The Inorganic Side of Paid Search

Zhongming Ma
Computer Information Systems Department
College of Business Administration
California State Polytechnic University
zma@csupomona.edu

Gautam Pant
School of Accounting and Information Systems
David Eccles School of Business
The University of Utah
{gautam.pant, olivia.sheng}@business.utah.edu

Abstract

Search engines have become a de facto gateway for finding information on the Web. The reliability of results from popular search engines is hence critical as these search engines determine what information people access. Major search engines, such as Google and Yahoo, provide two types of results – organic and paid. While paid results have been an advertising boon for search engines, are they comparable to organic results in reliability? We explore this question with a data set for 44 different digital camera models while using third-party reliability ratings (satisfactory/unsatisfactory) for companies appearing in search engine results. We find that paid results contain more unsatisfactory companies than organic results and that the unsatisfactory companies are ranked higher in paid results than in organic results. In addition, unsatisfactory companies are ranked lower than satisfactory companies in organic results. However, among the paid results, unsatisfactory companies are not necessarily ranked higher than satisfactory companies. Hence, while it is more likely to find unsatisfactory companies in the paid results than in the organic results, they do not completely overshadow the satisfactory companies.

Keywords: Paid search, Organic Search, Reliability, Search Engine Marketing, Internet Pricing

1. Introduction

Due to the large size of the Web, search engines have become a necessary and convenient tool for people to find information and shop online. This need and convenience has elevated search engines to the center stage of the online world and they now play a critical role of being the de facto gateway to content on the Web. Major search engines, such as Google and Yahoo, display search results in two categories. One consists of organic results that are generated by search engine’s proprietary algorithms, and the other is paid or sponsored results that appear on the top, right, and sometimes bottom of a search engine result page (SERP). The paid results appear because of content providers/advertisers placing bids on one or more terms in the search query (Jansen 2006, 2007). Paid search has enormous economic impact and has become a primary business model for Web search engines (Jansen 2006). For example, it was an $8 billion industry in 2004 (McCarthy 2005) and grew to $12 billion a year after (Jansen 2007). In 2004, 99% of Google’s and 84% of Yahoo’s income came from advertising (McCarthy 2005).

Popular search engines do not directly profit from organic results and they regularly block spamming and some other inappropriate search engine optimization (SEO) techniques that are used to manipulate the rankings of organic results. Hence, users can assume a certain level of reliability from organic search results. Can the users expect the same level of reliability from paid results? Does the commercial nature of paid search make it “inorganic” (unreliable)? This
research question has serious implications for the search engine companies, search advertising market providers, and even more importantly the search engine consumers. Our current analysis for this research question is based on digital camera resellers who buy digital cameras from the manufacturers and sell them to consumers. Using the reliability report from Better Business Bureau (BBB) we classify a reseller as satisfactory or unsatisfactory. Given this classification we test a series of hypotheses related with the appearance of satisfactory and unsatisfactory companies in organic and sponsored results.

2. Literature Review
Given the time gap between placing an order online and receiving the order, Cao et al. (2003-2004) separate the customers’ satisfaction associated with placing order and that with fulfilling order. More specifically, they examine the relationships between pricing, customers’ satisfaction with price, satisfaction with the ordering process, and satisfaction with fulfillment. They find a negative relationship between price satisfaction and satisfaction with fulfillment process. This situation suggests higher price satisfaction by lowering price, probably at the cost of lower level of service in shipping and customer support. This lower level of service can be expected to negatively affect satisfaction with the fulfillment process. Anderson (1996) has shown that there is a positive relationship between customer satisfaction with the firm and price tolerance (i.e., maximum price to pay before switching). In addition, higher competition is associated with lower price tolerance.

Several researchers have investigated market efficiency (problems including price and price dispersion) of online market. For example, Bailey (1998) as well as Brynjolfsson and Smith (2000) examine market efficiency for books, CDs, and software sold online (e-retailers) and through conventional retailers. Tang and Lu (2001) compare pure e-retailers with hybrid retailers (selling through both websites and physical locations). Pan et al. (2003) study evolution of price over time and investigate how service attributes affect price dispersion in e-retailers.

