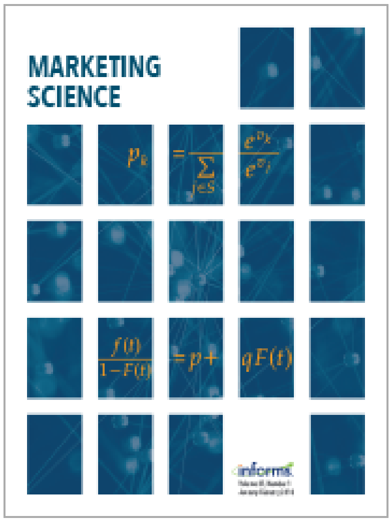


Marketing Science

Investigating the Spillover Effect of Keyword Market Entry in Sponsored Search Advertising

Shijie Lu, Sha Yang

To cite this article:
Shijie Lu, Sha Yang (2017) Investigating the Spillover Effect of Keyword Market Entry in Sponsored Search Advertising. Marketing Science 36(6):976-998. https://doi.org/10.1287/mksc.2017.1053

informs
http://pubsonline.informs.org/journal/mksc/

MARKETING SCIENCE
Vol. 36, No. 6, November–December 2017, pp. 976–998
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Investigating the Spillover Effect of Keyword Market Entry in Sponsored Search Advertising

Shijie Lu,[a] Sha Yang[b]

[a] Kenan–Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599; [b] Marshall School of Business, University of Southern California, Los Angeles, California 90089

Contact: shijie_lu@kenan-flagler.unc.edu, http://orcid.org/0000-0002-4180-6022 (SL); shayang@marshall.usc.edu, http://orcid.org/0000-0003-4190-5722 (SY)

Received: August 18, 2014
Revised: July 13, 2015; July 12, 2016; December 7, 2016
Accepted: December 19, 2016
Published Online in Articles in Advance: August 30, 2017

https://doi.org/10.1287/mksc.2017.1053

Abstract. As Internet advertising infomediaries nowadays provide rich competition information, sponsored search advertisers are becoming more strategic when selecting keywords. This paper empirically examines the spillover effects in advertisers' keyword market entry decisions, that is, how an advertiser's likelihood of using a keyword is affected by competitors' keyword entry decisions. We develop a structural model to characterize advertisers' keyword market entry decisions. We apply the model to a panel data set of 1,252 laptop-related keywords mainly used by 28 manufacturers, retailers, and comparison websites that advertise on Google. Our analysis leads to several interesting findings. First, an advertiser's expected position affects the nature of the competition. In particular, the spillover effect from below-ranked competitors is always positive, while the spillover effect from above-ranked competitors is either positive or negative. Second, the spillover effect from above-ranked ads is directionally affected by firms' product-line characteristics: the effect among firms offering homogenous products (e.g., comparison sites) is negative, whereas the effect among firms with more differentiated products (e.g., manufacturers and retailers) is positive. Third, the spillover effect from above-ranked ads is directionally affected by firms' positions in a distribution channel: the effect from upstream (downstream) on downstream (upstream) firms tends to be negative (positive). Finally, a downstream firm is more likely to learn new keywords from an upstream firm but not vice versa. Our counterfactual simulations demonstrate that the keyword-specific competition information provided by infomediaries can improve the search engine's revenue by about 5.7%.

History: K. Sudhir served as the editor-in-chief and Vishal Singh served as associate editor for this article.
Supplemental Material: Data are available at https://doi.org/10.1287/mksc.2017.1053.

Keywords: search advertising • competition • distribution channel • entry game • keyword selection • infomediary • Internet marketing • Bayesian estimation

1. Introduction

Sponsored search remains the most salient online advertising format today. According to a report from the Interactive Advertising Bureau (2015), sponsored search advertising has grown to be a $29 billion industry, accounting for about 49% of the annual digital ad spending (including both desktop and mobile) in the United States. This fast market expansion brings one noticeable difference in comparison to the past: more than just a few sponsored ads are often listed under the same search query, competing for user attention and transaction opportunity.

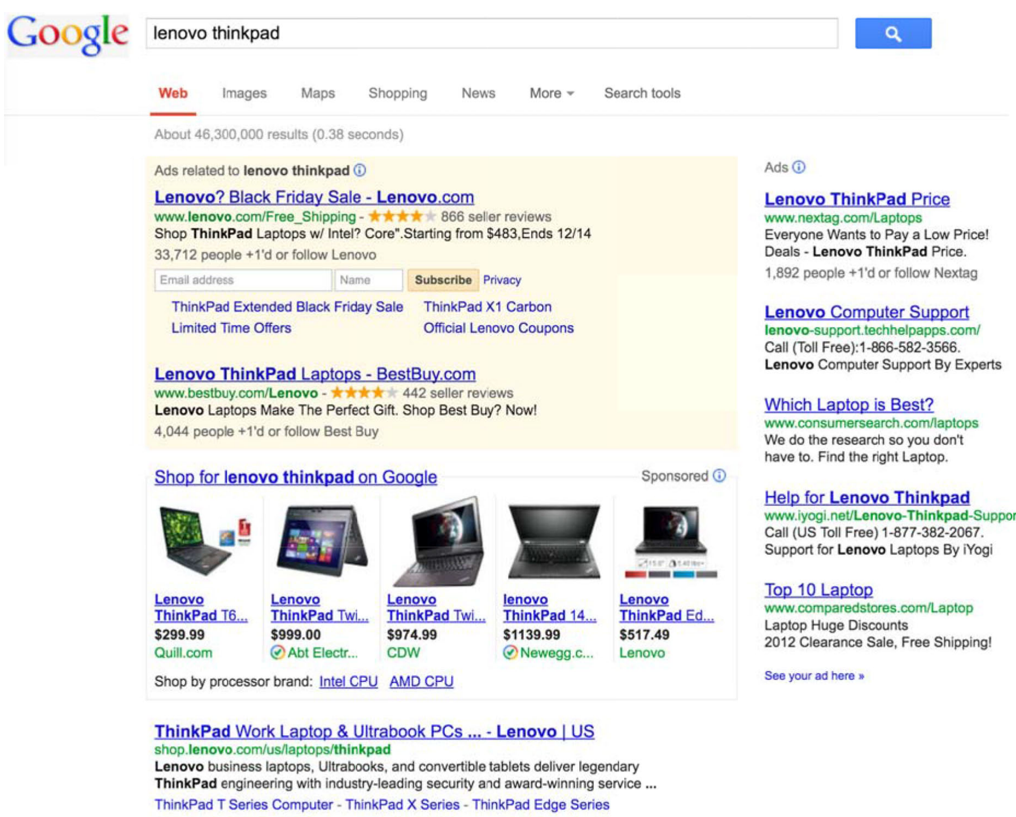
As competition escalates, the keyword choice decision turns more strategic. From a search advertiser's perspective, "entering" a particular keyword market means exposing its advertisement to users who share similar product interests and search intent, along with other entrant advertisers who also see value in this market. Intuitively, sponsored ads displayed for the same keyword, although placed in an order, affect the consumer's consideration set and the subsequent purchase, if any. Therefore, other companies that appear on the sponsored search results will affect users' click-through and conversion probabilities on a particular search ad and consequently the advertiser's expected payoff and entry probability.

The main objective of this research is to examine the spillover effect in advertisers' keyword market entry decisions, that is, how an advertiser's likelihood of using a keyword is affected by competitors' keyword entry decisions. The spillover effect in a keyword market is essentially driven by two opposing forces: the negative *business stealing effect* when firms are competing for users' limited attention and the positive *market expansion effect* when a cluster of ads in the same keyword market offers a higher expected benefit of ad viewing/clicking, thereby increasing the overall market demand. Hence, the net effect of spillovers of competitors' entry on an individual advertiser's entry can be either positive or negative depending on the relative strength of these two forces. The understanding of spillover effects in keyword markets can provide important managerial implications to both search advertisers and search engines. On one hand, it can help search advertisers more precisely predict competition intensity in a keyword market and therefore make better keyword choice decisions. On the other hand, given the linkage between competition and ad performances found in previous literature (Yang et al. 2014), quantifying spillover effects can help search engines assess the monetary value of providing competition data (e.g., firms' past keyword choices) to advertisers. In this paper, we focus on the following four research questions related to spillover effects and firms' use of competition data.

A unique feature of sponsored search ads is that they are ranked from top to bottom. This motivates us to investigate how the spillover effect of above-ranked ads differs from that from below-ranked ads as our first research question (RQ1). The position effect found in the previous literature (e.g., Ghose and Yang 2009, Narayanan and Kalyanam 2015) indicates that the click-through rate (CTR) decays with positions. Having an additional above-ranked competitor lowers the position of the focal ad and therefore decreases its CTR. This suggests that an increased number of above-ranked ads are most likely to steal the demand on the focal ad, thus lowering the focal advertiser's entry probability. However, a user's clicking propensity on the focal ad is usually affected only by above-ranked ads that this user has already viewed, because of the top-down search pattern of consumer information processing. This suggests that below-ranked ads are unlikely to affect a user's clicking propensity on the focal ad (i.e., unlikely to impose a business stealing effect). On the other hand, the market expansion effect is likely to apply to both above-ranked and below-ranked ads. Hence, we expect that the spillover effect of below-ranked ads is positive because the positive market expansion effect applies but the negative business stealing effect does not. However, the net spillover effect of above-ranked ads can be either positive or negative because both effects are likely to be present.

Another interesting observation of the keyword market is that ads from different types of firms in a distribution channel could appear in the same search listing. For example, *manufacturers* (Lenovo), *retailers* (Best Buy), and *comparison sites* (Nextag) all bought the keyword "Lenovo Thinkpad" on Google (Figure 1). Comparison sites are also called comparison engines, and

Figure 1. (Color online) Example of Sponsored Search Ads on Google



Source. Google and the Google logo are registered trademarks of Google Inc., used with permission.
Note. This snapshot was taken in November 2012.

list products of partnering retailers and charge these retailers a fee whenever directing consumers to retailers' websites. As comparison sites are distributors of retailers, they are downstream firms to retailers in the distribution channel. The spillover effect among different channel members has not been examined before since most manufacturers do not own physical stores. Nevertheless, the search advertising industry offers us a unique opportunity to understand the spillover effects between different channel members by examining the patterns of these firms' entry decisions in the virtual market of keywords.

Regarding the spillover effect among firms that are horizontal in a distribution channel (e.g., Lenovo versus Dell, or Target versus Best Buy), we suspect that the sign of the spillover effect is affected by the degree of product differentiation. Intuitively, if the products or services offered by firms are not differentiated, consumers will have less incentive to search and click on additional ads because of the lower marginal search benefit, which implies a strong business stealing effect and a weak market expansion effect. This motivates us to study whether a negative spillover effect is more likely to occur among firms offering more homogenous products because of the dominance of the business stealing effect (RQ2).

The spillover effect between vertical members in a channel can be asymmetric because upstream firms (e.g., manufacturers) directly benefit from increased demand fulfilled by downstream firms (e.g., retailers) but not necessarily the other way around. For instance, sales from retailers such as Best Buy could be cannibalized by sales from manufacturers' websites. Following this logic, we examine the potentially asymmetric spillover effects between different channel members (RQ3). We hypothesize that the keyword entry spillover from upstream firms on downstream firms is more likely to be negative compared with downstream firms on upstream firms.

Finally, we investigate whether past keyword choices of competitors were used by firms to update their keyword dictionaries (RQ4). With a growing demand for competition information, a new industry is emerging as third-party firms (known as keyword infomediaries) are starting to use screen-scraper technology to periodically track and record all sponsored ads associated with thousands of keywords on major search engines. Firms can now readily observe who they were competing with, and who was ranked above or below their own ads in previous periods, for a given keyword. Such information has been helping thousands of online advertisers evaluate their competitive landscape. The ability to observe competition information empowered by infomediaries not only drives firms to be more strategic in selecting keywords (i.e., a larger spillover effect) but also allows firms to learn new keywords from competitors' past keyword histories, especially for those with less knowledge about the product category. This is supported by anecdotal evidence that almost all major keyword infomediaries, such as Spyfu, AdGooroo, iSpoinage, and SemRush, are featuring the function of spying on competitors' secret keywords on their websites to attract and retain their clients. As downstream firms are usually less knowledgeable about product attributes than upstream firms, we suspect that downstream firms tend to update their keyword dictionaries by adding new keywords from upstream firms' past keyword histories, but not vice versa.

Our four research questions can be summarized as follows:

RQ1. How does the spillover effect vary by competitors' relative positions on ad listings?

RQ2. How does the spillover effect vary by the heterogeneity in firms' product offerings?

RQ3. What is the nature of the spillover effect between upstream and downstream firms in a distribution channel?

RQ4. How do firms update keyword dictionaries based on competitors' keyword histories?

To answer these questions, we develop a unified modeling framework to quantify the nature and magnitude of potential spillover effects in advertisers' keyword decisions. We build on the classical simultaneous-move game framework with incomplete information and extend the existing modeling setup in two important ways. First, we separately model the strategic interactions (i.e., spillover effects) with those ads ranked above those of the focal advertiser and with those ranked below to capture the ranking structure in a keyword market. Second, we characterize the set of potential entrants for a given keyword by modeling advertisers' keyword consideration process. This extension is crucial because forming a keyword consideration set or keyword dictionary (hereafter these two terms are used interchangeably) is a common industry practice because of advertisers' limited knowledge about all available keywords. Furthermore, the magnitude of estimates on spillover effects will suffer a downward bias without accounting for firms' keyword consideration process. We also model several postentry outcomes such as average click volume, average cost per click (CPC), and the ranking of ads for each keyword market. We link these variables to the keyword entry model by assuming that firms form rational expectations on postentry outcomes when deciding whether to purchase a keyword.

We apply the model to a novel data set obtained from a leading keyword infomediary in the United States, which contains advertisers' entry information and corresponding ad ranking on 1,252 popular keywords related to laptops and accessories on Google from September 2011 to April 2012. Our analysis focuses on 28 major advertisers including PC producers, retailers, and comparison websites. To overcome the multiple-equilibrium problem driven by firms' strategic interactions when making entry decisions and the high dimensionality of parameter space due to unobserved keyword consideration sets, we use a two-step estimation approach in conjunction with the Bayesian method to make model inferences.

Several key findings emerge from our analysis: (i) the spillover effect from below-ranked competitors is always positive, while the spillover effect from above-ranked competitors can be either positive or negative; (ii) the spillover among firms offering homogenous products (i.e., comparison sites) is negative, while the spillover among branded firms with more differentiated products (i.e., manufacturers and retailers) is positive; (iii) the spillover from upstream firms on downstream firms (i.e., manufacturers on comparison sites) tends to be negative, and the spillover from downstream firms on upstream firms (i.e., comparison sites on retailers) tends to be positive; (iv) a downstream firm is more likely to learn new keywords from an upstream firm than vice versa.

