Association for Information Systems
AIS Electronic Library (AISeL)

2010

ONLINE RETAIL KEYWORD CHARACTERISTICS AND SEARCH MARKETING PERFORMANCE

Han-fen Hu
The University of Utah, han-fen.hu@business.utah.edu

Olivia R. Liu Sheng
The University of Utah, olivia.sheng@business.utah.edu

Follow this and additional works at: http://aisel.aisnet.org/icis2010_submissions

ONLINE RETAIL KEYWORD CHARACTERISTICS AND SEARCH MARKETING PERFORMANCE

Completed Research Paper

Han-fen Hu
Department of Operations and Information Systems
David Eccles School of Business
The University of Utah
1645 E Campus Center Dr., Salt Lake City
UT 84112
han-fen.hu@business.utah.edu

Olivia R. Liu Sheng
Department of Operations and Information Systems
David Eccles School of Business
The University of Utah
1645 E Campus Center Dr., Salt Lake City
UT 84112
olivia.sheng@business.utah.edu

Abstract

To make judicious budget and resource allocation decisions, it is essential to understand the characteristics and performance of keywords used by target shoppers. Taking the approach of understanding consumers' information needs, we construct a goal-related keyword characterization framework. Most search keywords exhibit more than one of the following characteristics - retailer specific, brand specific, product specific, feature related, or shopping intention. We analyze search engine ranking, click-through, and revenue data associated with visitor-disclosed search keywords from a top Web-only retailer over one year. Our findings show the interesting impacts of both keyword characteristics and dual-appearance of paid and organic advertisements on search marketing performance. The contributions of the research include a comprehensive search keyword characterization framework, and the analysis of the relationships between keyword characteristics and search performance. These findings have strong implications for search marketing decisions.

Keywords: search keywords, keyword characteristics, keyword selections, paid and organic search engine rankings, search marketing revenue, click-throughs, regression models, lift analysis

Introduction

Due to the potential of reaching an increasingly large number of online consumers via search engines, search marketing has become the most popular type of online marketing. Search marketing that promotes a retailer's website in the search result listings of a search engine involves optimizing two types of search engine rankings: paid and organic search result ranks. Paid or sponsored search listings are advertising creatives displayed in a pre-specified region of a search result page. The search engine charges a placement fee when a search user clicks on a paid ad to visit an advertiser's website (i.e. a click-through). The placement fee, measured by cost-per-click (CPC) and the number of click-throughs the advertisement receives, is primarily determined by auction and the rank position (e.g., rank or position 1, rank or position 2, etc.) of a paid ad in a keyword's paid listing (Xing and Lin 2006). On the other hand, organic search result listings present websites that a search engine discovers using its native matching algorithm. Organic listings are free of charge, and the rankings on the result page reveal the relevancy of a website to searchers' queries.

The search marketing performance of an e-tailer depends on the effectiveness of reaching and attracting search engine users who are likely to visit and transact at the website. To reach a search engine user, the e-tailer's website needs to be "visible" in the top positions of a search keyword's result listings. In order to increase click-throughs and sales via visitors from search engines, it is essential to further analyze customers' behavior and preferences. In this context, search keyword analysis is considered instrumental, because search keywords can reveal a broad variety of user preferences and goals in only a few words (Chau et al. 2005; Mehta et al. 2007; Rose and Levinson 2004). Analyzing the relationships between search marketing performance and keywords with characteristics representative of user goals is essential for performance-based keyword selection decisions. Hence, a comprehensive user-goal framework for keyword characteristics is valuable to search marketing performance analysis.

Many advertisers invest in optimizing both paid and organic search result rankings as dual appearances of paid and organic search results on the same result page provides multiple exposures to a website. Marketing literature has shown that multiple exposures to an advertisement can enhance impression and increase conversion (McDonald 1996; Naples 1979; Von Gonten and Donius 1997). To make informed decisions on how to invest in paid and organic search marketing of the same keyword, it is important for e-tailers to discover the impact of dual search ad appearances on search marketing performance.

In this study, we are interested in the following questions related to search marketing strategies: What are the characteristics of the search keywords that help reveal user preferences or goals? What are the relationships amongst keyword characteristics, search result rankings, and search marketing performance? Do search engine users respond differently toward paid ads and organic listings? What is the impact of dual appearances of organic and paid search listings on search marketing performance? To address these questions, we construct a goal-related keyword characterization framework from previous research, with added aspects appropriate for an online retailer. A unique focus of the study is to analyze the keyword characteristics by consumers' information needs and loyalty orientation reflected in the keywords they enter (i.e. user-disclosed keywords). Over a one-year period, organic and paid search ranking data were extracted from Google on a daily basis for 4,000 user-disclosed keywords relevant to a top, Web-only, e-tailer. This data was then combined with keyword-based weekly visit and purchase information. To evaluate the impact of keyword characteristics and search result rankings, we built and tested regression models to evaluate their relationships to search marketing outcomes. To the best of our knowledge, our study is the first to cover a longitudinal time span and a large keyword set with detailed keyword characteristics and fine-grain search organic and paid search rank data. To evaluate the impact of using our regression models for keyword selections, we analyze the lift in click-throughs and revenue generated by keywords suggested by the regression models. Both the keyword characteristic framework and the empirical models can provide foundations for continuing studies on search keyword and marketing performance analysis.

Effects of Organic and Paid Search Advertisements: Related Research and Gap

In this section, we summarize the findings of prior studies on analyzing the relationships between search keywords and search marketing performance to identify the gap this study intends to fill.

Ghose and Yang (2008; 2009a) analyze the impact of retailer-specific keywords and brand-specific keywords on paid search marketing performance. They found that these two types of keywords show different influences on click-through rates, conversion rates, cost-per-click, search result rank, order values and profits. Many advertisers

note diverse characteristics of search keywords, and are interested in optimizing their search marketing performance by using a wide range of keyword characteristics. However, past findings have not expanded the potential to monetize on keywords with more characteristics. By codifying consumer queries into five categories, including intent to buy, product specificity, location specificity, company specificity and general information, Jansen (2007) found that a large number of search queries fell into the product-specific category (48.34%). However, Jansen's (2007) study did not examine the impact of product and other keyword characteristics on search marketing outcomes. Further studies are necessary to examine the relationships of these types of keyword characteristics to search marketing performance. In order to better characterize search keywords and analyze search keyword performance, we focus on keyword characteristics that help reveal user preferences or goals.

