into account in the design of the advertising mechanism.

Our paper has several limitations. These limitations arise primarily from the lack of data. For example, we do not have precise data on competition, because our data are limited to one firm. That is, we do not know the keyword ranks or other performance metrics of keyword advertisements of the competitors of the firm whose data we have used in this paper. Further, we do not have any knowledge of other information that was mentioned in the textual description in the space following a paid advertisement during consumers' search queries. Future work could integrate that information with our modeling approach to have more precise estimates. In addition, future work could examine product-specific characteristics to see how different kinds of products affect the click-through and conversion rates in different ways. This will help firms analyze which brands or products have higher conversions and lower costs per conversion. Future work could also examine firm-specific characteristics to see how differences in the size of the advertiser (small, medium, or large enterprises), type of advertiser (brick and mortar firms versus pure online firm), and size of search engine (Google, Yahoo, and Microsoft) affect consumer behavior and advertiser bidding strategies. Another area for future work is to study whether keyword advertising acts like a coupon by always inducing an immediate purchase or more like a regular ad that can induce a delayed purchase, as is often seen in traditional media. This sort of analysis requires access to highly granular consumer-level data that captures whether exposure to a sponsored ad in one session resulted in a conversion in the same session or in a subsequent session. We hope that this study will generate further interest in exploring this important and emerging interdisciplinary area.

7. Electronic Companion

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

Acknowledgments

The authors thank the anonymous company that provided data for this study. The authors are listed in alphabetical order and contributed equally. They are grateful to the department editor, associate editor, and three anonymous referees for extremely helpful comments. The authors also thank Susan Athey, Michael Baye, Gérard Cachon, Avi Goldfarb, Greg Lewis, Bill Greene, and participants at Carnegie Mellon University, McGill University, New York University, Purdue University, University of Calgary, University of Connecticut, University of California at Irvine, University of Pennsylvania, University of Washington, Microsoft Research, UT Dallas Marketing Conference, the 2008 IIO Conference, the 2008 NET Institute Conference, the 2008 Marketing Science Institute Conference, the WSDM 2008 Conference, the WISE 2007 Conference, and INFORMS-CIST 2007 for helpful comments. The first author acknowledges the generous financial support from NSF CAREER Award IIS-0643847. Funding for this project was also provided by the NET Institute, the Marketing Science Institute, and an NYU-Poly Seed Fund grant. The usual disclaimer applies.

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, in the application presented in the paper. We report below the MCMC algorithm for the simultaneous model of click-through rate, conversion rate, bid price, and keyword rank.

Step 1. Draw c_{ij}^p and c_{ij}^q:

As specified, the likelihood function of the number of clicks (n_{ij}) and number of purchases (m_{ij}) is

$$l(c_{ij}^p, c_{ij}^q \mid n_{ij}, m_{ij}) \propto \{p_{ij}q_{ij}\}^{m_{ij}}\{p_{ij}(1-q_{ij})\}^{n_{ij}-m_{ij}}\{1-p_{ij}\}^{N_{ij}-n_{ij}},$$

where

$$p_{ij} = \frac{\exp(c_{ij}^p)}{1+\exp(c_{ij}^p)}, \quad \frac{q_{ij} = \exp(c_{ij}^q)}{1+\exp(c_{ij}^q)};$$

$$c_{ij}^p = \bar{m}_{ij}^p + \varepsilon_{ij};$$

$$\bar{m}_{ij}^p = \beta_{i0} + \beta_{i1} Rank_{ij} + \alpha_1 Retailer_i + \alpha_2 Brand_i + \alpha_3 Length_i + \alpha_4 Time_{ij};$$

$$c_{ij}^q = \bar{m}_{ij}^q + \eta_{ij};$$

$$\bar{m}^q_{ij} = \theta_{i0} + \theta_{i1}Rank_{ij} + \delta_1 Retailer_i + \delta_2 Brand_i + \delta_3 Length_i + \delta_4 LandingPageQuality_i + \delta_5 Time_{ij}.$$