To help search engines assess the value of providing competition data to advertisers, we conduct counterfactual simulations under the assumption of (i) having no access to competition data, (ii) having access to aggregate competition data at the firm level, or (iii) having access to disaggregate competition data at both the firm and keyword levels. We find that the search engine can significantly benefit from the competition information provided by third-party keyword infomediaries, especially when the information is at the disaggregate level. Our results show that the search engine's revenue can be improved by 5.7% compared to the case without competition information.

The rest of this paper proceeds as follows. We begin with a literature review in Section 2, outlining some major prior work related to sponsored search advertising and competitive entry. We then describe the empirical context and data in Section 3, and the proposed model in Section 4 to capture the spillover effects in advertisers' keyword market entry. In Sections 5 and 6, we first discuss the estimation methods and empirical identification, and then present our empirical findings and managerial implications based on counterfactual simulations. We conclude with a summary of this study in Section 7.

2. Related Literature

This paper builds on the burgeoning literature on search advertising. As most search engines allocate and sell ad spaces via auctions, the theoretical literature of search advertising has mainly focused on understanding advertisers' bidding reactions to different sales mechanisms (e.g., Athey and Nekipelov 2012, Edelman et al. 2007, Varian 2007). By contrast, previous empirical work has focused on measuring the effectiveness of advertising performances. This line of research is exemplified by studies of position effects on various user response metrics (Agarwal et al. 2011, Ghose and Yang 2009, Narayanan and Kalyanam 2015), on the interplay between generic and branded keywords (Rutz and Bucklin 2011), on the interaction between search and organic results (Yang and Ghose 2010), on the substitution between online and offline advertising (Goldfarb and Tucker 2011, Joo et al. 2014), and on the role of keyword popularity on consumer clicking behavior (Jerath et al. 2014). In addition to measuring search advertising effectiveness, several papers have developed novel empirical methods to infer advertisers' bidding valuations in different auction mechanisms (e.g., Yang et al. 2014, Yao and Mela 2011).

A relatively underresearched yet highly important issue is advertisers' keyword decisions. Only a handful of papers have examined this phenomenon. For example, researchers in computer science have studied keyword choices by developing more efficient recommendation algorithms (e.g., Abhishek and Hosanagar 2007, Chen et al. 2008, Fuxman et al. 2008). However, the primary focus of our paper is on the strategic interactions among advertisers' keyword decisions, which has been largely ignored in the computer science literature.

A few marketing papers have examined this phenomenon. Desai et al. (2014) built a game-theoretical model to explain advertisers' purchasing of competitors' branded keywords, while taking into account the impact of this tactic on the firms' price competition on the advertised product. One interesting finding from this study is that advertisers may purchase their own branded keywords to offset the potential negative effect if competitors bought them instead. In another study, Sayedi et al. (2014) also examined the "piggybacking" practice in sponsored search advertising but from the perspective of budget allocation between online and offline advertising. Our study differs from and extends this strand of literature in several ways. First, our study is empirical and examines advertisers' choices on both branded and generic keywords. Second, we infer the strategic interaction via the observed keyword choice decisions and provide additional insights on keyword competition among different types of advertisers including manufacturers, retailers, and comparison sites. Finally, we adopt a structural modeling framework by assuming that advertisers' keyword entry decisions are affected by their expectations on the postentry outcomes, and this aspect has not been incorporated in previous studies.

Another relevant paper is Yang et al. (2014), which examined the impact of the number of competing ads on click volume and CPC, and the formation of a competition set. Our work differs from theirs in several major ways. First, unlike Yang et al. (2014), which treated advertisers as symmetric firms, our model includes different types of advertisers. This general framework allows us to flexibly capture the strategic interactions among advertisers. Second, the asymmetric assumption also enables us to study the effect of ads placed above and below, which has important implications for advertisers. Finally, we extend the simultaneous-move entry literature by modeling the set of potential entrants for each keyword. In this way, we can more accurately quantify the strategic interaction effects in advertisers' keyword decisions.

From a methodological perspective, our paper is related to the following two bodies of literature. First, we model advertisers' keyword choices as an incomplete-information simultaneous-move game. This framework has been widely adopted in the marketing and economics literature on discrete entry games (e.g., Datta and Sudhir 2013, Seim 2006, Vitorino 2012, Zhu and Singh 2009). Previous studies typically inferred the potential entrants based on the characteristics of the players and the market in a deterministic way. However, since many of these characteristics (such as distance from a firm to a market) are not applicable in the context of sponsored search advertising, it is rather difficult to identify the set of potential entrants a priori. Our paper extends the discrete game literature by modeling the set of potential entrants of a keyword as latently determined by a keyword-consideration process of advertisers. In a simulation study in Section A.4 in the appendix, we document that an incorrect specification of the set of potential entrants significantly contaminates the estimates on strategic interaction terms and causes a downward bias.

Second, we model advertisers' keyword consideration sets to be consistent with the industry's practices, and we capture advertisers' limited capacity in analyzing all existing keywords. There is a rich body of marketing literature on modeling consideration sets as exemplified by Andrews and Srinivasan (1995), Bronnenberg and Vanhonacker (1996), Chiang et al. (1999), Gilbride and Allenby (2004), and van Nierop et al. (2010). Two methods have been developed to model the consideration of J options. The first method is to model the probability distribution of all $2^J - 1$ possible consideration sets (Chiang et al. 1999). However, this approach is not feasible in our context because the number of keyword options is large. The second method breaks down the curse of dimensionality by modeling the marginal distribution of each option being considered (Bronnenberg and Vanhonacker 1996, van Nierop et al. 2010). As shown in van Nierop et al. (2010), the consideration sets retrieved by the second method accurately correspond to the actual consideration sets. We adopt the second approach to model advertisers' keyword considerations. Our study contributes to the consideration modeling literature by extending the concept of consideration sets to firms' decision making.

3. Empirical Context

We obtained data from a leading search-advertising keyword infomediary in the United States. This company uses screen-scraper technology to track all search ads associated with thousands of keywords on Google AdWords on a monthly basis. By analyzing the domain information related to each ad, the infomediary identifies the advertisers who purchased the corresponding keyword (i.e., entered the keyword market) and the corresponding positions of their ads. Our data include the identities and ranks of search ads on Google in the U.S. market that are associated with 1,252 keywords related to laptop products and accessories from September 2011 to April 2012. These keywords are popular ones, and each keyword was used by at least one advertiser during the data period. We regard each keyword as a market because a search keyword often reflects a certain type of user interest and purchase intention. Each advertiser who bought a keyword is called an *entrant* for that keyword market.

We focus on 28 major search advertisers, who account for 72% of the total exposures of these 1,252 keywords during the data period. The 28 search advertisers are classified into three types: manufacturers (M), retailers (R), and comparison sites (C). For each type, we list the 28 focal advertisers based on their entry frequencies from high to low. Our data provider confirmed with us that these 28 firms all have access to competition information provided by infomediaries. The firms are as follows:

- 8 manufacturers—Toshiba, Dell, HP, Sony, Apple, Acer, Samsung, Lenovo;
- 12 retailers—eBay, Microsoft Store, Amazon, Newegg, Best Buy, Walmart, Staples, OfficeMax, TigerDirect, Target, Sears, Office Depot;
- 8 comparison sites—Buycheapr, Bizrate, Nextag, Shopzilla, Pronto, Beso, PriceGrabber, Smarter.

We construct several time-invariant keyword characteristics: *Length* (number of words included in a keyword), *Specific* (number of words referring to the brand/model/serial number of a product), and *Promotional* (number of promotional terms). For example, a keyword of "Dell Latitude sale" includes two specific terms and one promotional term. In addition to these keyword-specific attributes, we also construct a keyword- and advertiser-specific dummy variable, *Match*, to capture the relevance of a keyword to a manufacturer. We define *Match* to be one if a keyword contains the brand name of the manufacturer or the keyword includes specific terms that are exclusively associated with this manufacturer. Table 1 presents the

Table 1. Summary Statistics of Time-Invariant Keyword Attributes

Variable	Mean	Std. dev.	Min.	Max.
Keyword specific				
Length	2.802	0.686	1	5
Specific	1.132	1.005	0	3
Promotional	0.138	0.350	0	2
Keyword and advertiser specific				
Match_Toshiba	0.200	0.400	0	1
Match_Dell	0.109	0.312	0	1
Match_HP	0.081	0.274	0	1
Match_Sony	0.056	0.230	0	1
Match_Apple	0.097	0.295	0	1
Match_Acer	0.069	0.253	0	1
Match_Samsung	0.019	0.137	0	1
Match_Lenovo	0.090	0.287	0	1

summary statistics of the keyword attributes. For example, 20% of the keywords are highly relevant to Toshiba, and only 2% are related to Samsung's products.

In addition to data on advertisers' keyword choices and ranks, we also have monthly and keyword-specific aggregate data on search volume, click volume, and CPC. However, for each keyword, we do not have click volume or CPC information for individual advertisers. We supplement our data of advertisers' keyword choices in the U.S. market with data on two additional sets of variables correlated with firms' keyword entry decisions: (1) advertisers' ranks on Google's organic search listings in the U.S. market for the same 1,252 keywords and (2) whether advertisers have bought the same keyword and consequently appeared on Google's sponsored search listings in the UK market. For each keyword in each month, we observe whether a firm was ranked in the top 50 on organic listings and its organic rank. Similarly, we observe whether a firm purchased this keyword on Google UK and the rank of its ad. Not all firms in our sample have a presence in the UK market: seven retailers (Newegg, Office Depot, OfficeMax, Sears, Target, TigerDirect, and Walmart) and two comparison sites (Beso and Smarter) did not purchase keywords from Google UK during our data period. We report the summary statistics of all time-variant variables in Table 2, where rank-related variables are summarized across firms that have appeared

Table 2. Summary Statistics of Time-Variant Keyword Attributes

Variable	Mean	Std. dev.	Min.	Max.
Sponsored search in the U.S.				
N	5.898	3.735	1	17
N(M)	1.762	1.781	0	8
N(R)	2.491	2.052	0	10
N(C)	1.646	1.455	0	7
Rank	6.420	5.123	1	22
CPC ($)	1.191	1.171	0.080	22.48
Daily Search Volume (1K)	2.282	20.130	0.002	370
Daily Clicks per Ad	16.160	109.057	0.050	2,716
Organic search in the U.S.				
Organic_dummy	0.169	0.375	0	1
Rank_organic	13.365	9.031	1	50
Sponsored search in the UK				
UK_entry	0.106	0.308	0	1
Rank_UK	4.927	4.437	1	22

on the listings of respective markets. The average number of ads per keyword is about six. The average CPC is $1.19, and, on average, each ad receives 16 clicks per day.

We report the correlations between variables related to the number of different types of firms on sponsored search listings in the United States, on organic listings in the United States, and on sponsored search listings in the United Kingdom in Table 3, where correlations significant at the 99% level are bolded. Note that the number of retailers who advertise in a keyword is significantly and positively correlated with the number of manufacturers and the number of comparison sites, suggesting the potentially positive spillover effects between manufacturers and retailers, and between retailers and comparison sites. Furthermore, for each type of firm, both the number of top-50 organically ranked ads (N^O) and the number of ads in the UK market (N^{UK}) are positively correlated with the number of ads in the U.S. market for the same keyword.

4. Model

4.1. Model Setup

We consider I advertisers and K keywords, indexed by i ($i = 1, \ldots, I$) and k ($k = 1, \ldots, K$), respectively. Time

Table 3. Correlation Matrix of the Number of Advertisers

	$N(M)$	$N(R)$	$N(C)$	$N^O(M)$	$N^O(R)$	$N^O(C)$	$N^{\mathrm{UK}}(M)$	$N^{\mathrm{UK}}(R)$	$N^{\mathrm{UK}}(C)$
$N(M)$	1								
$N(R)$	**0.31**	1							
$N(C)$	0.01	**0.36**	1						
$N^O(M)$	**0.69**	**0.24**	**−0.09**	1					
$N^O(R)$	**0.29**	**0.83**	**0.30**	**0.25**	1				
$N^O(C)$	**0.13**	**0.37**	**0.40**	**0.08**	**0.35**	1			
$N^{\mathrm{UK}}(M)$	**0.25**	**0.11**	**−0.15**	**0.21**	0.07	−0.04	1		
$N^{\mathrm{UK}}(R)$	**0.18**	**0.32**	**0.20**	**0.13**	**0.31**	**0.16**	**0.09**	1	
$N^{\mathrm{UK}}(C)$	**−0.18**	0.06	**0.39**	**−0.20**	**0.04**	**0.10**	**−0.16**	**0.23**	1

Note. Correlations significant at the 99% level are bolded.

is discrete and indexed by t ($t = 0, \ldots, S$), where $t = 0$ refers to the initial period of the data. We let I_{kt} stand for the set of potential entrants of keyword k at time t. Next we present a three-stage process to describe advertisers' keyword choices and several postentry observations including the rank of ads, average CPC, and average click volume for a keyword.

In the first stage, each firm forms a consideration set at time t, defined as the set of keywords it might purchase. This is consistent with a common practice in the search advertising industry where big advertisers typically form a keyword dictionary for each product category and then choose keywords to buy from this dictionary (Stokes 2008). This keyword dictionary contains all terms that a firm regards as relevant to its products and is updated periodically.

In the second stage, all advertisers who have considered a keyword k at time t form the set of potential entrants I_{kt} and decide whether to purchase the keyword or enter that market. We model the entry of advertisers as a simultaneous-move game with incomplete information. Building on the incomplete-information framework proposed by Seim (2006), we assume that each potential entrant $i \in I_{kt}$ receives a privately known profit shock when making the entry decision.

After the set of entrants, denoted by E_{kt}, is determined, postentry outcomes are realized. The position of an ad is determined via an auction mechanism, in which the search engine ranks all ads related to a keyword based on advertisers' weighted bids. Because we do not observe the bidding data for individual advertisers, we model ad positions as an ordering outcome of the vector of latent weighted bids for all entrant advertisers. Since one of our objectives is to assess the value of competition information to search engines, we also model the average CPC and the average click volume for a keyword.