Jansen and Resnick (2005) report that some users ignore sponsored results due to the lack of trust. After examining search results from three major search engines (Yahoo!, Google, and MSN), Jansen (2007) finds paid results are more relevant than organic results for e-commerce queries and the rank of paid results is not correlated with the relevance of those results. In this study we partition companies into two categories on the basis of third party (BBB) ratings on reliability. Unlike Jansen (2007), who measures the topical relevance of the results, we directly measure the reliability of the results based on the reliability of the corresponding companies. Hence, reliability (and not relevance) is the focus of our study.

The remainder of this paper is organized as follows. Section 3 describes data preparation. We introduce research questions in Section 4, and describe research methods and discuss results in Section 5. In Section 6 we provide our conclusions.

3. Data
In order to conduct our investigation we need the following data: (1) queries that may be used to search for digital cameras, (2) search engine result pages (SERPs) for each of the queries to obtain data about resellers, and (3) BBB ratings for each of the resellers identified from SERPs.
3.1 Search Queries
Using a keyword research tool from Google Adwords, for the initial seed keyword, “digital camera”, we ranked the Google suggested keywords by previous month’s (March 2007) search volume (the relative number of users searching for that keyword on Google) and identified nine more seed keywords such that each contains a different manufacturer name, such as “canon digital camera”. With each of the nine seed keywords, we repeated the above process and derived ten search keywords, such as “Canon PowerShot S2 IS”, which represent different digital camera models. Finally we obtained a total of 90 search keywords representing 90 different models from the nine manufacturers.

3.2 SERPs and Resellers
We programmatically submitted each of the 90 keywords to Google search engine and fetched content for the top three SERPs and all of the pages corresponding to the URLs in the three SERPs (on April 27, 2007). When examining results appearing in the SERPs, we consider only those results that are from websites of digital camera resellers selling brand new products, such as www.amazon.com and www.neweggs.com, and ignore links from other sites, such as camera review websites (e.g., www.dpreview.com), price comparison websites (e.g., www.pricescan.com), image-sharing sites (e.g., www.pbase.com), and personal websites or blogs. We also ignore results corresponding to non-US domains (e.g., amazon.com.uk). The product in this study can be a camera, camera body only, or camera kit (e.g., with lens). We consider only brand new products to make sure comparisons, such as price comparison, are fair. Moreover, there is only few companies selling refurbished items or having non-US domains in the top three SERPs.

If a result is from a reseller’s website and the page sells the corresponding model of camera, we record the result’s index in the SERP, URL, product name and price. We maintain the index of organic and paid results across the three SERPs. Hence, if an organic result appears as the third result on the second SERP and there are N₁ organic results on the first SERP, the result has an index of N₁+3. For paid results, we consider results listed on top to be ranked higher than those appearing on the right. For example, if first SERP has two top-listed results, then the index for the first right-listed result is 3. The price is identified from the reseller’s website and does not include shipping cost or tax. Among the 90 keywords, 44 keywords generate at least one search result that corresponds to a reseller selling the corresponding model of camera. For the 44 keywords we identified a total of 30 resellers and 139 price observations.

3.3 BBB Ratings
Founded in 1912, the Better Business Bureau (BBB) deals with consumers and businesses to resolve disputes, assist communication, and provide information on ethical business practices. For example, as of August 5 2007, BBB has recorded 586 complaints in last 36 months for the company named Best Price Cameras. Among those complaints, the company has failed to respond to 135 (23%) and the consumers remain dissatisfied for another 178 (30%) cases. This company receives an unsatisfactory rating from BBB that judges a company’ reliability on the basis of company’s size, volume of business, number of transactions, number and type of complaints, and how the company handled complaints.
There are some other company-rating websites, such as www.bizrate.com and www.ratingsplanet.com. Unlike the BBB system, they allow customers to post feedback for a company and share them with the public. A potential problem is that such comments can be that they are misleading and incorrect as comments are often not verified and even the users can be fabricated. As our focus is not on comparing and validating the quality of different rating sites, we choose BBB’s ratings given its informative and reliable records and a systematic approach in rating reliability. Among the 30 resellers identified through the SERPs, BBB provides explicit ratings (i.e., satisfactory or unsatisfactory) for 21 companies, has dispute and resolution information but no ratings for 5, and does not include 4 companies. We assign a rating of satisfactory for the 5 companies because 4 of them are large retailers (CircuitCity, CompUSA, Target, and Wal-Mart) and we do not include the 4 companies that do not exist in BBB in our analysis. Finally we get 133 price observations representing 44 camera models from 26 resellers.