We further define the following notations:

$$D = \Omega^*_{11} - \Omega^*_{12}\Omega^{*-1}_{22}\Omega^*_{21};$$

$$\Omega^*_{11} = \begin{bmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{bmatrix}, \quad \Omega^*_{22} = \begin{bmatrix}\Omega_{33} & \Omega_{34}\\ \Omega_{43} & \Omega_{44}\end{bmatrix},$$

$$\Omega^*_{12} = \Omega'^*_{21} = \begin{bmatrix}\Omega_{13} & \Omega_{14}\\ \Omega_{23} & \Omega_{24}\end{bmatrix};$$

$$u_{ij1} = \ln(CPC_{ij}) - (\omega_{i0} + \omega_{i1}Rank_{i,j-1} + \lambda_1 Retailer_i + \lambda_2 Brand_i + \lambda_3 Length_i + \lambda_4 LandingPageQuality_i + \lambda_5 Time_{ij});$$

$$u_{ij2} = \ln(Rank_{ij}) - (\phi_{i0} + \phi_{i1}CPC_{ij} + \bar{\phi}_2 CTR_{i,j-1} + \tau_1 Retailer_i + \tau_2 Brand_i + \tau_3 Length_i + \tau_4 Time_{ij});$$

$$E_{ij} = \Omega^*_{12}\Omega^{*-1}_{22}u_{ij}.$$

We use Metropolis-Hastings algorithm with a random walk chain to generate draws of $c_{ij} = (c^p_{ij}, c^q_{ij})$ (see Chib and Greenberg 1995, p. 330, method 1). Let $c^{(p)}_{ij}$ denote the previous draw; then the next draw $c^{(n)}_{ij}$ is given by

$$c^{(n)}_{ij} = c^{(p)}_{ij} + \Delta,$$

with the accepting probability α given by

$$\min\left\{\frac{\exp[-\frac{1}{2}(c^{(n)}_{ij} - \bar{m}_{ij} - E_{ij})' D^{-1}(c^{(n)}_{ij} - \bar{m}_{ij} - E_{ij})]\, l(c^{(n)}_{ij})}{\exp[-\frac{1}{2}(c^{(p)}_{ij} - \bar{m}_{ij} - E_{ij})' D^{-1}(c^{(p)}_{ij} - \bar{m}_{ij} - E_{ij})]\, l(c^{(p)}_{ij})}, 1\right\}.$$

Δ is a draw from the density Normal(0, 0.015I), where I is the identity matrix.

Step 2. Draw $b_i = [\beta'_i, \theta'_i, \omega'_i, \phi'_i]'$:

$$y_{ij1} = c^p_{ij} - (\alpha_1 Retailer_i + \alpha_2 Brand_i + \alpha_3 Length_i + \alpha_4 Time_{ij});$$

$$y_{ij2} = c^q_{ij} - (\delta_1 Retailer_i + \delta_2 Brand_i + \delta_3 Length_i + \delta_4 LandingPageQuality_i + \delta_5 Time_{ij});$$

$$y_{ij3} = \ln(CPC_{ij}) - (\lambda_1 Retailer_i + \lambda_2 Brand_i + \lambda_3 Length_i + \lambda_4 LandingPageQuality_i + \lambda_5 Time_{ij});$$

$$y_{ij4} = \ln(Rank_{ij}) - (\bar{\phi}_2 CTR_{i,j-1} + \tau_1 Retailer_i + \tau_2 Brand_i + \tau_3 Length_i + \tau_4 Time_{ij});$$

$$x_{ij} = \begin{bmatrix} x'_{ij1} & 0 & 0 & 0\\ 0 & x'_{ij2} & 0 & 0\\ 0 & 0 & x'_{ij3} & 0\\ 0 & 0 & 0 & x'_{ij4}\end{bmatrix}, \quad \Sigma = \begin{bmatrix}\Sigma^\beta & 0 & 0 & 0\\ 0 & \Sigma^\theta & 0 & 0\\ 0 & 0 & \Sigma^\omega & 0\\ 0 & 0 & 0 & \Sigma^\phi\end{bmatrix};$$