4.2. Modeling Ad Positions

Conditional on keyword entry, advertisers need to submit bids in an auction. The search engine then determines the rank of the ads based on these bids weighted by a quality score. For each entrant advertiser, $i \in E_{kt}$, let R_{kit} stand for advertiser i's unobserved weighted bid. Then the rank of advertiser i's ad $Rank_{kit}$ is a discrete value determined by the vector of $\{R_{ki't}\}_{i' \in E_{kt}}$ in descending order, that is, $Rank_{kit} < Rank_{kjt}$ iff $R_{kit} > R_{kjt}$. Here a smaller *Rank* refers to a position further toward the top. We model the latent weighted bid for each entrant advertiser as follows:

$$\begin{aligned} R_{kit} = {} & \alpha_i + \alpha_{1T} X_k + \alpha_{2T} Match_{ki} + \alpha_{3T} \ln(SV_{kt}) \\ & + \alpha_{4T} \ln(Click_{k,t-1}) + \alpha_{5T} \ln(CPC_{k,t-1}) \\ & + \alpha_{6T} Z_{kit} + \ln \sum_{l=0}^{t-1} \alpha_{7lT} W_{kil} + \xi_t^R + \varsigma_{kit}^R, \end{aligned} \tag{1}$$

where the weighted bid is assumed to depend on three groups of covariates: (i) keyword attributes including static keyword characteristics X_k, whether a keyword matches a particular manufacturer, search volume denoted by SV_{kt}, and lagged clicks and CPC; (ii) Z_{kit}, standing for the presence and the rank on organic listings in the United States and on sponsored search listings in the United Kingdom; and (iii) the firm's previous entry and ad rank (i.e., $W_{kit} = \{a_{kit}, \ln(Rank_{kit} + 1)\}$). We account for seasonal effects by including monthly dummies, ξ_t^R. The parameter α_i captures an advertiser-specific fixed effect, and the coefficients of other variables are all type specific. The error term ς_{kit}^R is assumed to follow a standard normal distribution. We normalize the variance of ς_{kit}^R to be one for identification purposes. Furthermore, since all covariates of keyword attributes do not vary across firms, we also normalize the associated coefficients to be zero for manufacturers for identification purposes. Based on the specifications described above, the probability for advertiser i's ad to rank above j's can be expressed as follows:

$$\begin{aligned} P(Rank_{kit} < Rank_{kjt}) &= P(R_{kit} > R_{kjt}) \\ &= \Phi\left(\frac{E(R_{kit}) - E(R_{kjt})}{\sqrt{2}}\right). \end{aligned} \tag{2}$$

4.3. Modeling Average Click Volume and Cost per Click

Since one of our research objectives is to assess the monetary value of competition information to search engines, we model the average click volume and the average CPC for a keyword. We follow the previous literature and model the evolution of these two outcomes as a first-order autoregressive (AR1) process (Rutz and Bucklin 2011). In addition to covariates of keyword attributes, we also include the number of firms from different types on both the organic listings and the sponsored ad listings. The average click volume and the average CPC of a keyword are specified as follows:

$$\begin{aligned} & \ln(Click_{kt}) \\ & = \beta_1^Q X_k + \beta_2^Q Match\_entry_{kt} + \beta_3^Q \ln(SV_{kt}) + \beta_4^Q \ln(Click_{k,t-1}) \\ & \quad + \sum_{T=\{M,R,C\}} \left[\beta_{5T}^Q \ln(N_{kt}(T)+1) + \beta_{6T}^Q \ln(N_{kt}^O(T)+1)\right] \\ & \quad + \xi_t^Q + \eta_k^Q + \upsilon_{kt}^Q, \end{aligned} \tag{3}$$

$$\begin{aligned} & \ln(CPC_{kt}) \\ & = \beta_1^{CPC} X_k + \beta_2^{CPC} Match\_entry_{kt} + \beta_3^{CPC} \ln(SV_{kt}) \\ & \quad + \beta_4^{CPC} \ln(CPC_{k,t-1}) \\ & \quad + \sum_{T=\{M,R,C\}} \left[\beta_{5T}^{CPC} \ln(N_{kt}(T)+1) + \beta_{6T}^{CPC} \ln(N_{kt}^O(T)+1)\right] \\ & \quad + \xi_t^{CPC} + \eta_k^{CPC} + \upsilon_{kt}^{CPC}, \end{aligned} \tag{4}$$

$$(\eta_k^Q, \eta_k^{CPC}) \sim \mathrm{MVN}(0, \Omega), \tag{5}$$

where $Match\_entry_{kt}$ is a dummy variable that equals one if this keyword is related to a product from any entrant manufacturers (e.g., "MacBook" bought by Apple); $N_{kt}(T)$ and $N^O_{kt}(T)$ are the number of type T advertisers who bought a keyword and the number of top-50 type T firms on the organic listing, respectively; υ^Q_{kt} and $\upsilon^{\mathrm{CPC}}_{kt}$ are measurement errors; ξ^Q_t and ξ^{CPC}_t control for monthly fixed effects; and the parameters η^Q_k and η^{CPC}_k capture the unobserved keyword heterogeneity in average click volume and CPC, which are assumed to be normally distributed and correlated.

The number of entrants $N(T)$ can be endogenous in models of average click volume and CPC because the keyword entry decisions may result from advertisers' expectations about these two metrics. We control for the potential endogeneity of $N(T)$ in Equations (3) and (4) by using the number of entrants in the UK market ($N^{\mathrm{UK}}(T)$) as instrument variables (IVs) and allowing the error terms in average click volume, CPC, and log of the number of manufacturers, retailers, and comparison sites to be correlated

$$\ln(N_{kt}(T)+1) = \beta_1^T X_k + \beta_2^T \ln(SV_{kt}) + \beta_3^T \ln(N_{k,t-1}(T)+1) + \beta_4^T \ln(N^{\mathrm{UK}}_{kt}(T)+1) + \xi_t^T + \upsilon_{kt}^T, \quad \text{where } T = M, R, C, \tag{6}$$

$$(\upsilon^Q_{kt}, \upsilon^{\mathrm{CPC}}_{kt}, \upsilon^M_{kt}, \upsilon^R_{kt}, \upsilon^C_{kt}) \sim \mathrm{MVN}(0, \Phi). \tag{7}$$

We believe that the number of entrants in the UK keyword market is a valid IV for that in the U.S. keyword market. The idea of finding IVs from outside markets has been commonly used in the economics literature (e.g., Hausman 1996, Nevo 2001). A similar approach of using advertising decisions in other countries as IVs for advertising decisions in the focal country was recently adopted by Lamey et al. (2012), Van Heerde et al. (2013), and Dover and Neslin (2015). In sponsored search advertising, advertisers typically incur the cost of designing effective ad copies (i.e., ad title and ad content) to cater to seasonal changes in demand at the time of a keyword market entry. For example, a firm might include words related to Thanksgiving in the ad copy in November and include words related to Christmas in the ad copy for the same keyword in December. This cost of ad copy design is also reflected by advertisers' constant effort in testing ad copy variations, commonly through A/B testing.

The validity of our proposed instrument N^{UK} hinges on two rationales. First, a firm's keyword choice in the UK market is often negatively associated with the firm's cost on ad copy design for the same keyword in the U.S. market during the same time period. This is true because a firm's ad copy used in the UK market can be readily served as a benchmark for conducting A/B tests in the U.S. market, and thus reduces the cost of ad design and experimentation. Second, the demand shocks associated with the same keyword are often uncorrelated across countries. This is true because firms typically recruit local ad agencies to manage advertising campaigns in different countries. As a result, users who search a keyword in the United Kingdom are unlikely to receive the same advertising message as those who search the same keyword in the United States. The discussion above suggests that a firm's use of a keyword in the UK market can serve as a valid instrument for the same firm's keyword decision in the U.S. market because (i) it is likely to affect the entry decision in the U.S. market for the same keyword given the correlation on cost in the two markets, and (ii) it is unlikely to affect the demand (such as clicks) for the same keyword in the U.S. market.

4.4. Modeling Keyword Entry

We model the entry decisions made by potential entrants of a keyword as a simultaneous-move game with incomplete information. We define the set of potential entrants as those advertisers who have considered this keyword at the stage of entry. Previous research on market entry typically defines the set of potential entrants deterministically, because it either focuses on an oligopoly or even a duopoly industry with only a few players (e.g., Berry 1992, Ciliberto and Tamer 2009, Jia 2008, Vitorino 2012, Zhu and Singh 2009) or is able to pin down the set of potential entrants using a geographical proximity assumption (e.g., Datta and Sudhir 2013, Seim 2006). However, in the context of sponsored advertising, there are more players and no well-defined rules that we can use to determine potential entrants for a keyword. Thus, we treat the set of potential entrants as latent variables and model the keyword consideration process.

Following Seim's (2006) incomplete information framework, we assume that each potential entrant $i \in I_{kt}$ forms rational expectations on others' entry probabilities, including those who do not consider entry and hence have zero entry probability. In other words, we assume that advertisers can rationally anticipate the set of potential entrants for each keyword, which is unobserved by researchers. This assumption can be justified in two ways. First, since all firms in our empirical context are large-sized companies, they have the financial resources to invest in research to help them identify their main competitors for each keyword market. Second, Google constantly encourages advertisers to run short-term experiments before making a change in an ad campaign. This experimentation process allows firms to actively monitor other sponsored links and hence infer the set of potential entrants.

Let a_{kit} be a binary indicator of advertiser i's use of keyword k at period t. Then the set of entrants is defined as $E_{kt} = \{i \mid i \in I_{kt} \text{ and } a_{kit} = 1\}$. We derive a firm's payoff of purchasing a keyword denoted by U_{kit} from a specification wherein U_{kit} depends on three sets of covariates: (1) expected outcomes including expected rank, expected average click volume, and CPC; (2) competition measured by the expected number of competitors, $E(N_{kit}) = \sum_{-i} \Pr(a_{k,-i,t})$; (3) other covariates denoted by X^U_{kit}, which include keyword attributes and the firm's presence on the associated organic listings and on Google UK

$$U_{kit} = \gamma_1 E(Click_{kt}) + \gamma_2 E(CPC_{kt}) + \gamma_3 E(Rank_{kit}) + \gamma_4 E(N_{kit}) + \gamma_5 X^U_{kit} + \varsigma^U_{kit}, \tag{8}$$

where ς^U_{kit} is an independent and identically distributed (i.i.d.) idiosyncratic profit shock observed by advertiser i but known to rivals up to its distribution. Note that the expected rank of a firm exactly equals the expected number of competitors ranked above (i.e., $E(Rank_{kit}) = E(N^a_{kit})$), and the total number of entrants equals the sum of the numbers of entrants ranked above and below (i.e., $N_{kit} = N^a_{kit} + N^b_{kit}$). Hence, we can rewrite Equation (8) with a parameterization of $\gamma'_3 = \gamma_3 + \gamma_4$ as follows:

$$U_{kit} = \gamma_1 E(Click_{kt}) + \gamma_2 E(CPC_{kt}) + \gamma'_3 E(N^a_{kit}) + \gamma_4 E(N^b_{kit}) + \gamma_5 X^U_{kit} + \varsigma^U_{kit}, \tag{9}$$

where the $E(N^a_{kit})$ and $E(N^b_{kit})$ are expressed in Equations (11) and (12) based on our model of ad ranking.

Note that the entry payoff in Equation (9) allows the strategic interaction or spillover effects from competitors' entry on an advertiser's payoff to vary with the expected order of ad positions. Furthermore, because of consumers' top-down search habit, firms are most likely to consider the type of advertisers who will rank above them rather than just the total number of competing advertisers when making an entry decision. This motivates us to assume the entry of different types of above-ranked competitors to have asymmetric effects on a firm's entry payoff, which leads to the following equations:

$$U_{kit} = \gamma_1 E(Click_{kt}) + \gamma_2 E(CPC_{kt}) + \sum_T \gamma_{3T} E(N^a_{kit}(T)) + \gamma_4 E(N^b_{kit}) + \gamma_5 X^U_{kit} + \varsigma^U_{kit}, \tag{10}$$

$$E(N^a_{kit}(T)) = \sum_{j \neq i,\, j \in I_{kt}} \left[P(a_{kjt}=1) P(Rank_{kjt} < Rank_{kit}) 1\{j \in T\}\right], \tag{11}$$

$$E(N^b_{kit}) = \sum_{j \neq i,\, j \in I_{kt}} [P(a_{kjt}=1) P(Rank_{kjt} \geq Rank_{kit})], \tag{12}$$

where $E(N^a_{kit}(T))$ represents the expected number of type T advertisers who rank above i.

We allow the coefficient in entry payoff to vary by firm type to capture the type-specific heterogeneity in these effects. We also use the log-transformed number of entrants justified by the improved model fitness found in a reduced-form analysis.[1] To capture the state-dependence effect on entry probability, we construct two covariates, $Elag_{kit}$ and $RlagInv_{kit}$, defined in Equations (14) and (15). Here, $Elag_{kit}$ and $RlagInv_{kit}$ stand for the exponentially weighted average of advertiser i's past entry and inverse of rank, respectively. The parameter $\rho \in (0,1)$ is the decay parameter capturing the diminishing effect of lagged variables on entry payoff. Finally, we capture the firm-specific fixed effect by γ_i and seasonal effect by ξ^U_t. This leads to the following equations of entry payoff:

$$\begin{aligned} U_{kit} = {} & \gamma_i + \gamma_{1T} E(Click_{kt}) + \gamma_{2T} E(CPC_{kt}) \\ & + \sum_{T'} \gamma_{3TT'} \ln(E(N^a_{kit}(T')) + 1) \\ & + \gamma_{4T} \ln(E(N^b_{kit}) + 1) \\ & + \gamma_{5T} X^U_{kit} + \gamma_{6T} Elag_{kit} + \gamma_{7T} RlagInv_{kit} \\ & + \xi^U_t + \varsigma^U_{kit}, \end{aligned} \tag{13}$$

$$Elag_{kit} = \rho\, Elag_{ki,t-1} + a_{kit}, \tag{14}$$

$$RlagInv_{kit} = \rho\, RlagInv_{ki,t-1} + RankInv_{kit}, \tag{15}$$

$$RankInv_{kit} = \begin{cases} Rank^{-1}_{kit}, & \text{if } a_{kit} = 1, \\ 0, & \text{if } a_{kit} = 0, \end{cases} \tag{16}$$

$$(\varsigma^R_{kit}, \varsigma^U_{kit}) \sim \text{MVN}\left(0, \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix}\right), \tag{17}$$

where a larger *RankInv* is associated with a higher ad position.

Note that both the expected average click volume and CPC are functions of the number of different types of entrants indicated by Equations (3) and (4). Hence, our specification of entry payoff enables us to assess the direct and the indirect impact of competition on entry payoff. The direct effect is measured by the coefficients of the number of competitors in the entry equation, while the indirect effect is measured by the impact of competition on expected click volume and CPC.

We allow unobserved error terms in the entry and ranking equations to be correlated in Equation (17) to control for the potential sample selection in the model of ranking. The unobserved component in a firm's weighted bid (ς^R_{kit}) could be potentially correlated with the firm's entry payoff for at least two reasons. On one hand, if ς^R_{kit} consists mainly of the quality score of the ad, we expect a positive δ because the quality score is positively associated with keyword relevance, which is further positively correlated with the firm's likelihood of entering a keyword market. On the other hand, if ς^R_{kit} consists mainly of a firm's bid, then a higher entry payoff suggests a higher quality score, which can lower the firm's incentive to raise their bid for a higher position and therefore lead to a negative δ.