Website visibility based on search result rank positions and counts also greatly influence the performance of search marketing (Sen 2005) by inducing more visits to the website. Several studies are interested in the effect of search result rankings on search marketing performance (e.g. Animesh et al. 2007; Agarwal et al. 2008; Ghose and Yang 2009a; 2009b). Ghose and Yang (2009a) find that sponsored keywords on the top of search result pages are associated with high click-through rates and conversion rates. Agarwal et al. (2008) also provide empirical evidences that while click-through rates of keywords decrease with keywords' paid search rankings, their conversion rates first increase and then decrease with paid search rankings. Animesh et al. (2007) found that result rankings of sponsored ads are positively associated with website visits, customer ratings and website inlinks; however, the relationships are more salient for vendors of search goods (i.e. low quality uncertainty goods), compared to that for experience and credence goods (i.e. high quality uncertainty goods). However, most of prior studies only focus on the rankings and performance of sponsored keywords.

A recent survey conducted by the Internet Retailer Association (Siwicki 2009) on the top 211 Web-only retailers, chain retailers, catalogers and consumer brand manufacturers indicates that 55.3% intend to increase spending on search engine optimization (SEO) for natural (organic) search. To obtain insights into optimizing organic search keyword performance, it is important to further examine the characteristics, rankings and performance of keywords in organic search listings. Some studies (Jansen 2007; Xing and Lin 2006) have indicated that search engine users prefer organic results because they are considered more objective and unbiased than sponsored results. On the other hand, some empirical evidences have shown that the conversion rates, order values and profits from paid search marketing are higher than those from organic natural search optimization (Ghose and Yang, 2008). Ghose and Yang (2009b) also model and estimate the inter-relationship between organic search listings and paid search advertisements. Their findings conclude that the respective probabilities to click on the advertiser's paid and organic listings in the same search result page are positively interdependent, and the positive interdependence leads to an increase in expected profits for the e-tailer However, past results did not address the impact of dual search ad appearances on visits and purchases via organic searches or paid searches.

Past studies use datasets that spanned up to 3 months to analyze search keyword performance (e.g., Ghose and Yang 2009a; 2009b). Since marketing campaigns of an e-tailer can vary a lot for seasonal events, and online customers change their search goals according to holidays or other time-sensitive needs, it is valuable to track the keyword performance over a one-year period and control for time effects in analyzing search marketing performance. Past studies primarily focus on websites that transact via both off-line and on-line channels. Their findings need further validations in order to be applicable to Web-only retailers. Since Web-only e-tailers are a fast growing segment of Internet Retailers (Siwicki, 2009), research based on data from Web-only e-tailers would have a profound impact on future research and practices of search marketing analysis.

The Conceptual Model

In order to address the identified research gap, we propose a research model including the two critical aspects of search marketing (as shown in Figure 1): (1) keyword characteristics of search user queries that reveal target consumers' information needs and online search goals, and (2) search result visibility that includes result rankings and dual appearances of paid and organic results on the same page.

Figure 1. Conceptual Model
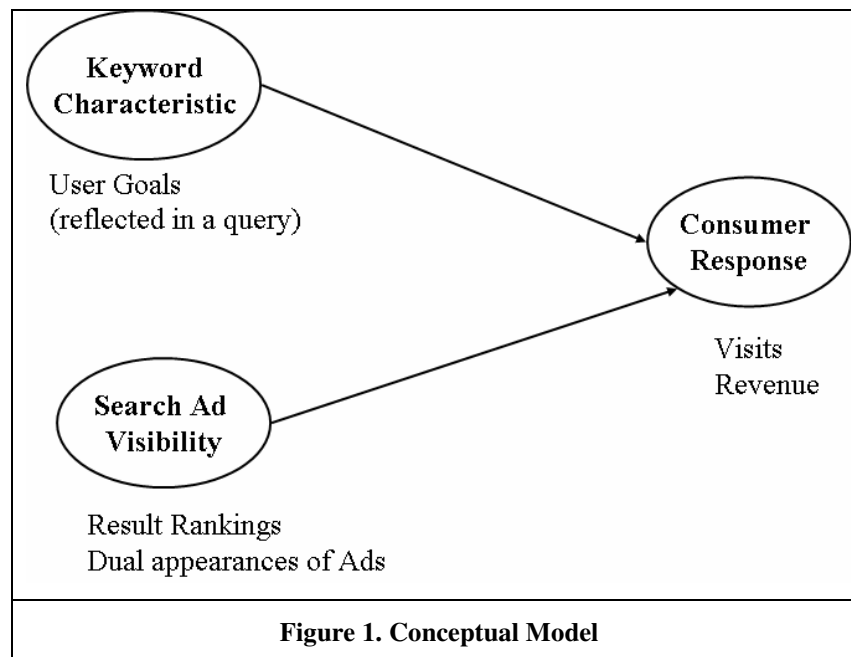
One of the most important motives of online users is information search for better shopping decisions, and therefore Web searches reflect a diverse set of underlying user goals (Rose and Levinson 2004). This functionalist approach allows researchers to articulate Internet users' motives, and to conceptualize how users perceive and process interactive advertisements (Rodgers and Thorson 2000). Such understandings help to develop more sophisticated models and make better predictions about users' Web-related attitudes and behaviors (Rodgers and Thorson 2000). Prior research in search engine logs also focuses on how users use search engines on the Web to satisfy their information needs (Chau et al. 2005). Based on the analysis of website search logs or other activity logs, several schemes of Web query classification have been proposed. Some examples are goal-oriented classification (i.e. navigational, informational, and transactional by Broder 2000) and shopping-related search classification (i.e., prepurchase and ongoing search by Bloch et al 1986). According to these prevalent frameworks, users with different search strategies are motivated by disparate goals and affected by dissimilar determinants.

Along this line, we construct a goal-related framework of search keyword classification based on previous research with added aspects that fit an online retailer's setting. This framework fully covers our data set and can be used in the keyword analysis for general online retailing websites. Most search keywords[1] submitted by users exhibit more than one characteristic. For example, the search keyword - "buy unlocked cell phones" contains product and feature related information, and the word – "buy" indicates shopping intention. Therefore, it would be more useful to codify search keywords with different characteristics rather than to categorize them into disjoint categories. We include the following characteristics in this framework based on past literature and common search marketing practitioners' considerations:

- Retailer specific: Keywords referencing the name of an online retailer, for example, "RetailerName", Retailer Name, RetailerName.com (Ghose and Yang 2009a) are often submitted by searchers who are familiar with, loyal to or interested in the retailer .

- Brand specific: Keywords containing a brand name, either a product specific or a company specific (Ghose and Yang 2009a; Rutz and Bucklin 2008), such as Nike, Wii, etc. Brand names become more important online than in the traditional market (Degeratu et al. 2000), serving as quality or style cues for products.

---

[1] A keyword may not be a single word; in stead, most of the keywords are combinations of multiple words. Search keywords are referenced interchangeably with search terms, search queries and search phrases in literature and practice.

◦ Product specific: Keywords containing specific product terms (Jansen 2007), for example, clothing or bed, that indicate searchers' product-related information needs.