$$x_{ij1} = x_{ij2} = [1, Rank_{ij}], \quad x_{ij3} = [1, Rank_{i,j-1}, Profit_{i,j-1}]',$$

$$x_{ij4} = [1, CPC_{ij}]';$$

$$\bar{b}_{i1} = \bar{\beta}_0, \quad \bar{b}_{i2} = \bar{\beta}_1 + \gamma_1 Retailer_i + \gamma_2 Brand_i + \gamma_3 Length_i;$$

$$\bar{b}_{i3} = \bar{\theta}_0, \quad \bar{b}_{i4} = \bar{\theta}_1 + \kappa_1 Retailer_i + \kappa_2 Brand_i + \kappa_3 Length_i + \kappa_4 LandingPageQuality_i;$$

$$\bar{b}_{i5} = \bar{\omega}_0, \quad \bar{b}_{i6} = \bar{\omega}_1 + \rho_{11} Retailer_i + \rho_{12} Brand_i + \rho_{13} Length_i + \rho_{14} LandingPageQuality_i;$$

$$\bar{b}_{i7} = \bar{\omega}_2 + \rho_{21} Retailer_i + \rho_{22} Brand_i + \rho_{23} Length_i + \rho_{24} LandingPageQuality_i, \quad \bar{b}_{i8} = \bar{\phi}_0;$$

$$\bar{b}_{i9} = \bar{\phi}_1 + \pi_1 Retailer_i + \pi_2 Brand_i + \pi_3 Length_i.$$

Then $b_i \sim \mathrm{MVN}(A_i, B_i)$:

$$B_i = [x'_i\Omega^{-1}x_i + \Sigma^{-1}]^{-1}, \quad A_i = B_i[x'_i\Omega^{-1}y_i + \Sigma^{-1}\bar{b}_i].$$

Step 3. Draw $a = [\alpha', \delta', \lambda', \bar{\phi}_2, \tau']'$:

$$y_{ij1} = c^p_{ij} - (\beta_{i0} + \beta_{i1}Rank_{ij});$$

$$y_{ij2} = c^q_{ij} - (\theta_{i0} + \theta_{i1}Rank_{ij});$$

$$y_{ij3} = \ln(CPC_{ij}) - (\omega_{i0} + \omega_{i1}Rank_{i,j-1});$$

$$y_{ij4} = \ln(Rank_{ij}) - (\phi_{i0} + \phi_{i1}CPC_{ij});$$

$$x_{ij} = \begin{bmatrix} x'_{ij1} & 0 & 0 & 0\\ 0 & x'_{ij2} & 0 & 0\\ 0 & 0 & x'_{ij3} & 0\\ 0 & 0 & 0 & x'_{ij4}\end{bmatrix};$$

$$x_{ij1} = [Retailer_i, Brand_i, Length_i, Time_{ij}];$$

$$x_{ij2} = [Retailer_i, Brand_i, Length_i, LandingPageQuality_i, Time_{ij}];$$

$$x_{ij3} = [Retailer_i, Brand_i, Length_i, LandingPageQuality_i, Time_{ij}];$$

$$x_{ij4} = [CTR_{i,j-1}, Retailer_i, Brand_i, Length_i, Time_{ij}];$$

$$\bar{a} = 0_{21x1}, \Sigma_0 = 100I.$$

Then $a \sim \mathrm{MVN}(A, B)$:

$$B = [X'\Omega^{-1}X + \Sigma^{-1}]^{-1}, \quad A = B[X'\Omega^{-1}Y + \Sigma_0^{-1}\bar{a}_0].$$