To complete the entry model, we normalize the outside profit to be zero so that an advertiser enters a keyword market iff $U_{kit} > 0$. Under these assumptions, we can characterize the Bayes–Nash equilibrium of the entry probability for potential entrants as follows:

$$P^*(a_{kit}=1)=\Phi[E(U(P^*(a_{-i}),X;\gamma))], \quad \forall i \in I_{kt}. \quad (18)$$

4.5. Modeling Keyword Consideration

We characterize the set of potential entrants I_{kt} by modeling advertisers' keyword consideration sets. By definition, an advertiser i is a potential entrant of keyword k if and only if keyword k belongs to its consideration set. We denote $c_{kit}=1$ if advertiser i considers keyword k at time t. We model a firm's consideration decision in a similar way to the consideration model proposed in Bronnenberg and Vanhonacker (1996) and as further developed in van Nierop et al. (2010). We model an advertiser's consideration decision by describing the latent consideration intensity V_{kit}, which indicates a consideration when V_{kit} exceeds a threshold that is normalized to be zero.

We assume advertisers to be nonstrategic at the consideration stage for several reasons. First, unlike the keyword entry game where a firm's advertising payoff is affected by others' entry decisions through position competition, there is no clear economic rationale for firms' strategic interaction at the consideration stage. Second, a firm's consideration of a keyword sometimes merely reflects a firm's awareness of a keyword rather than an active decision made from a strategic perspective. Finally, since we do not observe advertisers' keyword considerations, we are unable to empirically identify such strategic interactions, if any, at the consideration stage. We also assume that all advertisers have access to the competition information provided by the infomediary, as our data provider has confirmed that all 28 advertisers have subscribed to its service.

According to Stokes (2008), advertisers usually establish and update their keyword consideration set or keyword dictionary based on three sources of information: (i) the advertiser's own keyword history, (ii) keyword-recommendation tools relying on proximity-based algorithms (e.g., the free Keyword Tool provided by Google AdWords), and (iii) competitors' keyword histories as observed by infomediaries. To be consistent with these industrial routines, we allow an advertiser's keyword consideration set to be influenced by three sets of variables: *Keyword loyalty*, *Keyword similarity*, and *Keyword popularity*, the definitions of which are elaborated below. We model an advertiser's consideration intensity, V_{kit}, of a keyword as follows:

$$V_{kit}=\lambda_i+\lambda_{1T}\,Elag_{kit}+\lambda_{2T}\,KS_{kit} + \sum_{T'}\lambda_{3TT'}\ln[Nlag_{kit}(T')+1]+\varsigma^V_{kit}, \quad (19)$$

$$Sim(X_k,X_S)=\frac{X_k\cdot X_s}{\|X_k\|\|X_s\|}, \quad (20)$$

$$KS_{kit}=\frac{\sum_s Sim(X_k,X_s)\cdot RankInv_{si,t-1}}{\sum_s RankInv_{si,t-1}}, \quad (21)$$

$$Nlag_{kit}(T)=\rho Nlag_{ki,t-1}(T)+N_{kit}(T), \quad (22)$$

where λ_i accounts for the advertiser's fixed effect, and ς^V_{kit} is an unknown error term that follows i.i.d. standard normal distribution. Since we assume firms to be nonstrategic at the consideration stage, ς^V_{kit} is a measurement error that does enter into a firm's decision process. Furthermore, since firms' keyword consideration is unobserved, we assume there is no correlation between unobserved error terms in consideration and the counterparts in entry and ranking for the purpose of identification.

We use the advertiser's weighted average monthly entry frequency, $Elag_{kit}$, to measure its keyword loyalty. We create a variable, KS_{kit}, to measure the similarity between a keyword k and the set of keywords purchased by advertiser i in the last period based on the cosine similarity, which has been commonly used to measure the semantic similarity in various contexts such as sponsored search advertising and user-generated content in both computer science (e.g., Abhishek and Hosanagar 2007, Radlinski et al. 2008, Kakulapati et al. 2011) and marketing (e.g., Lee and Bradlow 2011). We first define the similarity between two keywords in Equation (20), where X_k is a vector of standardized keyword attributes, and $\|\cdot\|$ is the L^2 measure. We then define KS as the average of the similarity of a keyword to all keywords used by the firm in the last period. We further weight the similarity by the *RankInv* of each used keyword, which serves as the proxy of the relevance of a keyword to advertised content. We focus on this weighted KS because most search engines that provide keyword tools could easily incorporate relevance metrics to improve the quality of their keyword recommendation algorithms. We have also checked the robustness of estimation results by using KS with unweighted averages and found that estimated coefficients are qualitatively unchanged. Finally, keyword popularity is captured by $Nlag_{kit}(T)$, which is the weighted average number of type T competing advertisers who bought keyword k previously. We assume the decay parameter ρ to be the same in *Nlag*, *Elag*, and *RalgInv* to keep the model parsimonious. The parameters $\lambda_{3TT'}$ indicate how advertisers use previous competition information to construct their keyword consideration sets.

We have now completed the model of keyword selection, which leads to the following likelihood function of parameters γ and λ in keyword entry and consideration models, given the observations of advertisers' keyword entry:

$$L(\gamma,\lambda\mid E)=\prod_{t=1}^{S}\prod_{k=1}^{K}\left[\sum_{\forall I_{kt}}[P(I_{kt})P(E_{kt}\mid I_{kt})]\right], \quad (23)$$

$$P(I_{kt}) = \prod_{i=1}^{I} [P(V_{kit} > 0)^{1\{i \in I_{kt}\}} P(V_{kit} \le 0)^{1\{i \notin I_{kt}\}}], \quad (24)$$

$$P(E_{kt} \mid I_{kt}) = 1\{E_{kt} \subseteq I_{kt}\} \cdot \prod_{i \in I_{kt}} \left[P(U_{kit} > 0 \mid I_{kt})^{1\{i \in E_{kt}\}} P(U_{kit} \le 0 \mid I_{kt})^{1\{i \notin E_k\}}\right]. \quad (25)$$

5. Estimation Strategy

5.1. An Overview of the Estimation Method

We present an overview of our estimation strategy and discuss several econometric challenges. We first jointly estimate the equations of average click volume and average CPC along with the equations of the number of advertisers to deal with the potential endogeneity problem using IVs proposed in Section 4.3. We adopt the Bayesian estimation approach implemented via the Markov Chain Monte Carlo (MCMC) algorithm to make model inferences. We generate 10,000 iterations for each model and use every 10th of the last 5,000 draws for the inferences of parameter estimates. The iteration plots are monitored and inspected to determine convergence of the sampler. The detailed algorithm appears in Section A.1 in the appendix.

To control for the potential sample selection in the model of ad ranking, we estimate the equation of ad ranking jointly with equations of entry and consideration, which serve the selection equation indicating the set of entrants in a keyword market. Here we use a reduced-form specification of entry equation, which is related to the two-step approach, as we will discuss later in Section 5.2.

Our integrated model of entry and consideration equations presents two main challenges for estimation. The first problem is the curse of dimensionality. As noted in Equation (23), the likelihood function of keyword choice requires the summation across all possible sets of potential entrants for each keyword, the number of which can be as large as $2^I - 1 \approx 2.7 \times 10^8$ in our empirical context. Therefore, it is infeasible to integrate out all permutations of I_{kt} to compute the unconditional likelihood function. Second, conditional on I_{kt}, the vector of entry probabilities for potential entrants is implicitly determined under an equilibrium condition. However, the possibility of multiple equilibria prevents us from deriving the conditional likelihood of entry observations expressed in Equation (25).

We adopt the Bayesian estimation approach to cope with the first challenge. We follow van Nierop et al. (2010) in treating entry and consideration decisions as indicators of latent utilities, whose posteriors are either normally or truncated normally distributed. By augmenting the model with these latent variables, we have a deterministic set of potential entrants for each keyword during the MCMC iteration. Thus, the Bayesian method helps us bypass the necessity of integrating out all sets of potential entrants.

For the second challenge, previous literature suggested three approaches to tackle multiple equilibria in discrete games with incomplete information (for a more detailed discussion, see Ellickson and Misra 2011). The first method is the two-step estimator, which was originally introduced by Hotz and Miller (1993) in the dynamic discrete choice context. Recently, Bajari et al. (2010) proved the consistency of this estimator for a static discrete game and provided the identification conditions. The first stage is to find consistent estimates for those conditional choice probabilities (CCPs) such as the entry probability in Equation (18). Then, in the second stage, researchers can use these predicted CCPs as covariates to estimate the parameters of strategic interactions. The second method is the nested pseudolikelihood approach introduced by Aguirregabiria and Mira (2007). This method relies on repetitive procedures of best response iteration and pseudolikelihood maximization. The final estimation method is the mathematical programming with equilibrium conditions proposed by Su and Judd (2012), which treats the likelihood-maximization problem as a constrained optimization problem with equilibrium constraints. The last two methods each involve an optimization procedure, which encounters the first challenge we mentioned above. Therefore, we follow Bajari et al. (2010) and adopt the two-step estimation approach.

5.2. A Two-Step Estimation Approach

We first discuss why the standard two-step method cannot be directly applied to our empirical context. In a discrete game where the set of players is observed, researchers typically estimate the CCPs based on observed state variables using either a nonparametric method or a flexible parametric specification in the first step. Here state variables refer to the whole set of covariates that can affect a player's choice probability. Unfortunately, since we intend to be agnostic about the set of potential entrants in this study, we are unable to infer the CCPs based on the relationship between observed state variables and observed choice decisions. Next we discuss how we adapt the two-step method to estimate the proposed model. The details regarding the estimation algorithm for each stage are in Sections A.2 and A.3 of the appendix.

Step 1. The first-step estimation aims to predict the equilibrium entry probability $P^*(a_{kit} \mid I_{kt})$ for all possible I_{kt} and obtain consistent estimates of coefficients in the rank equation, which can be used to construct N^a_{kit} and N^b_{kit} in Equations (11) and (12). For each I_{kt}, we denote the vector of state variables by $S_{kt}(I_{kt})$, which includes all relevant keyword characteristics and previous entry and rank information of all potential entrants. We assume that the equilibrium entry probability $P^*(a_{kit} \mid I_{kt})$ can be approximated by a flexible probit specification with all $S_{kt}(I_{kt})$ as covariates, and there are over 300 covariates in our empirical context. We then employ the Bayesian approach to estimate this probit entry model jointly with the ranking model in Equation (1) and the consideration model in Equation (19).

Step 2. In the second step, we attempt to obtain consistent estimates on parameters of strategic interactions. Given the MCMC draws from the first step, we use every 20th of the last 10,000 of 20,000 draws from the posterior distributions to predict $\hat{P}^*(a_{kit} \mid I_{kt})$ and to construct the expected number of entrants ranked above $\hat{E}(N^a_{kit}(I_{kt}))$ and ranked below $\hat{E}(N^b_{kit}(I_{kt}))$, both of which appear in the right-hand side of entry utility in Equation (13). Furthermore, the expected average click volume and average CPC depend on the expected entry probabilities of firms and therefore are functions of the vector of $\hat{P}^*(a_{kit} \mid I_{kt})$. We plug the covariates of $\hat{E}(N^a_{kit}(I_{kt}))$, $\hat{E}(N^b_{kit}(I_{kt}))$, $\hat{E}(Click_{kt} \mid I_{kt})$, and $\hat{E}(CPC_{kt} \mid I_{kt})$, which are predicted from the first-step estimation, into the equation of entry payoff. We then reestimate the keyword entry and consideration equations jointly.

5.3. Identification

We next discuss the parameter identification in keyword entry and consideration equations. As in standard discrete choice models augmented with consideration decisions, the identification of nonstrategic parameters relies on the covariation between the explanatory variables and the revealed keyword selections. The separate identification of parameters in the consideration model from the entry model is achieved through variable exclusions. Specifically, the identification strategy is based on the assumption that keyword similarity (*KS*) affects only a firm's probability of considering a keyword and not the entry payoff. This assumption is consistent with the industrial practice that firms rely heavily on keyword recommendation tools such as those offered by Google AdWords to construct their keyword consideration sets. As most of these recommendation tools use proximity-based algorithms, a keyword that is semantically close to the set of keywords used previously by the firm (i.e., with a larger *KS*) is likely to be recommended and thus considered by the firm in the current period. Meanwhile, a keyword with a larger *KS* does not necessarily imply a higher payoff after controlling for other keyword characteristics. Thus, *KS* can serve as an exclusion restriction to help us separately identify parameters in the consideration set model and in the keyword entry model.

As established by Bajari et al. (2010), the identification of the structural parameters in the keyword entry model (i.e., $\gamma_{3TT'}$ and γ_{4T} in Equation (13)) rests on exclusion restrictions, which refer to variables that directly affect the payoffs of each individual but not the payoffs of other players. We propose using a firm's keyword choice in the UK market, denoted by *UK_entry*, as an exclusion restriction. We expect a firm's keyword decision in the United States to be positively correlated with *UK_entry*. The rationale is if a firm has advertised through a particular keyword in the UK market, this firm is likely to incur a lower cost in designing ad copy for the same keyword and therefore is more likely to use this keyword in its search advertising campaign in the U.S. market. However, since consumers in the United States do not observe ads associated with the same keyword in the UK market, a firm's keyword entry probability in the U.S. market should not be affected by its competitors' ads in the UK market, which makes *UK_entry* a valid exclusion restriction in our empirical context. As noted previously, *UK_entry* is unavailable to seven retailers and two comparison sites who have not bought keywords from Google UK. Nevertheless, since our research focuses on type-specific spillovers rather than the spillovers of individual firms' entry, we are still able to use *UK_entry* as an exclusion restriction for the identification of spillover between different types of firms under the assumption that those firms who have used Google UK are representative.

We conduct a reduced-form analysis to check the validity of our proposed variable as an exclusion restriction. Our estimation results from a probit model with random coefficients for the intercept in Table 4 show that the estimated coefficient of *UK_entry* is significantly positive. More important, after controlling for the competitors' actions in the U.S. market (i.e., the number of competitors and their average rank), the number of competitors in the United Kingdom has no significant impact on a firm's keyword entry probability. These results empirically justify our use of *UK_entry* as an effective exclusion restriction to help identify the strategic interaction parameters.