◦ Feature specific: Keywords used to specify certain attributes of products, including material, version, color, shape, size, style or the context; for instance, red, limited, Halloween, etc. In online consumer search, desired features of the products are also used as a way to narrow down choice sets and search results.

◦ Shopping intention: Keywords directly related to purchase behaviors. Some examples are buying and purchase (Jansen 2007), and promotional and discount related terms (e.g., coupon, sale, 20% off, free shipping, etc.). Consumers using such search keywords may have strong shopping intentions.

An advertiser can target consumers with different information needs based on the characteristics of the search keywords they enter in search engines by designing search advertisements related to the consumers search interests.

Another aspect of search marketing we would like to investigate is the visibility of an e-tailer in search result listings. The visibility of a website on the search result page is measured based on its position on the paid and organic listings for a given keyword, as well as the likelihood that the e-tailer appears in both paid and organic listings. It is commonly assumed that the search results at the top positions would attract more consumer attention. In addition, dual appearances of paid and organic ads of a website are considered a critical factor in search ad visibility. As both organic search result and advertisement space is limited, simultaneous exposure in both areas allows an e-tailer to substantially increase their website's prominence on the result page.

In this study, we focus on click-throughs (i.e. visits) and revenue that reveal "user reactions" or "behaviors that could impact a website's bottom-line" to evaluate the performance of search marketing. Click-through is a quantity of interest to search engine marketers as a metric for evaluating advertising effectiveness (Regelson and Fain 2006). Click-throughs from sponsored keyword impressions also have a direct impact on the cost of sponsored ads. Some researchers use profitability to assesses the performance of different search engine marketing strategies (e.g., Sen 2005). However, the cost of paid and organic search ads are not directly comparable and even may not be precisely measurable. Hence, we use revenue, which is product sales without shipping and handling fees generated from the visits associated with the keywords, as another search marketing metric. Some of the past studies (Ghose and Yang 2008; 2009a; 2009b; Agarwal 2008) measure search marketing performance by click-through rates and conversion rates. "Click-through rate" is the percent of searches using a keyword that actually lead to click-throughs, whereas "conversion rate" is the percent of click-throughs via a keyword actually resulting in purchases. As an e-tailer does not know the actual number of organic search result impressions in which its website has appeared, organic click-through rates tend to be estimated based on keyword search volumes made available by search engines. As the accuracy of click-through rate estimates is unknown, this study adopts click-throughs instead. To take into consideration the effect of popular keywords which are subject to high competition, we include cost-per-click (CPC) which is a surrogate for keyword popularity as a control variable in our models. Conversion rates of visits can be directly measured in our data set. However, we found that models based on revenue have higher explanation power than models based on conversion rates for our data. To account for the effect of product price on revenue, we include the average revenue per order as another control variable in our models. Prior studies (Ghose and Yang 2008; 2009a; 2009b; Agarwal 2008) have included the length of paid ads as an independent variable. However, we focus on user-specified keywords, rather than paid ads. The length of search keyword may represent the specificity of the consumer's information needs. Therefore, the length of search keywords is included as a control variable.

Hypotheses

Based on theories on customer loyalty, advertisement competition, and consumer attention, we have developed hypotheses regarding influences of keyword characteristics and website visibility in search results on search marketing performance.

A searcher who includes an e-tailer's name in search keywords may be a loyal customer of the e-tailer, may have strong interests in the e-tailer's products or may be familiar with the e-tailer. This familiarity with or interest in the e-tailer is likely to prompt the searcher to click on the e-tailer's organic search result or paid ad, rather than selecting a competitor's website. In addition, familiarity with a website can help the searcher to subjectively reduce uncertainty, simplify the interactions, build trust in the e-tailer, and then decide to transact (Gefen 2000). Therefore, we expect retailer-specific keywords are correlated to high click-throughs and revenue for the e-tailer. Ghose and Yang (2009a) confirmed that retailer-specific sponsored keywords are correlated to higher click-through rate and conversion rates. We, hypothesize that these positive relationships between retailer-specific keywords and search

marketing outcomes also should prevail for click-throughs and revenue generated via both organic and sponsored search listings. We hence test the following hypotheses:

H1: Retailer-specific search keywords generate more click-throughs and higher revenue than non-retailer-specific keywords.

> *H1a: Retailer-specific search keywords generate more click-throughs than non-retailer-specific keywords.*

> *H1b: Retailer-specific search keywords generate higher revenue than non-retailer-specific keywords.*

Search keywords related to a specific brand name are likely to come from more brand-loyal consumers who are driven by brand awareness (Ghose and Yang 2009a). The goal of search engine users who submit such queries tends to be collecting information about products by brand which can be sold by a number of different e-tailers. Branded keywords do not suggest searchers' preferences for a specific retailer, and hence are also subject to steep competition on the same head keywords with other retailers for search visibility (Green, 2003). Because of consumers' brand loyalty as well as the low search and switching cost on the Internet, branded keywords may not be correlated to high click-throughs for any one e-tailer. Ghose and Yang (2009a) find that the presence of brand-specific information in a paid search advertisement is associated with a decrease in its click-through and conversion rates. We believe the competitive conditions would exist in both sponsored and organic search result listings. Therefore, we hypothesize that branded keywords would be associated with lower click-throughs and revenue generated via both sponsored advertisements and organic listings in the following hypotheses:

H2: Brand-specific search keywords generate fewer click-throughs and lower revenue than non-brand-specific keywords.

> *H2a: Brand-specific search keywords generate fewer click-throughs than non-brand-specific keywords.*

> *H2b: Brand-specific search keywords generate lower revenue than non-brand-specific keywords.*

A consumer often uses some product words (e.g. TV, bedding, etc.) to reflect his/her explicit needs to collect information about products. Therefore, product related search keywords may lead to high click-throughs and revenue. In fact, queries containing product-related keywords are the most popular types of queries submitted to search engines, accounting for more than 45% of the queries according to Jansen (2007). In this vein, while the targeted consumers' interest and the keywords they may use are difficult to predict, an e-tailer may have a better chance to meet the searcher's information needs if it targets product-specific keywords on search engines. The enormous search volume on product-related keywords trumps the increased competition, and all visible websites can receive many click-throughs. Although product-specific keywords may induce fierce competitions, a website can still be rewarded with a great number of click-throughs and high revenue if its link and ads are easily visible (e.g., in top positions) relative to its competitors in the result page. We hence test the following hypotheses:

H3: Product-specific search keywords generate more click-throughs and higher revenue than non-product-specific keywords.