Step 4. Draw Ω:

$$y_{ij1} = c^p_{ij} - (\beta_{i0} + \beta_{i1}Rank_{ij} + \alpha_1 Retailer_i + \alpha_2 Brand_i + \alpha_3 Length_i + \alpha_4 Time_{ij});$$

$$y_{ij2} = c^p_{ij} - (\theta_{i0} + \theta_{i1}Rank_{ij} + \delta_1 Retailer_i + \delta_2 Brand_i + \delta_3 Length_i + \delta_4 LandingPageQuality_i + \delta_5 Time_{ij});$$

$$y_{ij3} = \ln(CPC_{ij}) - (\omega_{i0} + \omega_{i1}Rank_{i,j-1} + \lambda_1 Retailer_i + \lambda_2 Brand_i + \lambda_3 Length_i + \lambda_4 LandingPageQuality_i + \lambda_5 Time_{ij});$$

$$y_{ij4} = \ln(Rank_{ij}) - (\phi_{i0} + \phi_{i1}CPC_{ij} + \bar{\phi}_2 CTR_{i,j-1} + \tau_1 Retailer_i + \tau_2 Brand_i + \tau_3 Length_i + \tau_4 Time_{ij});$$

$$\Omega \sim \mathrm{IW}\left(\sum_i \sum_j y'_{ij}y_{ij} + Q_0, N + q_0\right);$$

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

Step 5. Draw Σ^{β}, Σ^{θ}, and Σ^{ω}:

$$\Sigma^{\beta} \sim \mathrm{IW}\left(\sum_i (\beta_i - \bar{\beta})'(\beta_i - \bar{\beta}) + Q_0, N + q_0\right);$$

$$Q_0 = 10I \text{ and } q_0 = 10; \quad n = \text{ no. of keywords;}$$

$$\Sigma^{\theta} \sim \mathrm{IW}\left(\sum_i (\theta_i - \theta)'(\theta_i - \theta) + Q_0, N + q_0\right);$$

$$Q_0 = 10I \text{ and } q_0 = 10; \quad n = \text{ no. of keywords;}$$

$$\Sigma^{\omega} \sim \mathrm{IW}\left(\sum_i (\omega_i - \bar{\omega})'(\omega_i - \bar{\omega}) + Q_0, N + q_0\right);$$

$$Q_0 = 10I \text{ and } q_0 = 10; \quad n = \text{ no. of keywords,}$$

where IW stands for the inverted Wishart distribution.

Step 6. Draw $f_1 = [\bar{\beta}_0, \bar{\beta}_1, \gamma_1, \gamma_2, \gamma_3]'$:

$$x_i = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & Retailer_i & Brand_i & Length_i \end{bmatrix};$$

$$a = 0_{5x1}, \Sigma_0 = 100I.$$

Then $f_1 \sim \mathrm{MVN}(A, B)$:

$$B = [X'\Sigma^{\beta^{-1}} X + \Sigma_0^{-1}]^{-1}, \quad A = B[X'\Sigma^{\beta^{-1}} \beta + \Sigma_0^{-1} \bar{a}_0].$$

Step 7. Draw $f_2 = [\bar{\theta}_0, \bar{\theta}_1, \kappa_1, \kappa_2, \kappa_3, \kappa_4]'$ similar to Step 6.

Step 8. Draw $f_3 = [\bar{\omega}_0, \bar{\omega}_1, \rho_{11}, \rho_{12}, \rho_{13}, \rho_{14}, \bar{\omega}_2, \rho_{21}, \rho_{22}, \rho_{23}, \rho_{24}]'$ similar to Step 6.

Step 9. Draw $f_4 = [\bar{\phi}_0, \bar{\phi}_1, \pi_1, \pi_2, \pi_3]'$ similar to Step 6.

References

Agarwal, A., K. Hosanagar, M. Smith. 2008. Location, location and location: An analysis of profitability of position in online advertising markets. Working paper, Carnegie Mellon University, Pittsburgh.

Aggarwal, G., J. Feldman, S. Muthukrishnan. 2006. Bidding to the top: VCG and equilibria of position-based auctions. *Approximation and Online Algorithms*. Lecture Notes in Computer Science. Springer, Berlin, 15–28.