To alleviate the concern of missing data in *UK_entry*, we gathered information on a new variable, *US_2yrs*, from the data provider. The term *US_2yrs* is a dummy variable indicating whether a firm bought the same keyword in the same month during the past two years from Google U.S. We believe that *US_2yrs* is correlated with a firm's current U.S. entry because the ad copy used in the same keyword and in the same month can help a firm reduce the cost of ad copy design and experimentation, and therefore increase the entry probability. The term *US_2yrs* is unlikely to be correlated with the demand shock in the current period as long as the serial correlation in demand shock, if any, is not too strong. This makes *US_2yrs* a valid exclusion restriction for U.S. entry. We checked the robustness of our results by estimating the model using both *UK_entry* and *US_2yrs* as exclusion restrictions in the entry equation. Our key results are qualitatively unchanged, suggesting that the missing data problem in *UK_entry* is not serious.

Table 4. Reduced-Form Estimation of Keyword Entry Model

	Parameter estimates
Keyword attributes	
Match	0.041 (0.029)
Length	**−0.022 (0.010)**
Specific	**−0.099 (0.008)**
Promotional	**0.067 (0.020)**
ln(*SV*)	**−0.042 (0.003)**
ln(*CPC*)	0.005 (0.009)
ln(*Click*)	0.005 (0.003)
State dependence	
$Entry_{t-1}$	**1.531 (0.023)**
$\ln(Rank_{t-1}+1)$	**−0.308 (0.012)**
Organic and UK presence	
Organic_dummy	**1.288 (0.036)**
ln(*Rank_organic* + 1)	**0.191 (0.014)**
UK_entry	**0.886 (0.032)**
Competition effects	
$\ln(N+1)$	**1.021 (0.021)**
ln(*AvgRank*+1)	**−0.711 (0.023)**
$\ln(N^{\mathrm{UK}}+1)$	0.010 (0.011)
Intercept	**−2.324 (0.037)**

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

We next discuss the identification for the effect of $E(N^a_{kit})$ and $E(N^b_{kit})$ on entry probability. Recall that the $\Pr(Rank_{kjt} < Rank_{kit})$ is identified based on firms' ranking conditional on their entry. The definitions of $E(N^a_{kit})$ and $E(N^b_{kit})$ suggest that the effect of these two variables can be regarded as the interaction effect of firms' relative ranking and the spillover of others' entry, the latter of which is identified through exclusion restrictions. Finally, we perform a simulation exercise to show the theoretical identification of our proposed joint model of entry and consideration in Section A.4 in the appendix.

6. Results and Counterfactuals

We select the best model by applying the proposed two-step estimation approach to the joint model of entry and consideration with different values of the decay factor ρ (i.e., autoregressive coefficient in Equations (14), (15), and (22)). We choose $\rho = 0.7$ based on the model fitness statistics reported in Table 5 by lowering ρ from 0.9 by a step size of 0.1.

To validate our proposed joint model of entry and consideration, we compare it with a benchmark model in which all advertisers are exogenously treated as potential entrants. In other words, there is no consideration stage in the benchmark model. The entry utility of this benchmark case is modeled in a similar way as in Equation (13) except for the inclusion of additional covariates that occur in our proposed consideration models. By doing this, we ensure that the total sets of covariates used in these two models are exactly the same. We estimate this benchmark model also using a two-step approach. We find that our proposed entry model with a consideration stage (log marginal = −58,993) outperforms the entry model without a consideration stage (log marginal = −61,604). This suggests the importance of accounting for firms' keyword consideration in the context of sponsored search advertising.

Table 5. Model Fitness Statistics of Joint Models of Entry and Consideration

Model specification	Log marginal
With consideration	
$\rho = 0.9$	−59,275
$\rho = 0.8$	−59,057
$\rho = 0.7$	−58,993
$\rho = 0.6$	−59,098
$\rho = 0.5$	−59,229
Without consideration	
$\rho = 0.7$	−61,604

6.1. Postentry Outcomes: Ad Position, Average Click Volume, and Average CPC

Table 6 reports the estimation results of the ad position model. Consistent with Google's practice of using relevance between ad and keyword as an important determinant of the ad's quality, our results indicate that a manufacturer obtains a higher position in a matched keyword. Regarding the effects of keyword characteristics, we find that both retailers and comparison sites tend to rank higher than manufacturers in less costly keywords with low search but high click traffic. We also find significant state-dependence effects of previous entry and rank. As expected, an advertiser who bought the keyword or achieved higher ranks in previous periods is generally more likely to maintain a higher position. We also find that a firm with a top-50-ranked organic presence tends to stay at a lower position on sponsored listings, perhaps because of the firm's declining incentive to raise their bid in auctions for this keyword. Finally, our results from the first-step estimation show that there is a significant and negative correlation ($\delta = -70\%$) between unobserved factors in the entry and ranking equations, suggesting the importance of controlling for sample selection in the ad position model.

We summarize our findings on the average click volume and the average CPC in Table 7. We note that these two advertising metrics vary across keywords with different attributes. For example, keywords with more promotional terms are associated with higher average click volume and CPCs. The regression results of endogenous variables (i.e., number of entrants) in

Table 6. Estimation Results of Position Model

	Manufacturers	Retailers	Comparison sites
Keyword attributes			
Match	**2.025 (0.050)**	N.A.	N.A.
Length	N.A.	**0.159 (0.018)**	**0.159 (0.024)**
Specific	N.A.	**0.177 (0.020)**	**−0.091 (0.022)**
Promotional	N.A.	−0.018 (0.036)	−0.015 (0.042)
$\ln(SV)$	N.A.	**−0.056 (0.007)**	**−0.083 (0.010)**
$\ln(CPC_{t-1})$	N.A.	**−0.064 (0.020)**	**−0.118 (0.025)**
$\ln(Click_{t-1})$	N.A.	**0.027 (0.007)**	**0.022 (0.009)**
State dependence			
$Entry_{t-1}$	**1.644 (0.049)**	**1.922 (0.043)**	**0.895 (0.064)**
$Entry_{t-2}$	**0.481 (0.049)**	**0.297 (0.047)**	**0.379 (0.070)**
$Entry_{t-3}$	**0.582 (0.055)**	**0.320 (0.055)**	**0.312 (0.079)**
$Entry_{t-4}$	**0.294 (0.062)**	**0.268 (0.060)**	0.009 (0.104)
$Entry_{t-5}$	**0.219 (0.071)**	**0.329 (0.070)**	**0.281 (0.106)**
$Entry_{t-6}$	**0.366 (0.077)**	0.022 (0.074)	**−0.316 (0.101)**
$\ln(Rank_{t-1}+1)$	**−0.553 (0.026)**	**−0.626 (0.020)**	**−0.436 (0.028)**
$\ln(Rank_{t-2}+1)$	**−0.173 (0.027)**	**−0.146 (0.024)**	**−0.206 (0.033)**
$\ln(Rank_{t-3}+1)$	**−0.222 (0.031)**	**−0.138 (0.028)**	**−0.208 (0.034)**
$\ln(Rank_{t-4}+1)$	**−0.106 (0.032)**	**−0.088 (0.029)**	**−0.120 (0.044)**
$\ln(Rank_{t-5}+1)$	−0.073 (0.041)	**−0.098 (0.033)**	**−0.147 (0.046)**
$\ln(Rank_{t-6}+1)$	**−0.146 (0.044)**	−0.010 (0.036)	0.091 (0.049)
Organic and UK presence			
Organic_dummy	**−1.766 (0.060)**	**−2.215 (0.053)**	**−1.335 (0.079)**
$\ln(Rank\_organic+1)$	**−0.066 (0.020)**	**−0.132 (0.018)**	**−0.059 (0.026)**
UK_entry	−0.032 (0.064)	−0.008 (0.052)	0.158 (0.081)
$\ln(Rank\_UK+1)$	−0.036 (0.040)	−0.043 (0.028)	−0.051 (0.040)

Notes. Estimates of the advertiser-specific fixed effect and the monthly fixed effect are not reported. Estimates that are significant at 95% are bolded.

Table 8 confirm the relevance condition of using N^{UK} as IVs. The variance–covariance estimates in Table 9 provide some evidence for the endogeneity on the number of comparison sites in the CPC equation.

After controlling for endogeneity, we find that the average click per ad is positively associated with the number of retailers and negatively associated with the number of manufacturers. This might happen because retailers tend to use search advertising to boost sales, while manufacturers' main advertising objective is to increase brand awareness. As a result, consumers are likely to have a higher (lower) clicking propensity on retailers' (manufacturers') ads. Our results also show that a keyword with more manufacturers and fewer comparison sites tends to have a higher CPC. We conjecture that this is because manufacturers (comparison sites) typically have higher (lower) weighted bids than average. This is consistent with our observation that the average ad positions are 4.16, 8.08, and 6.42 for

Table 7. Estimation Results of Click Volume and CPC Models

	$\ln(Click)$	$\ln(CPC)$
Keyword attributes		
Intercept	**−0.532 (0.099)**	**−0.153 (0.042)**
Length	0.064 (0.033)	**0.074 (0.013)**
Specific	0.020 (0.028)	**−0.095 (0.012)**
Promotional	**0.130 (0.050)**	**0.094 (0.021)**
$\ln(SV)$	**0.160 (0.021)**	−0.001 (0.004)
Match_entry	−0.004 (0.025)	−0.004 (0.010)
State dependence		
$\ln(CPC_{t-1})$	N.A.	**0.462 (0.010)**
$\ln(Click_{t-1})$	**0.742 (0.029)**	N.A.
Organic presence		
$\ln(N^O(M)+1)$	0.053 (0.034)	−0.011 (0.014)
$\ln(N^O(R)+1)$	0.018 (0.032)	0.021 (0.012)
$\ln(N^O(C)+1)$	0.029 (0.027)	0.021 (0.011)
Effect of competition		
$\ln(N(M)+1)$	**−0.143 (0.043)**	**0.097 (0.022)**
$\ln(N(R)+1)$	**0.202 (0.041)**	−0.056 (0.031)
$\ln(N(C)+1)$	−0.059 (0.067)	**−0.088 (0.026)**

Notes. Estimates of the monthly fixed effect are not reported. Estimates that are significant at 95% are bolded.

Table 8. Estimation Results of Number of Entrants

	$\ln(N(M)+1)$	$\ln(N(R)+1)$	$\ln(N(C)+1)$
Keyword attributes			
Intercept	**0.466 (0.024)**	**0.434 (0.026)**	**0.281 (0.024)**
Length	−0.005 (0.007)	**0.163 (0.009)**	**0.046 (0.008)**
Specific	**−0.137 (0.006)**	**−0.151 (0.006)**	**−0.017 (0.006)**
Promotional	**0.029 (0.013)**	**0.065 (0.015)**	**0.057 (0.015)**
$\ln(SV)$	−0.003 (0.002)	0.005 (0.003)	**−0.034 (0.009)**
State dependence			
$\ln(N_{t-1}+1)$	**0.564 (0.009)**	**0.332 (0.010)**	**0.349 (0.009)**
UK presence			
$\ln(N^{UK}+1)$	**0.054 (0.011)**	**0.145 (0.013)**	**0.229 (0.012)**

Notes. Estimates of the monthly fixed effect are not reported. Estimates that are significant at 95% are bolded.

Table 9. Variance–Covariance Estimates of Click Volume and CPC Models

Variance-covariance matrix Φ	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
Error term in ln(*Click*) (ϕ_1)	**0.959**	**0.062**	0.011	−0.025	−0.003
	(0.027)	**(0.005)**	(0.008)	(0.015)	(0.012)
Error term in ln(*CPC*) (ϕ_2)		**0.159**	−0.002	**0.013**	**0.018**
		(0.003)	(0.003)	**(0.006)**	**(0.005)**
Error term in $\ln(N(M)+1)$ (ϕ_3)			**0.158**	**0.043**	**0.031**
			(0.002)	**(0.002)**	**(0.002)**
Error term in $\ln(N(R)+1)$ (ϕ_4)				**0.227**	**0.082**
				(0.03)	**(0.003)**
Error term in $\ln(N(C)+1)$ (ϕ_5)					**0.208**
					(0.003)

Variance-covariance matrix Ω	ω_1	ω_2
Keyword-specific error term in ln(*Click*) (ω_1)	**0.163 (0.048)**	0.005 (0.004)
Keyword-specific error term in ln(*CPC*) (ω_2)		**0.039 (0.002)**

Note. Estimates that are significant at 95% are bolded.

manufacturers, comparison sites, and the full sample in our data, respectively.

6.2. Keyword Choice or Entry

Table 10 reports parameter estimates in the keyword entry equation.[2] We begin with a discussion on the effects of keyword attributes on advertisers' keyword choices. We find that all three types of advertisers have a preference for shorter keywords with smaller search volume. Comparison sites tend to use more specific keywords. One possible explanation is that comparison sites often have smaller advertising budgets than manufacturers and retailers, and therefore they favor niche keywords that are searched by users with higher purchase intention. We also find that both retailers and comparison sites are in favor of promotional keywords with low CPC.

Our results show a significant state dependence in advertisers' keyword decisions. In general, firms are more likely to advertise through a keyword that was frequently used or associated with a high ad position in the past. We also see that firms are more likely to use keywords with a top-50 organic presence, which can be explained by the positive interplay between payoffs from organic and paid results found in Yang and Ghose (2010). The significantly positive coefficients of firms' presence on sponsored listings in the UK support the validity of using *UK_entry* to provide exclusion restrictions.

Now let us turn to our first research question (RQ1) regarding the spillover effect. We hypothesized that while the spillover effect from above-ranked ads can be either positive or negative, the spillover effect from below-ranked ads is positive. Our hypothesis is supported by the empirical results. We find a positive spillover effect of below-ranked ads on an advertiser's entry probability, suggesting that given an advertiser's ranking, the larger the total number of paid search ads showing for a keyword, the higher the advertiser's entry probability. A larger number of below-ranked ads may signal a higher demand for the keyword market (i.e., market expansion effect) and consequently increase the advertiser's entry probability. At the same time, users often search information (i.e., ads) from top to bottom, which suggests that below-ranked ads are unlikely to affect a user's clicking propensity on the focal ad (i.e., no business stealing effect), and therefore do not impact the focal advertiser's entry likelihood. However, the business stealing effect applies to above-ranked ads. As expected, we find that the spillover from above-ranked ads is either positive or negative, since it is driven by the two opposing forces, namely, the negative business stealing effect and the positive market expansion effect.