> *H3a: Product-specific search keywords generate more click-throughs than non-product-specific keywords.*

> *H3b: Product-specific search keywords generate higher revenue than non-product-specific keywords.*

Past studies have shown that the search results at the top positions would attract more consumer attention, and therefore have high click-throughs. Search results rankings are also perceived as a reflection of the relevancy of the Web pages to searchers' queries, and therefore result links with top rankings are more likely to be visited by searchers. Prior studies on search marketing have also highlighted the effect of rankings on search performance (Ghose and Yang 2009a; 2009b; Agarwal et al. 2008). However, most of prior studies only examine the rankings

and performance of sponsored keywords. We would like to further conjecture the same relationship would also exist in organic search results, and be even more salient than that in the paid search listings. According to the cognitive load theory, the cognitive capacity of a human mind is limited (Sweller 1988), and therefore the amount of information a searcher can process during the information search process is also limited. Along this line, the number of search result links a searcher may read is finite. For the same query, the number of organic search results is often much greater than that of paid search listings. Hence, browsing organic search listings and browsing paid search listings impose different cognitive loads on searchers. By definition, high click-through rate does not equal high click-throughs. Similarly, high conversion rate does not equal high revenue. Thus, the relationship between rankings of paid ads and their performance indicated by prior research needs to be validated in this study. We also further hypothesize the negative association between rankings of organic listings and their performance. We test the following hypotheses:

H4: The search engine result ranking of an e-tailer's website for a search keyword is negatively associated with the click-throughs and revenue associated with the keyword.

> *H4a: The search engine result ranking of an e-tailer's website for a search keyword is negatively associated with the click-throughs associated with the keyword.*
>
> *H4b: The search engine result ranking of an e-tailer's website for a search keyword is negatively associated with the revenue associated with the keyword.*

On a search result page, paid and organic listings are shown side-by-side. Searchers can freely choose any result links from the two listings. However, studies have shown that organic results are more appealing to searchers (Xing and Lin, 2006; Jansen, 2007) and organic listings are chosen first by 70% of the searchers viewing search results (SEMPO survey, 2004). When a website's link also appears in the paid listings, it may help to make a deeper impression on consumers. The creative content of an e-tailer's paid ad may also provide additional information that is not shown in the organic counterpart, and therefore induce click-throughs of the organic result. On the other hand, when an e-tailer's link is shown amongst the top in the organic listings, it may make the e-tailer's ads appear more non-biased and objective (Xing and Lin, 2006), and improve the likelihood of the paid advertisement being visited.

Advertising literature also suggests that multiple exposures to an advertisement can lead to better results (McDonald 1996; Von Gonten and Donius 1997), as the repeated exposure can enhance the advertised messages and e-tailer images. Jansen and Spink (2007) have suggested that searchers are more likely to recall the name of a company from a search listing if it appears multiple times in search listings on the same page. In our context, the dual-appearance of paid and organic search results on the same page can lead to increased click-throughs and revenue via either paid or organic listings. Thus, we hypothesize that the dual-appearance of an e-tailer's search advertisement in both organic listing and paid listing on the same search result page of a keyword can increase click-throughs as well as revenue.

H5: Dual appearances of an e-tailer's search advertisement in both organic and paid search result listings on the same search result page of a keyword increase its click-throughs and revenue.

> *H5a: Dual appearances of an e-tailer's search advertisement in both organic and paid search result listings on the same search result page of a keyword increase its click-throughs.*
>
> *H5b: Dual appearances of an e-tailer's search advertisement in both organic and paid result listings on the same search result page of a keyword increase its revenue.*

Data

Similar to past studies on search marketing (e.g. Agarwal et al. 2008; Ghose and Yang 2008; 2009a; 2009b), we test the hypotheses on data collected from a single site. Our focal company is a full-service Web-only e-tailer that has been ranked amongst the top 20 e-tailers in the last three years according to the surveys by Internet Retailer (Siwicki

2010). An organic or sponsored referral[2] keyword is identified in the query used by the search engine user who clicked on the landing page link in the e-tailer's ad in the search results. These keywords disclose website visitors' goals and preferences. Our visitor-disclosed keywords consist of 2000 randomly sampled sponsored referral keywords and 2000 randomly sampled organic referral keywords during the month prior to the start of the 52-week data collection period. The size of each keyword set is slightly less than 1% of the total keywords from the same source.

We manually codified the characteristics of the search keywords. Two trained codifiers reviewed the keywords, and labeled their keyword characteristics according to a given guideline that described the characterization framework and definition of each characteristic. Their labeling results were then compared and validated by the authors in case of any discrepancy.

We obtained weekly aggregates of click-throughs and revenue from the Web Analytics software used by our focal e-tailer to select visitor-disclosed keywords. Table 1 summarizes the paid and organic search performance of the 4000 keywords over 52 weeks. The data indicates that our study's focal e-tailer has been more successful with search engine optimization to improve marketing performance of organic keywords than with PPC keyword bidding to generate customer responses to paid listings. It is worth noting that this performance trend is different from that reported in some of the prior studies (e.g., Ghose and Yang 2008; 2009a) on a mixed online player. Specifically, their keywords' paid listings outperformed the organic listings.

Table 1. Performance of visitor-disclosed keywords

		Maximum	Mean	Std. Deviation	t-test result
Click-throughs	Paid Search Keywords	4547	18.43	113.80	< 0.001
	Organic Search Keywords	14348	31.87	275.44	
Revenue	Paid Search Keywords	8499	21.26	154.75	< 0.001
	Organic Search Keywords	38938	46.97	415.24	

An e-tailer can track search rankings of keywords in the paid listings in Google Analytics (http://google.com/analytics/) or by a paid advertising agency. Popular Web analytics typically don't track or report organic search engine ranks because of the complexity of tracking such data. We submitted the 4000 search keywords using our own Web mining software to gather the ranks they received from Google[3] in the organic and the paid search listings daily over the 52-week period. We retrieved up to the top 30 paid and organic search results respectively. If the actual search ranks are greater than 30, or some unknown search engine server changes or network issues occurred, our software returned null values for the daily search ranks. Properly handling null values and the derivation of the weekly ranks are critical to our empirical analysis. If the keywords on average had been ranked amongst top 25 and then suddenly dropped out of the top 30 search results for only one or two days, we took a sliding window (five days) average to replace the missing (null-value) rank. If the keywords on average ranked between top 26 and top 30, we use the five-day average plus 3 times of the five-day variance to replace the missing rank[4], as we assume the search phrase just slightly fell out of the top 30 listing. We derive weekly ranks from our detailed daily rank data for our empirical analysis. By taking the mode (the value that occurs the most frequently) of the ranks for each week to generate the weekly ranks, we obtained weekly rankings for each keyword. Most researchers and agents use programs to crawl search result pages once per week, and apply the one-day rank

---

[2] A search referral is a user visit originates from sponsored or organic search results.

[3] According to the data collected by the Web Analytics used by the focal e-tailer, the click-throughs coming from Google account for 77.85% of the click-throughs from search engines, and the revenue generated from visits via Google accounts 73.97% of the revenue from visits via search engines. Therefore, we use the rankings on Google result pages to represent the visibility of the e-tailer's website on the search result page.