Animesh, A., V. Ramachandran, S. Viswanathan. 2009. Quality uncertainty and the performance of online sponsored search markets: An empirical investigation. *Inform. Systems Res.*, ePub ahead of print June 12, http://isr.journal.informs.org/cgi/content/abstract/isre.1080.0222v1.

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

Archak, N., A. Ghose, P. Ipeirotis. 2008. Deriving the pricing power of product features by mining consumer reviews. Working paper, New York University, New York.

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

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

Baye, M., R. Gatti, P. Kattuman, J. Morgan. 2009. Clicks, discontinuities, and firm demand online. *J. Econom. Management Strategy*. Forthcoming.

Börgers, T., I. Cox, M. Pesendorfer, V. Petricek. 2007. Equilibrium bids in sponsored search auctions: Theory and evidence. Mimeo, University of Michigan, Ann Arbor.

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

Broder, A. 2002. Taxonomy of Web search. *SIGIR Forum* **36** 3–10.

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

Brynjolfsson, E., A. Dick, M. Smith. 2004. Search and product differentiation at an Internet shopbot. Working paper, Massachusetts Institute of Technology, Cambridge.

Chandra, A. 2009. Targeted advertising: The role of subscriber characteristics in media markets. *J. Indust. Econom.* **57**(1) 58–84.

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

Chen, Y., C. He. 2006. Paid placement: Advertising and search on the Internet. Working paper, University of Colorado at Boulder, Boulder.

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

Danaher, P., G. Mullarkey. 2003. Factors affecting online advertising recall: A study of students. *J. Advertising Res.* (September) 252–267.

Dhar, V., A. Ghose. 2009. Sponsored search and market efficiency. Working paper, New York University, New York.

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

Drèze, X., F. Zufryden. 2004. The measurement of online visibility and its impact on Internet traffic. *J. Interactive Marketing* **18**(1) 20–37.

Edelman, B., M. Ostrovsky, M. Schwarz. 2007. Internet advertising and the generalized 2nd-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* **97**(1) 242–259.

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

Finkelstein, L., E. Gabrilovich, Y. Matias, E. Rivlin, Z. Solan, G. Wolfman, E. Ruppin. 2002. Placing search in context: The concept revisited. *ACM Trans. Inform. Systems* **20**(1) 116–131.

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

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.

Ghose, A., S. Yang. 2008. Modeling cross-category purchases in sponsored search advertising. Working paper, New York University, New York.

Goldfarb, A., C. Tucker. 2007. Search engine advertising: Pricing ads to context. Working paper, University of Toronto, Toronto.

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

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

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

Jansen, B., A. Spink. 2007. The effect on click-through of combining sponsored and non-sponsored search engine results in a single listing. *Proc. 2007 Workshop on Sponsored Search Auctions, Banff, AB, Canada*.

Katona, Z., M. Sarvary. 2007. The race for sponsored links: A model of competition for search advertising. Working paper, INSEAD, France.

Kilpatrick, D. 2003. Keyword advertising on a cost-per-click model. http://www.sitepronews.com.

Koschat, M., W. Putsis. 2002. Audience characteristics and bundling: A hedonic analysis of magazine advertising rates. *J. Marketing Res.* **39**(2) 262–273.

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

Liu, D., J. Chen, A. Whinston. 2009. Ex ante information and design of keyword auctions. *Inform. Systems Res.* Forthcoming.

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

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

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

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

Oneupweb. 2005. How keyword length affects conversion rates. http://www.seoxy.com/conversion/How-Keyword-Length-Affects-Conversion-Rates.php.

Rimm-Kaufman, A. 2007. PPC and your good name: Sales from brand searches aren't incremental. *Search Engine Land* (March 27), http://www.searchengineland.com.

Rossi, P., G. 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., R. E. Bucklin. 2007. A model of individual keyword performance in paid search advertising. Working paper, Yale University, New Haven, CT.

Rutz, O., R. Bucklin. 2008. From generic to branded: A model of spillover dynamics in paid search advertising. Working paper, Yale School of Management, New Haven, CT.

Ruud, P. A. 2000. *An Introduction to Classical Econometric Theory*. Oxford University Press, New York.