Our second research question (RQ2) examines how product-line characteristics of firms affect the impact from above-ranked ads. We start with spillovers between advertisers of the same type. Consistent with our conjecture, our results indicate that the spillovers between manufacturers and between retailers are positive, whereas the spillover between comparison sites is negative. One possible explanation is that the spillover effect is moderated by the degree of product differentiation offered by competing firms. Since the products or services offered by comparison sites are mostly homogenous, consumers incur a lower marginal search benefit when they search and click sequentially along the ad listings. This implies a larger business stealing effect and a smaller market expansion effect of competing ads on keyword entry. On the other hand, as most manufacturers and retailers we study are household brands, the products from these firms are often

Table 10. Estimation Results of Keyword Entry Model

	Manufacturers	Retailers	Comparison sites
Keyword attributes			
Match	**0.204 (0.068)**	N.A.	N.A.
Length	**−0.206 (0.017)**	**−0.117 (0.014)**	**−0.186 (0.015)**
Specific	−0.000 (0.022)	−0.001 (0.015)	**0.242 (0.016)**
Promotional	0.021 (0.027)	**0.153 (0.021)**	**0.110 (0.026)**
$\ln(SV)$	**−0.050 (0.006)**	**−0.028 (0.005)**	**−0.037 (0.006)**
$E(\ln(CPC))$	0.029 (0.027)	**−0.104 (0.021)**	**−0.070 (0.022)**
$E(\ln(Click))$	−0.024 (0.015)	−0.003 (0.004)	−0.009 (0.005)
State dependence			
Elag	**0.367 (0.022)**	**0.117 (0.019)**	**0.436 (0.017)**
RlagInv	**0.073 (0.032)**	**0.432 (0.047)**	**0.142 (0.066)**
Organic and UK presence			
Organic_dummy	**1.838 (0.069)**	**1.766 (0.076)**	**1.356 (0.108)**
$\ln(Rank\_organic+1)$	−0.053 (0.028)	**0.321 (0.028)**	0.058 (0.038)
UK_entry	**0.129 (0.034)**	**0.202 (0.028)**	**0.145 (0.027)**
Spillover effect of ads above			
$\ln(N^a(M)+1)$	**0.582 (0.050)**	−0.045 (0.033)	**−0.101 (0.035)**
$\ln(N^a(R)+1)$	0.088 (0.049)	**0.392 (0.045)**	−0.088 (0.046)
$\ln(N^a(C)+1)$	**−0.400 (0.054)**	**0.310 (0.047)**	**−0.167 (0.044)**
Spillover effect of ads below			
$\ln(N^b+1)$	**0.962 (0.031)**	**0.448 (0.039)**	**1.642 (0.043)**
Intercept		**−2.947 (0.068)**	

Notes. Estimates of the advertiser-specific fixed effect and the monthly fixed effect are not reported. Estimates that are significant at 95% are bolded.

horizontally differentiated from the consumers' perspective. The larger heterogeneity in products offered by different manufacturers and retailers suggests a smaller business stealing effect and a greater market expansion effect, which explain the positive spillovers found between same-type firms from these two types.

Our third research question (RQ3) examines how spillovers vary by different types of firms in a distribution channel. As expected, there exists a positive spillover from downstream on upstream firms' entry probability. To see this, the spillover from retailers on manufacturers (0.088) is positive and marginally significant, and the spillover from comparison sites on retailers (0.310) is significantly positive. This positive spillover is possibly driven by the fact that upstream firms generally benefit from the increased sales from downstream firms, but not vice versa. One exception here is that manufacturers are less likely to use keywords with a larger number of comparison sites' ads. This may occur because manufacturers do not directly work with comparison sites, and therefore the expanded demand by comparison sites, if any, might not benefit manufacturers. Consequently, the direct harm of business stealing from comparison sites' ads dominates the potential benefit of market expansion. As for the spillover from upstream on downstream firms, we find that the signs of coefficient estimates are all negative, suggesting the dominance of the business stealing effect.

6.3. Keyword Consideration

On our fourth research question (RQ4), we have two main findings on how advertisers utilize competition information to update their keyword dictionaries based on the estimated coefficients reported in the three-by-three matrix in Table 11. First, the positive coefficients on the diagonal suggest that manufacturers and retailers are more likely to consider keywords that have been used by competitors' of the same type. This is consistent with the conventional belief that the best way to obtain keyword knowledge is to learn from a firm's direct competitors.

Second, the positive coefficients on the upper triangular matrix indicate that for the retailer–manufacturer pair, and the comparison site–retailer pair, the former downstream firm is more likely to consider keywords previously purchased by the latter upstream firm. Conversely, the upstream firm is less likely to consider keywords that have been used by the downstream firm in the past, as suggested by the negative coefficients on the lower triangular matrix. This finding is likely driven by the keyword knowledge structure of firms at different positions of a channel. In general, manufacturers stay upstream in a distribution channel, comparison sites are downstream, and retailers are in the middle. Since manufacturers are closely involved in producing the product, they tend to be more knowledgeable about the appropriate keywords for their business than comparison sites and retailers who are mainly involved in distributing the product along with

Table 11. Estimation Results of Keyword Consideration Model

	Manufacturers	Retailers	Comparison sites
Keyword loyalty (*Elag*)	**1.738 (0.072)**	**1.265 (0.085)**	**1.642 (0.058)**
Keyword similarity	**1.722 (0.144)**	**1.461 (0.104)**	**1.319 (0.119)**
Keyword popularity			
$\ln(Nlag(M)+1)$	**0.309 (0.088)**	**0.263 (0.043)**	0.055 (0.066)
$\ln(Nlag(R)+1)$	**−0.908 (0.106)**	**0.532 (0.048)**	**0.443 (0.080)**
$\ln(Nlag(C)+1)$	**−0.748 (0.081)**	**−0.710 (0.057)**	0.073 (0.100)
Intercept		**1.762 (0.223)**	

Notes. Estimates of the advertiser-specific fixed effect are not reported. Estimates that are significant at 95% are bolded.

thousands of other products from a variety of product categories. Therefore, manufacturers are likely to be at a more advanced stage in optimizing the keyword dictionary by improving the relevance of keywords, yet comparison sites are more likely to be at a discovery stage in expanding the keyword dictionary, and retailers' situation may be in-between. This knowledge hierarchy may explain why downstream firms are inclined to expand the list of keywords based on past keyword choices by upstream firms, but not vice versa.

Finally, our analysis suggests that keyword loyalty is a significant factor in keyword consideration for both manufacturers and retailers. We also find that all three types of firms are likely to consider terms that are similar to terms in their previous keyword portfolio, which is consistent with the industrial practice of using recommendation tools to generate keyword ideas.

6.4. Counterfactual Simulations

The structural nature of our model can help search engines evaluate the implications of the existence of third-party keyword infomediaries on their revenue. Intuitively, if the spillover in a keyword market is mostly positive, the search engine is likely to benefit from the competition data provided by infomediaries. When competition data are available, a firm can observe its competitors' past keyword choices and use this information to better assess competitors' entry probability in the current period. Since a firm's past keyword decisions often indicate its higher likelihood of using the same keyword in the current period, each firm will have a higher belief on others' entry probability compared to the case without competition data. Positive spillovers suggest that a firm is more likely to buy a keyword if it believes that competitors have a higher probability of buying the same keyword. Hence, the positive spillover effect leads to market expansion (i.e., more entrants per keyword market), which benefits the search engine. By contrast, the dominance of negative spillover in a keyword market is expected to hurt search engines when competition data are provided. Since our results indicate that both positive and negative spillovers exist, it is rather difficult to make a directional prediction on the impact of the availability of competition data on search engines' revenue a priori.

To quantify the monetary value of providing competition data, we consider two counterfactual scenarios in which advertisers either have no access to competition information or have access to only aggregate competition information. This can happen if search engines such as Google start to restrict the commercial use of competitive intelligence by third-party infomediaries. We compare the search engine's expected revenue under three scenarios: (i) when infomediaries do not exist, advertisers have no access to past competition information; (ii) when infomediaries do exist, but they only provide aggregate competition information so that each firm can only see the number of keywords purchased by each individual competitor and its average ad rank every month; and (iii) when infomediaries do exist, they provide detailed competition information so that each firm can observe the complete keyword purchase histories and the associated ad ranks of competitors.

For each scenario, we simulate advertisers' keyword choices in the last month in our data period by constructing advertisers' keyword considerations and then solving the simultaneous-move game of keyword entry. The availability and the granularity of past competition information can affect firms' keyword choices in two ways: (i) they affect how firms utilize others' past keyword choices when forming their own keyword consideration sets (i.e., the construction of *Nlag* in Equation (19)), and (ii) they affect firms' predictions on competitors' keyword entry probability (i.e., the construction of *Elag* and *RlagInv* in Equation (13)). After solving the entry equilibrium, we use the average click volume and CPC under the new equilibrium to compute the search engine's advertising revenue. To alleviate the possibility of multiple equilibria of the entry game, we randomize the starting value for the vector of entry probability in each keyword during each of our 1,000 simulations.

Table 12 reports findings from this counterfactual experiment. We find that as keyword infomediaries provide aggregate competition data, both retailers and comparison sites will buy 7.7% and 21% more keywords, respectively, whereas manufacturers will buy 14% fewer. The keyword expansion effect is the largest for comparison sites, consistent with the notion that comparison sites use infomediaries as a tool to learn new keywords from others. Compared with the aggregate information, the availability of precise competition information will encourage manufacturers to buy and discourage retailers and comparison sites from buying more keywords from the search

Table 12. Counterfactual Results of Advertisers' Keyword Choices and Search Engine Revenue

	Manufacturers	Retailers	Comparison sites	Search engine
	Average number of ads per keyword			
Without competition info	2.37	2.63	1.15	N.A.
Aggregate competition info	2.05 (−14.4%)	2.83 (7.67%)	1.39 (21.0%)	N.A.
Accurate competition info	2.02 (−13.4%)	2.80 (6.49%)	1.36 (18.2%)	N.A.
	Average number of keywords purchased per advertiser			
Without competition info	371	274	180	N.A.
Aggregate competition info	321 (−14.4%)	295 (7.67%)	218 (21.0%)	N.A.
Accurate competition info	316 (−13.4%)	292 (6.49%)	213 (18.2%)	N.A.
	Search engine revenue per keyword per day ($)			
Without competition info	N.A.	N.A.	N.A.	34.3
Aggregate competition info	N.A.	N.A.	N.A.	34.7 (1.12%)
Accurate competition info	N.A.	N.A.	N.A.	36.2 (5.68%)

Notes. The case without competition information is the benchmark. The percentage change of each variable with respect to the benchmark is reported in parentheses.

engine. Note that we observe the average click volume and CPC for each keyword but not for each individual advertiser. Because of this data limitation, we are unable to assess the impact of policy change on each advertiser's profit. Nevertheless, the search engine revenue can be computed by multiplying of average click volume, average CPC, and number of entrants, each of which can be predicted in our counterfactual. Our results indicate that the search engine will expect to increase its advertising revenue by 5.7% (1.1%) if the keyword-level (aggregate-level) competition data are accessible to advertisers. For consumers, an ideal way of assessing their benefits is based on conversions. Unfortunately, we do not observe conversions in our data. Instead, we use clicks as a proxy to gauge consumer benefit under the assumption that consumers are more likely to click when ads are more relevant. Our results indicate that the total number of clicks increases by 7.2% (3.2%) if the keyword-level (aggregate-level) competition data are available, suggesting an improvement in consumer benefits by the service offered by keyword infomediaries.

Taken together, our findings suggest that infomediaries create value to the search engine, help retailers and comparison sites discover more keywords, allow manufacturers to refine their keyword dictionaries, and increase consumers' clicks. One potential caveat concerning the application of our counterfactual results is that we model the postentry outcomes (e.g., CPC, click volume, ranking) in a reduced form instead of based on a first principle, mainly because of the data limitation. Despite this limitation, we believe that this analysis is still meaningful by serving as the first step to evaluating the value of competition information in sponsored search advertising.

7. Conclusion

Despite the importance of keyword management in search engine marketing, it has not received enough attention from marketing researchers. In this paper, we regard each keyword as a market and develop a structural model to examine the spillover effects in advertisers' keyword market entry decisions by addressing the following four questions: (RQ1) How does the spillover effect vary by competitors' relative positions on ad listings? (RQ2) How does the spillover effect vary by the heterogeneity in firms' product offerings? (RQ3) What is the nature of the spillover effect between upstream and downstream firms in a distribution channel? (RQ4) How do firms update keyword dictionaries based on competitors' keyword histories? The answers to these questions have important implications for both search advertisers and the search engine to fully understand the nature of clustering patterns in a keyword market, predict the competition intensity, and assess the optimal granularity of competition information to release.

Our model is built on the incomplete-information simultaneous-entry framework, with two novel extensions to handle the complexities that arise in the sponsored search advertising context. First, we allow the strategic interactions to be advertiser-type-specific and dependent on the advertisements' relative positions. Second, we formally model the construction of advertisers' keyword dictionaries or keyword consideration sets and allow them to vary with the observable historical competition knowledge for the same keyword.

We find a presence of both positive and negative spillovers effects in keyword markets. For spillovers from ads ranked below, the effect is always positive for all three types of advertisers, suggesting that an advertiser's entry probability increases with the expected number of search ads ranked below. For spillovers from ads ranked above, we find both positive and negative effects depending on the advertiser type combinations. In particular, the spillover among firms offering homogenous products (i.e., comparison sites) is negative, while the spillover among firms offering more differentiated products (i.e., manufacturers and retailers) is positive. Furthermore, the spillover is found to be asymmetric between upstream and downstream firms in a distribution channel. The spillover from upstream on downstream firms (i.e., manufacturers on comparison sites) tends to be negative, and the spillover from downstream on upstream firms (i.e., comparison sites on retailers) tends to be positive. As for keyword consideration, both manufacturers and retailers are likely to learn from the same types of other advertisers, and a downstream firm who typically has less product knowledge tends to learn new keywords from an upstream firm.

Our study has limitations that open doors for future research. First, jointly modeling advertisers' bidding and keyword entry decisions can provide a more comprehensive understanding of the role of competition in search engine marketing. Unfortunately, we are unable to examine this issue because of a lack of data on individual entrant advertisers' bids. Second, while we focus on spillovers among advertisers' entry decisions on individual keywords, cross-keyword spillover effects could also exist for an individual advertiser. Recovering the keyword consideration set is a first step we take toward that direction. Third, our research can be extended to account for strategic collusions among channel members like manufacturers and retailers. If so, we need more institutional details to guide the characterization of advertisers' payoff functions. Finally, while this research has studied spillovers among advertisers' keyword choices on the same search engine, future work could examine advertisers' keyword competition across different platforms. Overall, we hope this study stimulates further interest in advertisers' keyword strategies, as sponsored search advertising continues to thrive and grow.

Acknowledgments

The authors thank the anonymous company that provided data for this study. The authors are grateful to the editor-in-chief, associate editor, and two anonymous referees for their constructive comments. The authors also thank participants at the 2015 Marketing Science Conference, Ohio State University, UNC Marketing Camp, Central South University, University of Frankfurt, Erasmus University, Georgetown University, UC-Davis Marketing Camp, Southern Methodist University, and University of British Columbia for their valuable comments.

Appendix. The MCMC Algorithm

A.1. Estimation Algorithms for Average Click Volume and Average CPC

We ran the MCMC chain for 10,000 iterations and use every 10th draw of the last 5,000 iterations to compute the mean and standard deviation of the posterior distribution of model parameters. We report the MCMC algorithm for the joint models of average click volume and average CPC.