[4] We found that average plus 3 times standard deviation covers 95% of the data without missing rankings, and therefore considered 3 times standard deviation can give use a good augmentation of the abnormally missed rankings.

snapshot as the weekly rankings for the search keywords. This method is stated to be reasonable in estimating keyword rankings because rankings do not vary much within a week (Ghose and Yang, 2009b); however, it can be biased when the rank variations within a week are significant. The alternative rank metrics we propose in this study - the weekly average search ranking based on seven daily rankings is considered especially valuable for the vendors who advertise on keywords in high competition situations and hence may experience high rank variations. With normal web browser settings, the first five search results can be seen without scrolling and can be read in a single glance, and are considered to have comparable attractiveness to the users. We re-scale the top 5 rankings as rank class 1, the top 6 to 10 rankings as rank class 2, and so on. The missing rank class is scaled as 10. Re-scaled rank classes allow our models to streamline the relationships between search ranks and search marketing outcomes.

Table 2 shows the distribution of the search keyword characteristics and ranks. About 67% of the organic keywords' daily ranks are among the top 30, which is about 7 % more than that of paid keywords ranked among the top 30. Furthermore, organic keywords are more likely to be ranked in the top half of the first listing page than paid keywords. 84.55% of the paid search keywords contain product related information, followed by 51.90% of paid keywords with feature related information. Only 7.00% of paid search keywords are related to retailer information. Similar but slightly lower distributions are observed in organic search keywords as well. It is interesting to note that 31% of organic keywords include brand related information, which is higher than the percent of branded paid search keywords by 5%. Most of the search keywords are of multiple characteristics. Only 24.60% of the paid search phrases and 23.36% of organic search phrases are of single characteristic keywords. This pattern is similar to the distribution reported in Jenson's (2007) study, and echoes the need to further examine the overlap among characteristics to understand the intent of searchers (Jensen 2007).

Table 2. Keyword statistics by keyword characteristics and ranks					
		Paid Search Keywords		Organic Search Keywords	
		# of keywords	%	# of keywords	%
Retailer		140	7.00	116	5.80
Brand		520	26.00	620	31.00
Product		1691	84.55	1612	80.64
Feature		1038	51.90	1016	50.80
Shopping		296	14.80	259	12.95
Single Characteristic		492	24.60	467	23.36
Rank Class	1	23034	22.18	37159	35.78
	2	24160	23.27	14805	14.26
	3	14905	14.35	7263	6.99
	4	226	0.22	4990	4.81
	5	6	0.01	3572	3.44
	6			2179	2.10
	7			134	0.13
	10	41253	39.73	33741	32.49

Regression Models and Results

We use regression models to examine the effects of search keyword characteristics, search advertisement visibility, and performance metrics. We use six dummy variables corresponding to (1) the five keyword characteristics in the conceptual model, and (2) whether an e-tailer's search advertisement in organic listing and paid listing dual-appear on the same search result page of a keyword. The rank class is included in the model as ordinal numbers. We also include three additional control variables, average revenue per order, the length of search queries entered by users, and the cost-per-click (CPC) of the keyword in paid placement. The CPC of a keyword represents the competition condition in paid ads auction, and is also considered a proxy of the competition for keyword visibility in search

marketing. The time variables representing the week number are also included in our models. Because of the high variance in click-throughs and revenue among keywords, we adopt log-transformations of these performance metrics for dependent variables as exemplified in past related studies (Agarwal et al. 2008; Chau et al. 2005; Ghose and Yang 2009a; Jansen et al. 2000; Regelson and Fain 2006; Spink et al. 2001). Ordinary least square (OLS) regression models with time fixed effect (i.e. include the week numbers – e.g., week 1 or 2 in the model as a fixed effect controlling for the seasonal fluctuation) were constructed to evaluate the relationships between keyword characteristics and performance. Table 3 shows the results from the log-transformation regression models.

Table 3. Log-transformation regression on click-throughs and revenue rate				
	Paid	Organic	Paid	Organic
Dependent Variable	log(Click-throughs)		log(revenue)	
Model	1	2	3	4
R^2	0.494	0.345	0.487	0.513
Adjusted R^2	0.489	0.340	0.482	0.509
(Constant)	1.917 **	1.853 **	1.349 **	1.843 **
log (Click-throughs)			0.243 **	0.206 **
Retailer	0.228 **	0.451 **	0.143 **	0.038
Brand	0.083 *	-0.410 **	-0.012	-0.169 **
Product	0.066	0.381 **	0.073 **	0.015
Feature	-0.369 **	-0.337 **	0.020	0.002
Shopping	-0.257 **	0.075	0.096 **	-0.028
Rank	-0.003	-0.024 **	0.001	-0.007 **
Dual-appearance	0.046 *	0.008	0.038 **	0.045 **
CPC	0.070 **	0.113 **	0.062 **	0.067 **
Average revenue per order	0.001 **	0.001 **	0.002 **	0.002 **
Length	-0.137 **	-0.400 **	0.013	-0.086 **

**: $p < 0.01$; *: $p < 0.05$.

Model 1 and Model 2. We first report interesting relationships between keyword characteristics and click-throughs revealed in the results. The log-transformed regression model for click-throughs generated by all of the paid listings explains a substantial amount of the variances (49.4%) in the click-throughs, and predicts better than the organic counter part (34.5%). All of the three control variables – CPC, average revenue per order and keyword length have significant relationships with click-throughs via sponsored or organic paid keywords. It is interesting to note that both the positive effect of CPC and the negative effect of keyword length on organic keyword click-throughs are stronger than those on sponsored keyword click-throughs. Hence, it is important to consider their effects when examining the relationships between click-throughs and the main constructs. CPC is a surrogate for both keyword competition and popularity. High keyword competition could adversely impact click-throughs, and yet high keyword popularity implies more searches which could increase click-throughs. With the aid of good organic search visibility obtained by getting an e-tailer's website listed in the top organic search results, the harvest from a keyword's popularity outweighs the adverse impact of fierce competition.

Retailer-specific keywords have significant positive relationships with click-throughs via paid and organic listings. *Brand-specific* keywords have significant negative relationships with click-throughs via organic listings but not for paid listings. One reason could be that the CPCs and hence competitions for such paid branded keywords (mean= 0.105) are lower than the CPCs for organic branded keywords (mean=0.169; t = 12.828, p<0.0001). Since the competition condition for sponsored branded keywords is not as fierce, the proposed negative relationship caused by the competition can be not as salient. *Product-specific* keywords have significant positive relationships with click-throughs via organic listings, but not with click-throughs via paid ads. A likely cause is the high search volumes

combined with good search ranks of organic product-specific keywords. These results also exhibit different consumer reactions toward paid and organic search listings of branded or product keywords. In terms of result visibility, our results show that keywords for which the e-tailer is ranked amongst the top (i.e., highly visible) in organic search listings can generate more click-throughs than organic keywords for which the e-tailer is less visible. This confirms the conventional wisdom about the importance of utilizing search engine optimization to improve organic search ranks. However, the effect is not significant for paid ads. The coefficient for *Dual-appearance* is significant and positive for paid results only, but not for organic listings. The asymmetric effect of the dual-appearance of paid and organic results (i.e. only the click-through via paid results increases when the two types of results appear on the same page) could be caused by the difference in cognitive load between browsing in the paid search result listings and browsing in the organic listings.