Sherman, L., J. Deighton. 2001. Banner advertising: measuring effectiveness and optimizing placement. *J. Interactive Marketing* **15**(2) 60–64.

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

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

Tellis, G., R. Chandy, P. Thaivanich. 2001. Which ad works, when, where, and how often? Modeling the effects of direct television advertising. *J. Marketing Res.* **37**(1) 32–46.

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

Wilbur, K. C. 2008. A two-sided, empirical model of television advertising and viewing markets. *Marketing Sci.* **27**(3) 356–378.

Xu, L., J. Chen, A. Whinston. 2009. To place better or price cheaper? Bidding and pricing under keyword advertising. Working paper, University of Texas at Austin, Austin.

Yang, S., A. Ghose. 2008. Analyzing the relationship between organic and sponsored search advertising: Positive, negative or zero interdependence? Working paper, New York University, New York.

MANAGEMENT SCIENCE
DOI 10.1287/mnsc.1090.1054ec

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets" by Anindya Ghose and Sha Yang, *Management Science*, DOI 10.1287/mnsc.1090.1054.

Online Appendix: Diagonal Ω

Table A1: Coefficient Estimates on Click-through Rate

	Intercept	Retailer	Brand	Length
Intercept	$\overline{\beta}_0$ **-2.528** **(0.043)**	α_1 **1.505** **(0.117)**	α_2 **-0.178** **(0.058)**	α_3 -0.008 (0.048)
Rank	$\overline{\beta}_1$ **-0.079** **(0.008)**	γ_1 **-0.116** **(0.048)**	γ_2 -0.006 (0.014)	γ_3 **0.023** **(0.009)**
Time	α_4 **0.008** **(0.002)**			

Table A2: Coefficient Estimates on Conversion Rate

	Intercept	Retailer	Brand	Length	Landing Page Quality
Intercept	$\overline{\theta}_0$ **-5.461** **(0.098)**	δ_1 **1.623** **(0.213)**	δ_2 **-0.917** **(0.151)**	δ_3 -0.011 (0.106)	δ_4 **0.235** **(0.063)**
Rank	$\overline{\theta}_1$ **-0.114** **(0.016)**	κ_1 **0.178** **(0.080)**	κ_2 0.059 (0.040)	κ_3 0.004 (0.023)	κ_4 -0.002 (0.018)
Time	δ_5 **0.035** **(0.008)**				

Note: Posterior means and posterior standard deviations (in the parenthesis) are reported, and estimates that are significant at 95% are bolded in Tables A1 – A18.

Table A3: Coefficient Estimates on CPC

	Intercept	Retailer	Brand	Length	Landing Page Quality
	$\overline{\omega_0}$	λ_1	λ_2	λ_3	λ_4
Intercept	**-1.650**	**-0.732**	**0.165**	-0.027	**-0.038**
	(0.024)	**(0.072)**	**(0.032)**	(0.025)	**(0.016)**
	$\overline{\omega_1}$	ρ_{11}	ρ_{12}	ρ_{13}	ρ_{14}
LagRank	**-0.041**	**0.036**	-0.012	**0.019**	-0.001
	(0.004)	**(0.010)**	(0.008)	**(0.005)**	(0.004)
	λ_6				
Time	**-0.020**				
	(0.001)				

Table A4: Coefficient Estimates on Keyword Rank

	Intercept	Retailer	Brand	Length
	$\overline{\phi_0}$	τ_1	τ_2	τ_3
Intercept	**1.734**	**-0.530**	**-0.299**	**-0.236**
	(0.031)	**(0.085)**	**(0.040)**	**(0.033)**
	$\overline{\phi_1}$	π_1	π_2	π_3
CPC	**-1.881**	**0.673**	**0.322**	0.039
	(0.093)	**(0.325)**	**(0.128)**	(0.098)
	$\overline{\phi_2}$			
Lag_CTR	**-0.091**			
	(0.030)			
	τ_5			
Time	**0.025**			
	(0.001)			