1. Draw β^Q.

We rewrite the model of average click volume in the form $\ln(Click_{kt}) = \beta^Q X^Q_{kt} + \eta^Q_k + \upsilon^Q_{kt}$. The posterior distribution of β^Q depends on the conditional distribution of υ^Q_{kt}, given $(\upsilon^{CPC}_{kt}, \upsilon^M_{kt}, \upsilon^R_{kt}, \upsilon^C_{kt})$. We define the following notations:

$$(\upsilon^Q_{kt}, \upsilon^{CPC}_{kt}, \upsilon^M_{kt}, \upsilon^R_{kt}, \upsilon^C_{kt})' \sim \text{MVN}\left(0, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right), \quad \text{(A.1)}$$

where $\Sigma_{11} = \Phi_{11}$, $\Sigma_{12} = (\Phi_{12}, \Phi_{13}, \Phi_{14}, \Phi_{15},)$, $\Sigma_{21} = \Sigma'_{12}$,

$$\Sigma_{22} = \begin{bmatrix} \Phi_{22} & \cdots & \Phi_{25} \\ \vdots & \ddots & \vdots \\ \Phi_{52} & \cdots & \Phi_{55} \end{bmatrix}.$$

Then, we have

$$\upsilon^Q_{kt} \mid \upsilon^{CPC}_{kt}, \upsilon^M_{kt}, \upsilon^R_{kt}, \upsilon^C_{kt} \sim N(\bar{\mu}, \bar{\sigma}^2), \quad \text{(A.2)}$$

where $\bar{\mu} = \Sigma_{12}\Sigma_{22}^{-1}(\upsilon^{CPC}_{kt}, \upsilon^M_{kt}, \upsilon^R_{kt}, \upsilon^C_{kt})'$ and $\bar{\sigma}^2 = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. We assume the prior distribution of β^Q as the following:

$$\beta^Q \sim \text{MVN}(\beta_0, \Sigma_0). \quad \text{(A.3)}$$

Then $\beta^Q \mid Click, X^Q, \eta^Q, \bar{\mu}, \bar{\sigma}^2 \sim \text{MVN}(A, B)$, where $\beta_0 = 0$, $\Sigma_0 = 100I$, $B = [X^{Q'}X^Q\bar{\sigma}^2 + \Sigma_0^{-1}]^{-1}$, and $A = B[X^{Q'}(\ln(Click) - \eta^Q - \bar{\mu}) + \Sigma_0^{-1}\beta_0]$.

2. Draw η^Q_k.

We update η^Q_k for each keyword k, respectively. The conditional distribution of η^Q_k given η^{CPC}_k is the following:

$$\eta^Q_k \mid \eta^{CPC}_k \sim N(\hat{\mu}, \hat{\sigma}^2), \quad \text{(A.4)}$$

where $\hat{\mu} = \Omega_{12}\Omega_{22}^{-1}\eta^{CPC}_k$ and $\hat{\sigma}^2 = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$. We keep the notations $\bar{\mu}, \bar{\sigma}^2$ defined in Step 1. The posterior distribution of η^Q_k is

$$\eta^Q_k \mid Click, X^Q, \beta^Q, \eta^{CPC}_k, \hat{\mu}, \hat{\sigma}^2, \bar{\mu}, \bar{\sigma}^2 \sim N(A, B), \quad \text{(A.5)}$$

where $B = (S\hat{\sigma}^2 + \bar{\sigma}^2)^{-1}\hat{\sigma}^2\bar{\sigma}^2$ and $A = B[(\sum_{t=1}^{S}(\ln(Click_{kt}) - \beta^Q X_{kt} - \bar{\mu}))/\bar{\sigma}^2 + \hat{\mu}/\hat{\sigma}^2]$.

3. Draw Φ.

We assume that the prior of Φ follows an inverted-Wishart distribution:

$$\Phi \sim \text{IW}(w_0, W_0). \quad \text{(A.6)}$$

Then the posterior of Φ can be shown as follows:

$$\Phi \mid \upsilon^Q_{kt}, \upsilon^{CPC}_{kt}, \upsilon^M_{kt}, \upsilon^R_{kt}, \upsilon^C_{kt}, w_0, W_0 \sim \text{IW}(w_0 + S \cdot K, W_0 + \upsilon\upsilon'), \quad \text{(A.7)}$$

where $\upsilon = (\upsilon^Q, \upsilon^{CPC}, \upsilon^M, \upsilon^R, \upsilon^C)$, S is the number of time periods, K is the number of keywords, $w_0 = 10$, and $W_0 = 10I$.

4. Draw Ω.

We assume that the prior of Ω follows an inverted-Wishart distribution:

$$\Omega \sim \text{IW}(w_0, W_0). \quad \text{(A.8)}$$

Then the posterior can be shown as follows:

$$\Omega \mid \eta^Q_k, \eta^{CPC}_k, w_0, W_0 \sim \text{IW}(w_0 + K, W_0 + \eta\eta'), \quad \text{(A.9)}$$

where $\eta = (\eta^Q, \eta^{CPC})$, $w_0 = 10$, and $W_0 = 10I$.

5. Draw β^{CPC} as in Step 1.
6. Draw η^{CPC}_k as in Step 2.
7. Draw β^M as in Step 1.
8. Draw β^R as in Step 1.
9. Draw β^S as in Step 1.

A.2. First-Step Estimation Algorithm for Keyword Entry and Consideration Models

We ran the MCMC chain for 20,000 iterations and use every 20th draw of the last 10,000 iterations to compute the mean and standard deviation of the posterior distribution of model parameters. Note that we control for the potential sample selection by allowing the unobserved error terms in ad ranking and keyword entry equations to be potentially correlated. Therefore, we need to jointly estimate the models of ad position, keyword entry, and keyword consideration in the first-step estimation. Following the previous literature (e.g., Erdem et al. 2008), we set the initial value of $Elag_{ki0}$ as $Elag_{ki0} = \sum_{t=1}^{S} a_{kit}/S \cdot (1-\rho)$, where S is the total number of periods. The initial values of $RlagInv$ and $Nlag$ are defined in a similar way.

We use a flexible probit specification to describe an advertiser's keyword entry decision, which is expressed as follows:

$$\begin{aligned} U_{kit} &= \gamma_i + \gamma_{1T} X_k + \gamma_{3T} Match_{ki} + \gamma_{3T}\ln(SV_{kt}) \\ &\quad + \gamma_{4T}\ln(Click_{k,t-1}) + \gamma_{5T}\ln(CPC_{k,t-1}) + \gamma_{6T} Z_{kit} \\ &\quad + \sum_{l=0}^{t-1} \gamma_{7lT} a_{kil} + \sum_{l=0}^{t-1} \gamma_{8lT} \ln(RankInv_{kil}) \\ &\quad + \sum_{l=0}^{t-1}\left(\sum_{T'=\{M,R,C\}} \gamma_{9TT'} \bar{a}_{k,-i,l}(T', c_{kt})\right) \\ &\quad + \sum_{l=0}^{t-1}\left(\sum_{T'=\{M,R,C\}} \gamma_{10TT'} \ln\left[\overline{RankInv}_{k,-i,l}(T', c_{kt})\right]\right) \\ &\quad + \xi_t^U + \varsigma_{kit}^U, \end{aligned} \tag{A.10}$$

$$\bar{a}_{k,-i,l}(T, c_{kt}) = \frac{\sum_{j\neq i,\, j\in T} a_{kjl} c_{kjt}}{\sum_{j\neq i,\, j\in T} c_{kjt}}, \tag{A.11}$$

$$\overline{RankInv}_{k,-i,l}(T, c_{kt}) = \frac{\sum_{j\neq i,\, j\in T} RankInv_{kjl} c_{kjt}}{\sum_{j\neq i,\, j\in T} c_{kjt}}, \tag{A.12}$$

where we allow advertiser i's payoff to depend on three sets of variables: (i) keyword characteristics, (ii) the focal advertiser's entry and rank history, and (iii) the average entry and average rank of potential entrants from each type in each of the previous periods. We also include advertiser-specific and time-specific fixed effects. Let c_{kt} denote the vector of advertisers' consideration decisions, i.e., $c_{kt} = (c_{k1t}, \ldots, c_{kIt})$. Then $\bar{a}_{k,-i,l}(T, c_{kt})$ and $\overline{RankInv}_{k,-i,l}(T, c_{kt})$ stand for the average entry frequency and average inverse of rank of type T potential entrants in period l.

For brevity, we simplify the notations by rewriting the utility of ad position, keyword entry, and keyword consideration as the following:

$$R_{kjt} = \alpha X_{kjt}^R + \varsigma_{kjt}^R, \tag{A.13}$$

$$U_{kit} = \gamma X_{kit}^U + \varsigma_{kit}^U, \tag{A.14}$$

$$V_{kit} = \lambda X_{kit}^V + \varsigma_{kit}^V, \tag{A.15}$$

$$(\varsigma_{kjt}^R, \varsigma_{kit}^U, \varsigma_{kit}^V) \sim \mathrm{MVN}\left(0, \begin{bmatrix} 1 & \delta & 0 \\ \delta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right), \tag{A.16}$$

$$c_{kit} = 1\{V_{kit} > 0\}, \tag{A.17}$$

$$a_{kit} = c_{kit} \cdot 1\{U_{kit} > 0\}, \tag{A.18}$$

$$X_{kit}^U = (X_{kit}^{U1}, X_{kit}^{U2}(c_{kt})), \tag{A.19}$$

where $X_{kit}^{U2}(c_{kt})$ refer to those covariates in Equation (A.10) that depend on the vector of firm's consideration decisions. Since c_{kit} is an indicator function of V_{kit}, we can also express $X_{kit}^{U2}(c_{kt})$ as $X_{kit}^{U2}(V_{kt})$, where V_{kt} is a vector of V_{kit}. Next we describe the MCMC algorithm for the first-step estimation.

1. Draw R_{kjt}.

For expositional convenience, we index advertisers in a keyword by their ranks (i.e., R_{kjt} stands for the weighted bid of the advertiser at the jth position in keyword k at time t). Given ς_{kjt}^U and δ, the conditional distribution of ς_{kjt}^R is the following:

$$\varsigma_{kjt}^R \mid \varsigma_{kjt}^U, \delta \sim N(\delta\varsigma_{kjt}^U, 1-\delta^2). \tag{A.20}$$

We can then describe the posterior distribution of augmented variables R as follows:

$$\Pr(R \mid Rank) \propto \sum_{t=1}^{S}\sum_{k=1}^{K}\left(1\{R_{k1t} > R_{k2t} > \cdots > R_{kN_{kt}t}\} \cdot \prod_{j=1}^{N_{kt}} \exp\left[-\frac{(R_{kjt} - \bar{R}_{kjt} - \delta\varsigma_{kjt}^U)^2}{2(1-\delta^2)}\right]\right), \tag{A.21}$$

where N_{kt} stands for the number of entrant advertisers, $1\{R_{k1t} > R_{k2t} > \cdots > R_{kN_{kt}t}\}$ is an indicator function that equals one if the vector of weighted bids is in a descending order, and $\bar{R}_{kjt} = \alpha X_{kjt}^R$.

The posterior conditional distribution of R_{kjt} given $(R_{k1t}, \ldots, R_{k,j-1,t}, R_{k,j+1,t}, \ldots, R_{kN_{kt}t})$, denoted by $R_{k,-j,t}$, is the truncated normal distribution, expressed as follows:

$$R_{kjt} \mid R_{k,-j,t} \sim \begin{cases} N(\bar{R}_{kjt} + \delta\varsigma_{kjt}^U, 1-\delta^2) \times I(R_{k,j+1,t}, +\infty) \\ \quad \text{if } j = 1, \\ N(\bar{R}_{kjt} + \delta\varsigma_{kjt}^U, 1-\delta^2) \times I(R_{k,j+1,t}, R_{k,j-1,t}) \\ \quad \text{if } 2 \le j \le N_{kt} - 1, \\ N(\bar{R}_{kjt} + \delta\varsigma_{kjt}^U, 1-\delta^2) \times I(-\infty, R_{k,j-1,t}) \\ \quad \text{if } j = N_{kt} - 1. \end{cases} \tag{A.22}$$

2. Draw α.

We assume the prior distribution of α as

$$\alpha \sim \mathrm{MVN}(\alpha_0, \Sigma_0). \tag{A.23}$$

Then, the posterior distribution of α is also normal

$$\alpha \mid R, X^R \sim \mathrm{MVN}(A, B), \tag{A.24}$$

where $B = (1-\delta^2)[X^{R'}X^R + \Sigma_0^{-1}]^{-1}$, $A = B[X^{R'}(R - \delta\varsigma_{kjt}^U) + \Sigma_0^{-1}\alpha_0]$, $\alpha_0 = 0$, and $\Sigma_0 = 100I$.

3. Draw U_{kit}.

The conditional distribution of U_{kit} given ς_{kjt}^R and δ is the following:

$$U_{kit} \mid \varsigma_{kjt}^R, \delta \sim N(\gamma X_{kit}^U + \delta\varsigma_{kjt}^R, 1-\delta^2). \tag{A.25}$$

We update U_{kit} given a_{kit} and c_{kit} as the following:

$$U_{kit} \mid a_{kit}, c_{kit} \sim \begin{cases} N(\gamma X_{kit}^U + \delta\varsigma_{kjt}^R, 1-\delta^2) \times I(0, +\infty) \\ \quad \text{if } a_{kit} = 1, \\ N(\gamma\gamma X_{kit}^U + \delta\varsigma_{kjt}^R, 1-\delta^2) \times I(-\infty, 0) \\ \quad \text{if } a_{kit} = 0 \text{ and } c_{kit} = 1, \\ N(\gamma X_{kit}^U + \delta\varsigma_{kjt}^R, 1-\delta^2) \\ \quad \text{if } a_{kit} = 0 \text{ and } c_{kit} = 0. \end{cases} \tag{A.26}$$

4. Draw V_{kit} and update c_{kit} and X^U_{kit}.

We use the Metropolis–Hastings algorithm with a random walk chain to generate draws (see Chib and Greenberg 1995). Let $V^{(p)}_{kit}$ denote the previous draw; then the next draw $V^{(p)}_{kit}$ is given by the following:

$$V^{(n)}_{kit} = V^{(p)}_{kit} + \Delta, \tag{A.27}$$

with the accepting probability given by

$$\text{Prob} = \begin{cases} \min\left\{\dfrac{\exp[-(V^{(n)}_{kit} - \lambda X^V_{kit})^2/2]1\{V^{(n)}_{kit} > 0\}}{\exp[-(V^{(p)}_{kit} - \lambda X^V_{kit})^2/2]}, 1\right\} \text{ if } a_{kit} = 1, \\ \min\left\{\dfrac{\exp[-(V^{(n)}_{kit} - \lambda X^V_{kit})^2/2]}{\exp[-(V^{(p)}_{kit} - \lambda X^V_{kit})^2/2]}, 1\right\} \\ \qquad\qquad \text{if } a_{kit} = 0 \text{ and } V^{(n)}_{kit} \cdot V^{(p)}_{kit} > 0, \\ \min\{(\exp[-(V^{(n)}_{kit} - \lambda X^V_{kit})^2/2 - (U_{kit} - \gamma X^U_{kit}(V^{(n)}_{kit}) \\ \quad - \delta\varsigma^R_{kjt})^2/2(1-\delta^2)]) \cdot (\exp[-(V^{(p)}_{kit} - \lambda X^V_{kit})^2/2 \\ \quad - (U_{kit} - \gamma X^U_{kit}(V^{(p)}_{kit}) - \delta\varsigma^R_{kjt})^2/2(1-\delta^2)])^{-1}, 1\} \\ \qquad\qquad \text{if } a_{kit} = 0 \text{ and } V^{(n)}_{kit} \cdot V^{(p)}_{kit} \le 0, \end{cases} \tag{A.28}$$

where Δ is a draw from the density $N(0, 2.25)$. Given an accepted V_{kit}, we update c_{kit} and X^U_{kit} accordingly.