Table 4. Summary of hypotheses testing result.		Paid Search Keywords	Organic Search Keywords
	Hypotheses	Paid Search Keywords	Organic Search Keywords
H1a	Retailer-specific search keywords generate more click-throughs than non-retailer-specific keywords.	Supported	Supported
H1b	Retailer-specific search keywords generate higher revenue than non-retailer-specific keywords.	Supported	Not Supported
H2a	Brand-specific search keywords generate fewer click-throughs than non-brand-specific keywords.	Not Supported	Supported
H2b	Brand-specific search keywords generate lower revenue than non-brand-specific keywords.	Not Supported	Supported
H3a	Product-specific search keywords generate more click-throughs than non-product-specific keywords	Not Supported	Supported
H3b	Product-specific search keywords generate higher revenue than non-product-specific keywords.	Supported	Not Supported
H4a	The search engine result ranking of an e-tailer's website for a search keyword is negatively associated with the click-throughs associated with the keyword.	Not Supported	Supported
H4b	The search engine result ranking of an e-tailer's website for a search keyword is negatively associated with the revenue associated with the keyword.	Not Supported	Supported
H5a	Dual appearances of an e-tailer's search advertisement in both organic and paid search result listings on the same search result page of a keyword increase its click-throughs	Supported	Not Supported
H5b	Dual appearances of an e-tailer's search advertisement in both organic and paid result listings on the same search result page of a keyword increase its revenue.	Supported	Supported

Model 3 and Model 4. We found that the models on revenue have more explanation power than that of the log-transformed click-through counterparts as indicated by their R-squares. The effects of the control variables are similar to those in Models 1 and 2.

Retailer-specific keywords have significant positive relationships with revenue from paid ads, but not with revenue via organic listings. One reason could be that the e-tailer's competitors rarely place paid search ads associated with our focal e-tailer's name or URL. Thus, consumers visiting the e-tailer via paid search ads are not distracted by the competitors' ads, and therefore more likely to transact on the e-tailer's website. On the other hand, the search engine would provide a large set of organic search results based on its algorithm that contains result links similar to the e-tailer's website (i.e., similar pages). Consumers' attention can be distracted, and may even discover the e-tailer's competitors from organic search results. *Brand-specific* keywords have significant, negative relationships with revenue from organic listings, but not with revenue from paid listings. This can be also attributed to the less steep

competition condition in paid branded keywords, as compared to organic branded keywords. *Product-specific* keywords have significant positive relationships with revenue from paid ads, but not organic listings. One reason could be that the CPCs and hence competition for such paid product-specific keywords (mean= 0.293) are higher than the CPCs for organic product-specific keywords (mean=0.250; t = 10.163, p<0.0001). Since the search volume for organic product-specific keywords is not as high, the proposed positive relationship therefore is not significant. Dual-appearance is significant for revenue from both organic and paid listings. The multiple exposures of search results not only attract consumers to click on the result links, but also induce the consumer to transact on the website. Table 4 summarizes the hypotheses testing results.

Performance Lift Analysis

Lift analysis is popular in the marketing area to evaluate the usefulness of predictive analytics such as regression models (Hughes 2000; Kim et al. 2005; Padmanabhan et al. 2006). For this analysis, we use the regression models to estimate the performance of the 2000 paid keywords and the 2000 organic keywords in the first 26 weeks, and then rank order the keywords by their estimated click-throughs and conversion rates. We then calculate the average performance of keywords in the last 26 weeks and compare it to the performance of the keywords if randomly selected. Figure 2 shows the performance curves (diamonds for selected keywords and squares for random keywords) from these comparisons.

The results imply that if the e-tailer invests in the first 50 paid keywords selected by our model instead of random choices, the number of average visits per keyword increases by 21.1 times, while the revenue per keyword increases by 32.4 times. Similarly, the first 50 organic keywords selected by our model instead of random choices, the number of average visits per keyword increases by 74.2 times, while the revenue per keyword increases by 173.1 times. This is indicative of the high payoff of keyword performance predictive models.

Figure 2. Lift Analysis of Click-throughs and Revenue of Suggested Keywords
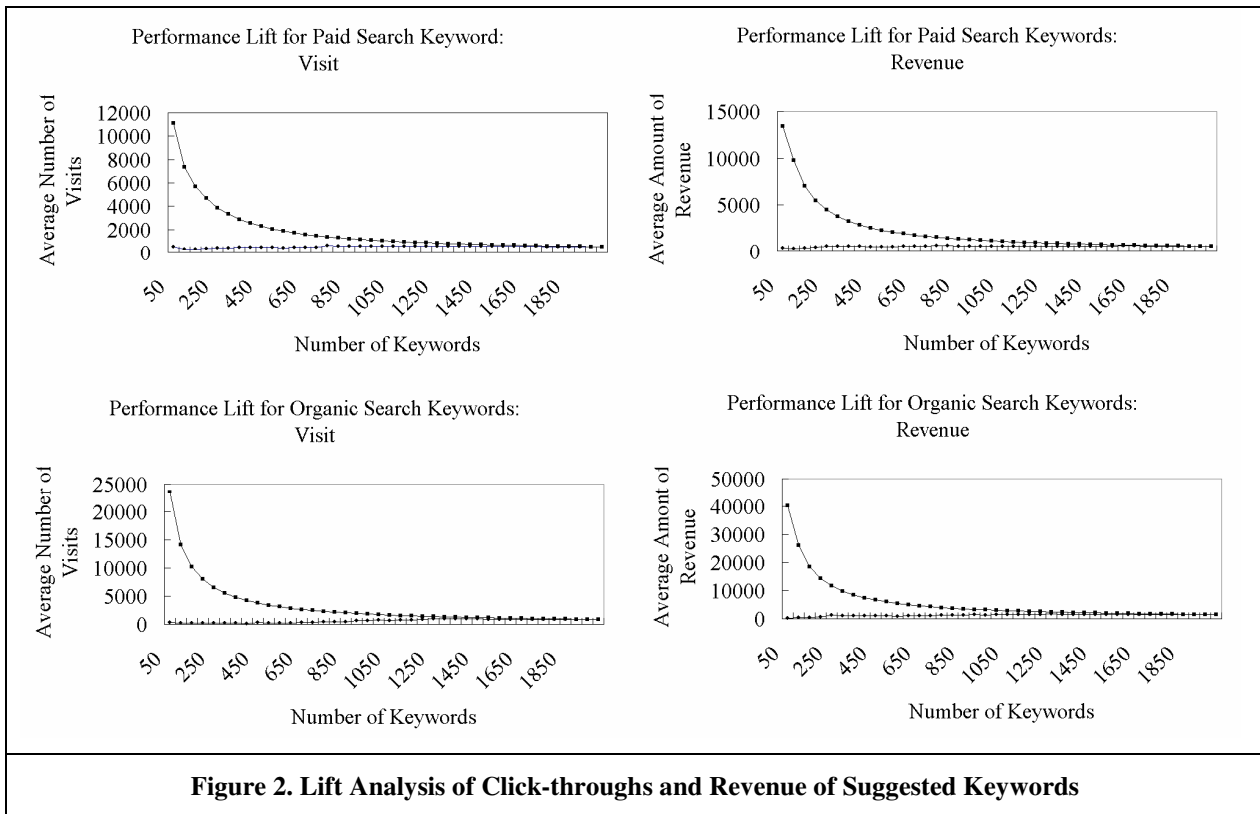
Conclusion