Adding Competitor Price and Lag Profit

Table A5: Coefficient Estimates on Click-through Rate

	Intercept	Retailer	Brand	Length
Intercept	$\overline{\beta_0}$ **-1.692** **(0.054)**	α_1 **1.283** **(0.113)**	α_2 **-0.307** **(0.059)**	α_3 **-0.095** **(0.044)**
Rank	$\overline{\beta_1}$ **-0.256** **(0.013)**	γ_1 **-0.198** **(0.026)**	γ_2 **-0.044** **(0.014)**	γ_3 -0.007 (0.009)
Time	α_4 **0.050** **(0.003)**			

Table A6: Coefficient Estimates on Conversion Rate

	Intercept	Retailer	Brand	Length	Landing Page Quality
Intercept	$\overline{\theta_0}$ **-4.431** **(0.085)**	δ_1 **1.022** **(0.217)**	δ_2 **-0.875** **(0.138)**	δ_3 -0.024 (0.096)	δ_4 **0.191** **(0.059)**
Rank	$\overline{\theta_1}$ **-0.305** **(0.023)**	κ_1 0.042 (0.099)	κ_2 0.021 (0.044)	κ_3 -0.025 (0.021)	κ_4 0.004 (0.013)
Time	δ_5 **0.071** **(0.008)**				

Table A7: Coefficient Estimates on CPC

	Intercept	Retailer	Brand	Length	Landing Page Quality	Competitor Price
Intercept	$\overline{\omega}_0$ **-1.660** **(0.027)**	λ_1 **-0.735** **(0.073)**	λ_2 **0.090** **(0.033)**	λ_3 -0.009 (0.024)	λ_4 **-0.030** **(0.014)**	λ_5 -0.004 (0.013)
LagRank	$\overline{\omega}_1$ **-0.041** **(0.006)**	ρ_{11} **0.037** **(0.010)**	ρ_{12} -0.005 (0.009)	ρ_{13} **0.018** **(0.005)**	ρ_{14} -0.001 (0.003)	ρ_{15} 0.002 (0.002)
LagProfit	$\overline{\omega}_2$ **-0.036** **(0.010)**	ρ_{21} **0.044** **(0.021)**	ρ_{22} **0.020** **(0.009)**	ρ_{23} -0.002 (0.009)	ρ_{24} 0.005 (0.006)	ρ_{25} -0.003 (0.005)
Time	λ_6 **-0.020** **(0.001)**					

Table A8: Coefficient Estimates on Keyword Rank

	Intercept	Retailer	Brand	Length	Competitor Price
Intercept	$\overline{\phi}_0$ **1.942** **(0.031)**	τ_1 **-0.199** **(0.077)**	τ_2 **-0.275** **(0.036)**	τ_3 **-0.174** **(0.029)**	τ_4 **0.022** **(0.010)**
CPC	$\overline{\phi}_1$ **-2.008** **(0.091)**	π_1 **0.389** **(0.301)**	π_2 0.120 (0.109)	π_3 -0.018 (0.084)	π_4 **0.084** **(0.039)**
Lag_CTR	$\overline{\phi}_2$ **-1.271** **(0.048)**				
Time	τ_5 **0.031** **(0.001)**				

Table A9: Estimated Covariance across Click-through, Conversion, CPC and Rank (Ω)

	Click-through	Conversion	CPC	Rank
Click-through	**0.9241** **(0.052)**	**1.090** **(0.050)**	**-0.072** **(0.008)**	**0.462** **(0.022)**
Conversion	**1.090** **(0.050)**	**2.550** **(0.134)**	**-0.213** **(0.020)**	**-0.542** **(0.025)**
CPC	**-0.072** **(0.008)**	**-0.213** **(0.020)**	**0.217** **(0.004)**	**-0.005** **(0.002)**
Rank	**0.462** **(0.022)**	**0.542** **(0.025)**	**-0.005** **(0.002)**	**0.316** **(0.007)**