5. Draw δ.

We also use a Metropolis–Hastings algorithm with a random walk chain to generate draws for δ. We first transform δ to $\mu = \log(\delta/(1-\delta))$ so that μ has an infinite support. Let $\mu^{(p)}$ and $\delta^{(p)}$ denote the previous draws, then the new draws of μ and δ are given by the following:

$$\mu^{(n)} = \mu^{(p)} + \Delta, \tag{A.29}$$

$$\delta^{(n)} = \frac{\exp(\mu^{(n)})}{1 + \exp(\mu^{(n)})}, \tag{A.30}$$

with the accepting probability given by

$$\begin{aligned} \text{Prob} &= \min\left\{\left(\prod_{k,i,t}\left[(1-\delta^{(n)2})^{-1/2}(\varsigma^R_{kjt}, \varsigma^U_{kjt})\begin{bmatrix}1 & \delta^{(n)} \\ \delta^{(n)} & 1\end{bmatrix}(\varsigma^R_{kjt}, \varsigma^U_{kjt})'\right]\right)\right. \\ &\quad \left. \cdot \left(\prod_{k,i,t}\left[(1-\delta^{(p)2})^{-1/2}(\varsigma^R_{kjt}, \varsigma^U_{kjt})\begin{bmatrix}1 & \delta^{(p)} \\ \delta^{(p)} & 1\end{bmatrix}(\varsigma^R_{kjt}, \varsigma^U_{kjt})'\right]\right)^{-1}, 1\right\}, \end{aligned} \tag{A.31}$$

where $\varsigma^R_{kjt} = R_{kit} - \alpha X^R_{kit}$, $\varsigma^U_{kjt} = U_{kit} - \gamma X^U_{kit}$, and Δ is a draw from the density $N(0, 0.004)$.

6. Draw γ as in Step 2.
7. Draw λ as in Step 2.

A.3. Second-Step Estimation Algorithm for Keyword Entry and Consideration Models

Given the MCMC draws from the first step, we use every 20th draw of the last 10,000 of 20,000 draws from the posterior distributions to predict $\hat{P}^*(a_{kit} \mid I_{kt})$ and to construct the expected number of entrants ranked above $\hat{E}(N^a_{kit}(I_{kt}))$ and ranked below $\hat{E}(N^b_{kit}(I_{kt}))$, both of which appear on the right-hand side of entry utility in Equation (13). Furthermore, the expected average click volume and average CPC depend on the expected entry probability of firms and therefore are functions of the vector of $\hat{P}^*(a_{kit} \mid I_{kt})$. We plug the covariates of $\hat{E}(N^a_{kit}(I_{kt}))$, $\hat{E}(N^b_{kit}(I_{kt}))$, $\hat{E}(Click_{kt} \mid I_{kt})$, and $\hat{E}(CPC_{kt} \mid I_{kt})$, which are predicted from the first-step estimation, into the equation of entry payoff. We then reestimate the keyword entry and consideration equations jointly in a similar fashion as the algorithm outlined in Section A.2.

A.4. Simulation Studies

We perform a simulation exercise with two objectives. First, we want to test the theoretical identification of our proposed joint model of entry and consideration. Second, we hope to demonstrate the potential bias in the estimate of strategic interaction effect when researchers incorrectly regard potential entrants as the entire set of players.

We consider a simplified version of our entry and consideration model by assuming that there are three players who might be interested in entering a market. All three firms have a probability of being a potential entrant in the consideration stage. After that, the firms who are potential entrants compete in a simultaneous-move entry game. If only one firm considers entry, the game is reduced to an individual decision-making problem. We include an intercept and two covariates (X_k, X_{ki}) in the entry utility, where X_{ki} serves as the exclusion restrictions. As for the consideration intensity, we include a new covariate, W_{ki}, in addition to the intercept to warrant the model identification. We generate X_k, X_{ki}, and W_{ki} from an i.i.d. standard normal distribution. We also allow a firm's entry utility to be affected by the entry of the number of competitors when more than two firms are considering market entry.

We simulate the data for 1,000 markets and estimate the model with two different approaches. We first estimate the model using the proposed two-step method outlined in Section 5.2. Second, we estimate the model with a standard two-step method without the consideration stage. Specifically, we assume that all three firms are potential entrants and use $(1, X_k, X_{ki}, W_{ki})$ as covariates in the specification of entry payoff. We report two sets of simulation results for the model with either positive or negative strategic interaction effect.

The estimation results reported in Table A.1 suggest that all parameters in our models of keyword entry and keyword consideration are identifiable and can be recovered within a 95% posterior confidence level. Furthermore, we find that the estimate on the strategic interaction effect incurs severe *downward bias* if researchers ignore the consideration decision and treat the entire set of players as potential entrants. The reasoning for such a downward bias can be briefly described as follows. Let us assume the entry utility of a firm is $U_{ki} = \beta X_{ki} + \gamma \sum_{j \ne i} P(a_j = 1) + \epsilon_{ki}$, where X_{ki} includes the exclusion restriction and ϵ_{ki} stands for the private profit shock. Then, when the consideration stage is not accounted for, the $\hat{P}(a_j = 1)$ predicted from the first-step estimation in a two-step method will be biased upward. Therefore, the magnitude of the estimated strategic interaction effect γ obtained from the second step will tend to be smaller than its true value.

Table A.1. Simulation Results of Entry and Consideration Models

Sign of strategic interaction effect	Negative		Positive	
	True value	Estimates	True value	Estimates
With consideration (true model)				
β_1 (Consideration)	1, 1	**1.008**, **1.006** (0.103) (0.082)	1, 1	**0.994**, **0.953** (0.041) (0.047)
β_2 (Entry)	1, 1, 1	**1.050**, **1.054**, **0.966** (0.113) (0.067) (0.056)	1, 1, 1	**0.981**, **1.062**, **1.048** (0.162) (0.093) (0.087)
γ (Spillover)	−1	**−1.042** (0.106)	1	**1.171** (0.121)
Without consideration				
β_3 (Entry)	N.A.	**0.250** (0.115) **0.387** (0.046) **0.687** (0.060) **0.599** (0.043)	N.A.	0.039 (0.100) **0.635** (0.031) **0.265** (0.029) **0.256** (0.027)
γ (Spillover)	−1	**−0.580** (0.127)	1	**0.404** (0.072)

Endnotes

[1] Please see the description of the probit entry model in Section 5.3. The model using log-transformed numbers of entrants as covariates (log marginal = −69,640) outperforms the model with nontransformed covariates (log marginal = −70,290).

[2] Since the first-step estimation results are used to approximate equilibrium entry probabilities, we report the estimation results from only the second step in this paper.

References

Abhishek V, Hosanagar K (2007) Keyword generation for search engine advertising using semantic similarity between terms. Gini M, Kauffman RJ, eds. *Proc. Ninth Internat. Conf. Electronic Commerce* (ACM, New York), 89–94.

Agarwal A, Hosanagar K, Smith MD (2011) Location, location and location: An analysis of profitability of position in online advertising markets. *J. Marketing Res.* 48(6):1057–1073.

Aguirregabiria V, Mira P (2007) Sequential estimation of dynamic discrete games. *Econometrica* 75(1):1–53.

Andrews RL, Srinivasan TC (1995) Studying consideration effects in empirical choice models using scanner panel data. *J. Marketing Res.* 32(1):30–41.

Athey S, Nekipelov D (2012) A structural model of sponsored search advertising auctions. Working paper, Harvard University, Cambridge, MA.

Bajari P, Hong H, Krainer J, Nekipelov D (2010) Estimating static models of strategic interactions. *J. Bus. Econom. Statist.* 28(4): 469–482.

Berry ST (1992) Estimation of a model of entry in the airline industry. *Econometrica* 60(4):889–917.

Bronnenberg BJ, Vanhonacker WR (1996) Limited choice sets, local price response, and implied measures of price competition. *J. Marketing Res.* 33(2):163–174.

Chen Y, Xue GR, Yu Y (2008) Advertising keyword suggestion based on concept hierarchy. Najork M, ed. *Proc.* 2008 *Internat. Conf. Web Search Data Mining* (ACM, New York), 251–260.

Chiang J, Chib S, Narasimhan C (1999) Markov Chain Monte Carlo and models of consideration set and parameter heterogeneity. *J. Econometrics* 9(1–2):223–248.

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

Ciliberto F, Tamer E (2009) Market structure and multiple equilibria in airline markets. *Econometrica* 77(6):1791–1828.

Datta S, Sudhir K (2013) Does reducing spatial differentiation increase product differentiation? Effects of zoning on retail entry and format variety. *Quant. Marketing Econom.* 11(1):83–116.

Desai PS, Shin W, Staelin R (2014) The company that you keep: When to buy a competitor's keyword. *Marketing Sci.* 33(4):485–508.

Dover Y, Neslin S (2015) Amplifying the impact of digital advertising: The role of published sales rankings. Working paper, Dartmouth College, Hanover, NH.

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

Ellickson PB, Misra S (2011) Estimating discrete games. *Marketing Sci.* 30(6):997–1010.

Erdem T, Keane MP, Sun B (2008) A dynamic model of brand choice when price and advertising signal product quality. *Marketing Sci.* 27(6):1111–1125.

Fuxman A, Tsaparas P, Achan K, Agrawal R (2008) Using the wisdom of the crowds for keyword generation. *Proc.* 17*th Internat. Conf. World Wide Web* (ACM, New York), 61–70.

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

Gilbride TJ, Allenby GM (2004) A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* 23(3): 391–406.

Goldfarb A, Tucker C (2011) Advertising bans and the substitutability of online and offline advertising. *J. Marketing Res.* 48(2):207–227.

Hausman JA (1996) Valuation of new goods under perfect and imperfect competition. Bresnahan TF, Gordon RJ, eds. *The Economics of New Goods* (University of Chicago Press, Chicago), 207–248.

Hotz VJ, Miller RA (1993) Conditional choice probabilities and the estimation of dynamic models. *Rev. Econom. Stud.* 60(3):497–529.

Interactive Advertising Bureau (2015) IAB Internet advertising revenue report. New York, http://www.iab.com/wp-content/uploads/2016/04/IAB_Internet_Advertising_Revenue_Report_FY_2015-final.pdf.

Jerath K, Ma L, Park YH (2014) Consumer click behavior at a search engine: The role of keyword popularity. *J. Marketing Res.* 51(4): 480–486.

Jia P (2008) What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica* 76(6):1263–1316.

Joo M, Wilbur KC, Cowgill B, Zhu Y (2014) Television advertising and online search. *Management Sci.* 60(1):56–73.

Kakulapati V, Kolikipogu R, Revathy P, Karunanithi D (2011) Improved web search engine by new similarity measures. Abraham A, Mauri JL, Buford JF, Suzuki J, Thampi SM, eds. *Advances in Computing and Communications*, Comm. Comput. Inform. Sci., Vol. 193 (Springer, Berlin Heidelberg), 284–292.

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

Lamey L, Deleersnyder B, Steenkamp JBEM, Dekimpe MG (2012) The effect of business-cycle fluctuations on private-label share: What has marketing conduct got to do with it? *J. Marketing* 76(1):1–19.

Lee TY, Bradlow ET (2011) Automated marketing research using online customer reviews. *J. Marketing Res.* 48(5):881–894.

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

Nevo A (2001) Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69(2):307–342.

Radlinski F, Broder A, Ciccolo P, Gabrilovich E, Josifovski V, Riedel L (2008) Optimizing relevance and revenue in ad search: A query substitution approach. Chua T-S, Leong M-K, eds. *Proc. 31st Annual Internat. ACM SIGIR Conf. Res. Development Inform. Retrieval* (ACM, New York), 403–410.

Rutz OJ, Bucklin RE (2011) From generic to branded: A model of spillover dynamics in paid search advertising. *J. Marketing Res.* 48(1): 87–102.

Sayedi A, Jerath K, Srinivasan K (2014) Competitive poaching in sponsored search advertising and its strategic impact on traditional advertising. *Marketing Sci.* 33(4):586–608.

Seim K (2006) An empirical model of firm entry with endogenous product-type choices. *RAND J. Econom.* 37(3):619–640.

Stokes R (2008) *Mastering Search Advertising: How the Top 3% of Search Advertisers Dominate Google AdWords* (iUniverse, Bloomington, IN).

Su CL, Judd KL (2012) Constrained optimization approaches to estimation of structural models. *Econometrica* 80(5):2213–2230.

Van Heerde HJ, Gijsenberg MJ, Dekimpe MG, Steenkamp J-BEM (2013) Price and advertising effectiveness over the business cycle. *J. Marketing Res.* 50(2):177–193.

van Nierop E, Bronnenberg B, Paap R, Wedel M, Franses PH (2010) Retrieving unobserved consideration sets from household panel data. *J. Marketing Res.* 47(1):63–74.

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

Vitorino MA (2012) Empirical entry games with complementarities: An application to the shopping center industry. *J. Marketing Res.* 49(2):175–191.

Yang S, Ghose A (2010) Analyzing the relationship between organic and sponsored search advertising: Positive, negative or zero interdependence? *Marketing Sci.* 29(4):602–623.

Yang S, Lu S, Lu X (2014) Modeling competition and its impact on paid-search advertising. *Marketing Sci.* 33(1):134–153.

Yao S, Mela CF (2011) A dynamic model of sponsored search advertising. *Marketing Sci.* 30(3):447–468.

Zhu T, Singh V (2009) Spatial competition with endogenous location choices: An application to discount retailing. *Quant. Marketing Econom.* 7(1):1–35.