Data driven research on search marketing strategies is an emerging research area. Motivated by the research gap identified, we propose a broad set of goal-oriented keyword characteristics that can offer analyses on complex relationships between keyword characteristics and performance. Because multi-channel and Web-only e-tailers have different online marketing strategies and Web-only e-tailers are among the fastest growing categories of Internet retailers, we perform hypothesis testing and regression analysis using data gathered on a top Web-only full-service e-tailer. Our major findings include (1) in general, retailer specific keywords improves performance of paid ads, as well as organic listings; (2) keywords with brand and product information improve performance of paid ads, but hurt the performance of organic listings; (3) Keywords with high CPC (i.e., the competition in PPC is high) are associated with high performance; (4) organic search results shown on the top of the result page increase click-throughs and conversion rate, but this effect is not seen with paid ads; and (5) when search results of the e-tailer appear in both paid ads and organic listings on the same page, the search word performs better.

In turn, this study makes several research contributions. First, we propose a comprehensive search keyword characterization framework and the analysis of the relationships between keyword characteristics and search performance. This framework sufficiently covers all the search keywords entered by searchers, and therefore enables the advertisers to comprehensively monetize on keywords. Second, the visibility metrics proposed in this study (i.e. consider both rankings and dual-appearance of paid and organic listings on the same result page) provide further insights into the need for advertisers to invest in both paid and organic search marketing. A larger variety of keyword characteristics and the visibility metrics of search advertisement also afford us the opportunities to analyze increasingly complex relationships keywords have with search performance. Third, this study responds to the importance of analyzing and comparing search marketing performance of both paid ads and organic search results. To extend our understanding on the characteristics, rank and performance of keywords in organic search listings, we analyze, track the performance of organic search results and compare it with paid ads. Fourth, whereas most previous search marketing research concentrates on advertiser specified keywords in paid search advertising, we target visitor-disclosed keywords to advance our understanding of the users' actual information needs, and how search engine advertisements capitalize on various types of users' information needs. The use of visitor-disclosed keywords also makes the resulting implementations sustainable.

Based on longitudinal keyword performance data, our findings have strong implications for future research and practices. The keyword selection for studies on search keyword performance is critical. It is important to use a repeatable, systematic and scalable approach intuitive to practitioners to select keywords as done in this study. We focus on the search goals disclosed in the visitors' keywords from both the organic and sponsored search listings. As such, the analysis can provide broader insights into their characteristics and effects on search marketing performance. Our models can help e-tailers discover effective optimization search marketing strategies. If an e-tailer invests in the keywords suggested by our model, the performance of paid search words could increase by 21.1 times in number of visits, and 32.4 times in revenue over randomly selected keywords; similarly, the performance of organic search words could increase by 74.2 times in number of visits, and 173.1 times in revenue. Our lift analysis also shows that this approach can generate the most profitable result when a small set of keywords are targeted (i.e. 5% to 10% of the training dataset). For advertisers who have limited marketing budget and want to maximize returns on investments in search marketing, this study provides an approach that is easy to implement, with easy to interpret results that facilitate marketing decision making in keyword selection.

Although organic search results are free of charge and are preferred by most online users, advertisers should invest in both paid ads and organic search listing, in order to achieve the optimal search marketing performance. As our empirical results show, the dual-appearance of paid ads and organic listings on the same result page benefits the performance of both paid ads and organic listings, with positive influences seen in both higher click-throughs and increased revenue.

In addition, prior research has indicated that the optimal advertisement strategies for offline and online stores are not the same (Zhang and Wedel 2009), and the marketing strategies for multichannel retailers have a different effect when compared with conventional single-channel retailing (Zhang 2009). Along this line, the online advertisement strategies for mix-play retailers may have to also align with the offline advertising campaign, while the offline promotion and sales may also impact the performance of search engine advertisements. The insights from prior research using offline or multichannel advertiser data may not be as applicable to Web-only e-tailers. Web-only e-tailers are the fastest growing category compared to their retail chain, cataloger, and direct manufacturer counter parts generating almost $37 billion Web sales according to a recent survey by the Internet Retailer Association

(Brohan 2009). The insights from analyzing the relationships between keyword characteristics and performance for a top Web-only of full-service e-tailer can be of timely importance.

Our study is limited in several ways. First, we primarily focused on the predictive abilities of keyword characteristics on keyword performance in this study. Ghose and Yang's (2009a) study shows that the landing page quality also has an impact on the search result performance. The purpose of this study is to first understand the impact and the explanation powers of the search phrase characteristics without the landing page factors. In future studies, the landing page factors as well as the listing factors (e.g. the title, URL, and summary of each result links) on the search result page can be included to understand the effects and the potential interactions with the search phrase characteristics. Second, click-through rate is widely used in the research in the online advertisement area. The panel data we used in this study do not include the number of impression of each search phrase, and therefore we could not include the click-through rate as one of the performance metrics in our models. Third, our lift analysis mainly is used to validate the usefulness of the predictive model, and therefore we compare the predictive results with randomly selected keywords. In the future, if we can compare our predictive model with the selection method currently used by the e-tailer in its search marketing, the value of the predictive model can be further verified. Fourth, we use the rankings on Google search result pages as a proxy of an e-tailer's visibility in search results. Future efforts in consolidating rankings on different search engine result pages are needed to provide a precise visibility metric. In addition, our findings are based on the single dataset used in this study. Further testing and validation using data from websites in other categories can help to generalize the findings of this study.

References

Agarwal, A., Hosanagar, K., and Smith, M.D. "Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets," Working paper, available at SSRN: http://ssrn.com/abstract=1151537, 26 November 2008.