4. Research Questions
We describe our research objectives and present our specific research questions in the form of hypotheses. We call a company with a satisfactory rating a satisfactory company, and unsatisfactory company is a reseller with an unsatisfactory rating. Our hypotheses directly compare organic and sponsored results. In particular, we compare them on the basis of average item price, number of resellers, average price dispersion, and average rank on SERPs. We define the average price dispersion of an item as \( \frac{\text{range of item price}}{\text{mean item price}} \) which is same as suggested in Ratchford et al. (2003). When calculating average item price, the item’s model and color are the same across different resellers. The hypotheses are tested across products but the basic unit of analysis is a single and identical product.

**H1:** The average item price offered by unsatisfactory companies is lower than that by satisfactory companies

It seems reasonable that companies with a bad rating may sell the same item at a lower price in order to attract customers. This hypothesis compares satisfactory with unsatisfactory companies on average item price.

**H2a:** The average rank/index of unsatisfactory resellers in sponsored results is higher than that in organic results  
**H2b:** The number of unsatisfactory resellers in sponsored results is larger than that in organic results

Unlike sponsored results that depend on bidding for keywords, organic search results are based on search engine algorithms that consider a lot of factors relating to a website, such as in-links and reputation of in-linked pages. As a result, resellers in top organic results can be generally considered to be more important (Brin and Page 1998), trustworthy (Palmer et al. 2000), and authoritative (Kleinberg 1999), and thus are less likely to be firms with unsatisfactory ratings. Hence we assume the average rank/index of unsatisfactory resellers in sponsored results is higher than that in organic results. For the same reason, we expect the number of unsatisfactory resellers in sponsored results to be larger than that in organic results.

**H3a:** The average item price in organic results is higher than that in sponsored results.  
**H3b:** The average price dispersion in organic results is smaller than that in sponsored results.
We expect that sponsored results include more unsatisfactory companies which may sell items at lower prices and cause greater price dispersions in sponsored results. For the same reason, we expect that the average item price in organic results is higher than that in sponsored results and average price dispersion in organic results is smaller than that in sponsored results.

**H4a:** Among organic results, the average rank of unsatisfactory resellers is lower than that of satisfactory resellers.

**H4b:** Among sponsored results, the average rank of unsatisfactory resellers is higher than that of satisfactory resellers.

We do not expect unsatisfactory companies to fare well among the organic results. However we assume that in paid results they can be ranked better by bidding higher than satisfactory companies. For H4b, we want to find out whether unsatisfactory companies on average “dominate” the sponsored results in terms of being ranked higher than satisfactory companies.

### 5. Research Methods and Results

Table 1 shows each of the hypotheses described above and their results. If the statement corresponding to a given hypothesis is true for a product (keyword) we mark the observation as 1; otherwise, we mark it as 0. We report the observed numbers of 0s and 1s across all unique products and the significance values of a two-tailed Binomial test. The total number of 0s and 1s in Table 1 may not always be 44 (number of unique products) since for some keywords no data to support or negate the hypothesis statement is available in SERPs. To illustrate, for hypothesis H1, we find that for 13 products the average item price offered by unsatisfactory companies is lower than by satisfactory companies and for 3 products this is not true. None of the remaining products are sold by companies in both organic and paid results (in the first 3 SERPs). Other hypotheses are tested in a similar manner and so we do not include the details.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Number of 1s</th>
<th>Number of 0s</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: The average item price offered by unsatisfactory companies is lower</td>
<td>13</td>
<td>3</td>
<td>0.021</td>
</tr>
<tr>
<td>than that by satisfactory companies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2a*: The average rank/index of unsatisfactory resellers in sponsored</td>
<td>15</td>
<td>4</td>
<td>0.019</td>
</tr>
<tr>
<td>results is higher than that in organic results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2b*: The number of unsatisfactory resellers in sponsored results is</td>
<td>15</td>
<td>3</td>
<td>0.008</td>
</tr>
<tr>
<td>larger than that in organic results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H3a: The average item price in organic results is higher than that in</td>
<td>18</td>
<td>10</td>
<td>0.185</td>
</tr>
<tr>
<td>sponsored results.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H3b: The average price dispersion in organic results is smaller than that</td>
<td>10</td>
<td>8</td>
<td>0.815</td>
</tr>
<tr>
<td>in sponsored results.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H4a: Among organic results, the average rank of unsatisfactory resellers</td>
<td>38</td>
<td>1</td>
<td>0.000</td>
</tr>
<tr>
<td>is lower than that of satisfactory resellers.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H4b: Among sponsored results, the average rank of unsatisfactory resellers</td>
<td>11</td>
<td>24</td>
<td>0.041</td>
</tr>
<tr>
<td>is higher than that of satisfactory resellers.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* normalizing the measure (i.e., rank in H2a and number in H2b) by the total number of results in top 3 SERPs provides similarly significant results