Adding Squared 'Rank' Term in CTR and Conversion Equations

Table A10: Coefficient Estimates on Click-through Rate

	Intercept	Retailer	Brand	Length
Intercept	$\overline{\beta_0}$ **-1.380** **(0.047)**	α_1 **1.149** **(0.111)**	α_2 **-0.392** **(0.059)**	α_3 **-0.152** **0.042)**
Rank	$\overline{\beta_1}$ **-0.331** **(0.008)**	γ_1 **-0.196** **(0.031)**	γ_2 **-0.046** **(0.015)**	γ_3 -0.013 0.008)
Time	α_4 **0.060** **(0.002)**			
$Rank^2/100$	α_5 **0.167** **(0.008)**			

Table A11: Coefficient Estimates on Conversion Rate

	Intercept	Retailer	Brand	Length	Quality
Intercept	$\overline{\theta_0}$ **-4.351** **(0.098)**	δ_1 **1.158** **(0.252)**	δ_2 **-1.118** **(0.103)**	δ_3 -0.129 (0.077)	δ_4 **0.179** **(0.039)**
Rank	$\overline{\theta_1}$ **-0.311** **(0.022)**	κ_1 -0.117 (0.070)	κ_2 0.042 (0.028)	κ_3 0.015 (0.016)	κ_4 -0.008 (0.010)
Time	δ_5 **0.068** **(0.006)**				
$Rank^2/100$	δ_6 **0.149** **(0.026)**				

Table A12: Coefficient Estimates on CPC

	Intercept	Retailer	Brand	Length	Quality	Comp_Price
Intercept	$\overline{\omega_0}$ **-1.673** **(0.026)**	λ_1 **-0.731** **(0.068)**	λ_2 **0.095** **(0.032)**	λ_3 -0.005 (0.025)	λ_4 **-0.034** **(0.017)**	λ_5 -0.003 (0.014)
LagRank	$\overline{\omega_1}$ **-0.040** **(0.005)**	ρ_{11} **0.036** **(0.010)**	ρ_{12} -0.003 (0.009)	ρ_{13} **0.018** **(0.005)**	ρ_{14} 0.000 (0.003)	ρ_{15} 0.002 (0.003)
LagProfit	$\overline{\omega_2}$ **-0.038** **(0.010)**	ρ_{21} **0.041** **(0.021)**	ρ_{22} **0.020** **(0.010)**	ρ_{23} -0.002 (0.009)	ρ_{24} 0.004 (0.006)	ρ_{25} -0.004 (0.006)
Time	λ_6 **-0.020** **(0.001)**					

Table A13: Coefficient Estimates on Keyword Rank

	Intercept	Retailer	Brand	Length	Comp_Price
Intercept	$\overline{\phi_0}$ **1.941** **(0.033)**	τ_1 **-0.157** **(0.063)**	τ_2 **-0.291** **(0.036)**	τ_3 **-0.183** **(0.026)**	τ_4 **0.022** **(0.010)**
CPC	$\overline{\phi_1}$ **-1.860** **(0.090)**	π_1 0.330 (0.283)	π_2 0.129 (0.099)	π_3 -0.021 (0.077)	π_4 **0.079** **(0.037)**
Lag_CTR	$\overline{\phi_2}$ **-1.436** **(0.033)**				
Time	τ_5 **0.034** **(0.001)**				

Table A14: Estimated Covariance across Click-through, Conversion, Bid Price and Rank (Ω)

	Click-through	Conversion	CPC	Rank
Click-through	**1.135**	**1.224**	**-0.101**	**0.560**
	(0.039)	**(0.062)**	**(0.008)**	**(0.016)**
Conversion	**1.224**	**2.519**	**-0.239**	**0.589**
	(0.062)	**(0.126)**	**(0.024)**	**(0.034)**
CPC	**-0.101**	**-0.239**	**0.217**	**-0.017**
	(0.008)	**(0.024)**	**(0.004)**	**(0.005)**
Rank	**0.560**	**0.589**	**-0.017**	**0.341**
	(0.016)	**(0.034)**	**(0.005)**	**(0.007)**