Animesh, A., Ramachandran, V., and Viswanathan, S. "An Empirical Investigation of the Performance of Online Sponsored Search Markets," in *Proceedings of the 9th International Conference on Electronic commerce*, Minneapolis, Minnesota, 19-22 August 2007, pp. 153-160.

Borgs, C., Chayes, J., Etesami, O., Immorlica, N., Jain, K., and Mahadian, M. "Dynamics of Bid Optimization in Online Advertisement Auctions," in *Proceedings of the 16th International Conference on World Wide Web*, Banff, Alberta, Canada, 8-12 May 2007, pp. 153-160.

Brohan, M. "The Top 500 Guide Overview," in *Internet Retailer Top 5000 Guide*, 2009, pp. 6-16.

Chau, M., Fang, X., and Sheng, O.R.L. "Analysis of the Query Logs of a Web Site Search Engine," *Journal of the American Society for Information Science and Technology* (56:13), November 2005, pp. 1363-1376.

Degeratu, A.M., Rangaswamy, A., and Wu, J. "Consumer Choice Behavior in Online and Traditional Supermarkets: The Effects of Brand Name, Price, and Other Search Attributes," *International Journal of Research in Marketing* (17:1), March 2000, pp. 55-78.

Feldman, J., Muthukrishnan, S., Pal, M., and Stein, C. "Budget Optimization in Search-Based Advertising Auctions," in *Proceedings of the 8th ACM Conference on Electronic Commerce*, San Diego, California, 11-15 June 2007, pp. 40-49.

Gefen, D. "E-Commerce: The Role of Familiarity and Trust," *Omega*, (28:6), December 2000, pp. 725-737.

Ghose, A. and Yang, S. "Comparing Performance Metrics in Organic Search with Sponsored Search Advertising," in *Proceedings of the 2nd International Workshop on Data Mining and Audience Intelligence for Advertising*, Las Vegas, Nevada, 24 August 2008, pp. 18-26.

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

Ghose, A., and Yang, S. "Analyzing the Relationship between Organic and Sponsored Search Advertising: Positive, Negative or Zero Interdependence?" *Marketing Science* (55:10), October 2009, pp. 1605-1622.

Green, D.C. "Search Engine Marketing: Why it Benefits Us All," *Business Information Review* (20:4), December 2003, pp. 195-202.

Hughes, A.M. *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable Customer-Based Marketing Program*. McGraw-Hill, Columbus, Ohio, 2000.

Jansen, B., and Spink, A. "The Effect on Click-Through of Combining Sponsored and Non-Sponsored Search Engine Results in a Single Listing," in *Proceedings of the 2007 Workshop on Sponsored Search Auctions*, Banff, Alberta, Canada, 8-12 May 2007.

Jansen, B.J. "The Comparative Effectiveness of Sponsored and Nonsponsored Links for Web E-Commerce Queries," *ACM Transactions on the Web* (1:1), May 2007, pp. 1-25.

Jansen, B.J., Brown, A., and Resnick, M. "Factors Relating to the Decision to Click on a Sponsored link," *Decision Support Systems* (44:1), November 2007, pp. 46-59.

Jansen, B.J., Spink, A., and Saracevic, T. "Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web," *Information Processing and Management* (36:2), March 2000, pp. 207-227.

Kim, Y., Street, W.N., Russell, G., Menczer, G. "Customer Targeting: A Neural Network Approach Guided by Genetic Algorithms," *Management Science* (51:2), February 2005, pp. 264-276.

McDonald, C. *Advertising Reach and Frequency. Maximizing Advertising Results through Effective Frequency*, 2nd editon, Association of National Advertisers and NTC Business Books, New York and Lincolnwood, Illinois, 1996.

Mehta, A., Saberi, A., Vazirani, U., andVazirani, V. "AdWords and Generalized Online Matching," *Journal of the ACM* (54:5), 2007, Article No. 22.

Naples, M. *Effective Frequency: The Relationship between Frequency and Advertising Effectiveness*. Association of Nation Advertisers, New York, 1979.

Nicosia, F. *Consumer Decision Processes*, Prentice Hall, Englewood Cliffs, 1966.

Padmanabhan, B., Zheng, Z., Kimbrough, S.O. "An Empirical Analysis of the Value of Complete Information for eCRM Models," *MIS Quarterly* (30:2), June 2006, pp. 247-267.

Regelson, M., and Fain, D.C. "Predicting Click through Rate Using Keyword Clusters," in *Proceedings of the 2nd Workshop on Sponsored Search Auctions*, Ann Arbor, Michigan, 11 June 2006,

Rodgers, S., and Thorson, E. "The Interactive Advertising Model: How Users Perceive and Process Online Ads," *Journal of Interactive Advertising* (1:1), Fall 2000, pp. 42-61.

Rose, D.E., and Levinson, D. "Understanding User Goals in Web Search," in *Proceedings of the 13th International Conference on World Wide Web*, New York, 17-22 May, 2004.

Rutz, O.J. and Bucklin, R.E. "From Generic to Branded: A Model of Spillover Dynamics in Paid Search Advertising," Working paper, available at SSRN: http://ssrn.com/abstract=1024766, 8 May 2008.

Sen R. "Optimal Search Marketing Strategy," *International Journal of Electronic Commerce* (10:1), Fall 2005, pp. 9-25.

Siwicki, B. "Search Marketing, Especially Optimization, Is a Buoy in the Economic Storm," in *The Search Marketing Guide*, 2010 edition, Internet Retailer, Chicago, Illinois, 2009, pp. 29-32.

SEMPO.org. The state of search engine marketing 2004.

Spink, A., Wolfram, D., Jansen, B.J., and Saracevic, T. "Searching the Web: The Public and Their Queries," *Journal of the American Society for Information Science and Technology* (52:3), 2001, pp. 226-234.

Sweller, J. "Cognitive Load during Problem Solving: Effects on Learning," *Cognitive Science* (12:2), April-June 1988, pp. 257-285.

Von Gonten, M.F., and Donius, J.F. "Advertising Exposure and Advertising Effects: New Panel-Based Findings," *Journal of Advertising Research* (37:4), 1997, pp. 51-60.

Xing, B., and Lin, Z. "The Impact of Search Engine Optimization on Online Advertising Market," in *Proceedings of the 8th International Conference on Electronic Commerce*, Fredericton, New Brunswick, Canada, 14-16 August, 2006, pp. 519-529.

Zhang, J. and Wedel, M. "The Effectiveness of Customized Promotions in Online and Offline Stores," *Journal of Marketing Research* (46:2), April 2009, pp. 190-206.

Zhang, X. "Retailers' Multichannel and Price Advertising Strategies," *Marketing Science* (28:6), November 2009, pp. 1080-1094.