| Table 1 Hypotheses H1 to H4b and their significance values |

For H4b, for each product (sold by either a satisfactory or unsatisfactory company), we identify the indices of the companies among the sponsored results, calculate the average indices for the two types of companies, and compare the indices to produce a binary value. For a product, if the
average index of unsatisfactory resellers is smaller (i.e., ranked higher) than that of satisfactory companies, we denote the result as 1, and otherwise as 0. If, for a product, only satisfactory companies appear within the sponsored results, we treat the observation as 0 because the best index that an unsatisfactory company can achieve is the total number of sponsored results in top three SERPs + 1 which must be larger (i.e., ranked lower) than the index for satisfactory companies.

As shown in Table 1, the hypothesis H1 is supported \((p < 0.05)\) and the unsatisfactory resellers indeed provide lower prices than satisfactory resellers, which suggests that at least some unsatisfactory resellers try to lure customers by offering lower prices. However as suggested by Cao et al. (2003-2004), lower price strategy doesn’t benefit a company in the long run if there is lower satisfaction with the fulfillment process which is a key driver of customer loyalty. Their result indicates that for our study those unsatisfactory companies, which try to attract customers by offering lower prices and many of which pay search engine companies in order to be displayed on the paid search, may not be able to build strong long-term customer relationship due to failing to handle customer complaints properly and therefore succeed in the long run. In synch with commercial nature of paid results, unsatisfactory companies are ranked higher \((p < 0.05)\) in sponsored results than in organic results, and paid results include more unsatisfactory resellers \((p < 0.01)\). Thus we accept H2a and H2b. However we do not observe significant differences for average item price and average price dispersion between organic and sponsored results, and therefore reject H3a and H3b. The results indicate that resellers in organic results can provide competitive prices as well. For organic results, we find an unsatisfactory reseller is ranked lower than satisfactory companies \((p < 0.001)\), and so H4a is supported. Although unsatisfactory companies are more likely to appear in paid results than in organic results, we do not find that they are ranked higher than satisfactory resellers among the paid results (the number of 1s is smaller than number of 0s). Hence, we reject H4b.

6. Conclusions and Future Directions

Paid results have been the focus of several media reports and industry studies that have indicated a certain level of consumer distrust towards paid results. We present empirically supported and direct evidence that unsatisfactory companies do find their way more easily into the paid results than into the organic results. We also find that the unsatisfactory companies do not overwhelm the paid results by grabbing all of the top ranks as compared to satisfactory companies. Our findings should motivate search engine companies to factor in reliability information while presenting paid results to gain greater trust from consumers. Also, our findings have implications for search consumers since it provides them with an empirically verified warning about greater unreliability of paid results. We realize that a company’s performance and its demographics, such as the size, affect the frequency and timing of its bids for paid search. Since our data was collected at one time point, we would have missed out on some companies.

In the future, we plan to verify our results on larger data sets by covering multiple time points, including different types of products besides digital cameras, taking company demographic information into account, using other search engines, looking for additional resources such as sites’ ratings by BizRate (http://www.bizrate.com/) that provide richer reliability data, and examining how sites’ SEO qualities correlate with companies’ reliability. We also plan to apply the lessons learned from this empirical analysis to design a classification system that can
automatically detect unsatisfactory merchants among the paid and organic results. For example, the empirical analysis presented here informs us that the prior probability of finding an unsatisfactory merchant is higher for paid results which may be utilized in a Bayesian classification model. However, the priors are expected to vary with product category and may be extremely skewed making this classification a difficult problem. Another challenge would be with respect to identifying appropriate attributes (such as price, product features, linkage structure etc.) for classification. Investigating, comparing, and classifying different company-rating websites and price comparison sites can also be an interesting future direction.

References