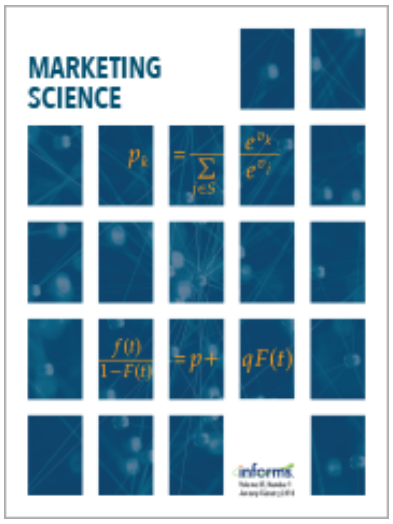

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Role of Search Engine Optimization in Search Marketing

Ron Berman, Zsolt Katona,

To cite this article:
Ron Berman, Zsolt Katona, (2013) The Role of Search Engine Optimization in Search Marketing. Marketing Science 32(4):644-651. https://doi.org/10.1287/mksc.2013.0783

MARKETING SCIENCE
Vol. 32, No. 4, July–August 2013, pp. 644–651
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2013.0783

The Role of Search Engine Optimization in Search Marketing

Ron Berman, Zsolt Katona
Haas School of Business, University of California, Berkeley, Berkeley, California 94720
{ron_berman@haas.berkeley.edu, zskatona@haas.berkeley.edu}

This paper examines the impact of search engine optimization (SEO) on the competition between advertisers for organic and sponsored search results. The results show that a positive level of search engine optimization may improve the search engine's ranking quality and thus the satisfaction of its visitors. In the absence of sponsored links, the organic ranking is improved by SEO if and only if the quality provided by a website is sufficiently positively correlated with its valuation for consumers. In the presence of sponsored links, the results are accentuated and hold regardless of the correlation. When sponsored links serve as a second chance to acquire clicks from the search engine, low-quality websites have a reduced incentive to invest in SEO, giving an advantage to their high-quality counterparts. As a result of the high expected quality on the organic side, consumers begin their search with an organic click. Although SEO can improve consumer welfare and the payoff of high-quality sites, we find that the search engine's revenues are typically lower when advertisers spend more on SEO and thus less on sponsored links. Modeling the impact of the minimum bid set by the search engine reveals an inverse U-shaped relationship between the minimum bid and search engine profits, suggesting an optimal minimum bid that is decreasing in the level of SEO activity.

Key words: search engine marketing; electronic commerce; marketing contests
History: Received: January 22, 2011; accepted: January 21, 2013; Preyas Desai served as the editor-in-chief and Chakravarthi Narasimhan served as associate editor for this article. Published online in *Articles in Advance* May 22, 2013.

1. Introduction

Consumers using a search engine face the option of clicking organic or sponsored links. The organic links are ranked according to their relevance to the search query, whereas the sponsored links are allocated to advertisers through a competitive auction. Because consumers tend to trust organic links more, advertisers often try to increase their visibility in the organic list by gaming the search engine's ranking algorithm using techniques collectively known as *search engine optimization* (SEO).[1]

A notable example of the dramatic impact that an SEO campaign can have is that of JCPenney, an American retailer. This retailer's organic links skyrocketed during the 2010 holiday shopping season and suddenly climbed to the top of the search results for many general keywords such as "dresses," "bedding," and "furniture" (Segal 2011). JCPenney eventually fired their SEO contractor after finding out that they used black hat techniques that eventually led to a punitive response from Google. Search engine optimization is widespread in the world of online advertising; a 2010 survey of 1,500 advertisers and agencies revealed that 90% of them engaged in SEO, compared with 81% who purchased sponsored links (Econsultancy 2010). In the past few years, search engine optimization has grown to become a multibillion dollar business (VanBoskirk 2009).

This paper explores the economics of the SEO process and its effects on consumers, advertisers, and search engines. Using a game-theoretical model, we fully characterize the incentives and trade-offs of all players in the ecosystem. Our model consists of (i) advertisers with exogenous qualities and potentially correlated valuations for clicks, competing for the attention of consumers; (ii) a search engine that offers both organic and sponsored links and can set minimum bids; and (iii) consumers who engage in costly search to find the highest quality site. To capture the effect of SEO, we model the imperfections in the algorithms used by search engines, assuming that there is a measurement error that prevents the search engine from perfectly ordering links according to quality. Advertisers can, in turn, manipulate the potentially erroneous quality observations to their advantage through SEO and improve their ranking. A key parameter of our model is the effectiveness of SEO, determining the extent to which SEO efforts by advertisers affect the organic results.

[1] We focus only on "black hat" SEO, which does not improve the actual relevance of the Web page to the query, but just games, the ranking algorithm.

We first ask how SEO changes the organic results and whether these changes are always detrimental to consumers and high-quality advertisers. The interest in this question stems from the strong stance that search engines typically take against SEO by emphasizing the potential downside on organic link quality. To justify their position, search engines typically claim that their manipulation results hurt consumer satisfaction and decrease the welfare of "honest" sites. In contrast, search engines also convey the message that the auction mechanism for sponsored links ensures that the best advertisers will obtain the links of highest quality, resulting in higher social and consumer welfare. This reasoning suggests that consumers should trust sponsored links more than organic links in equilibrium and would prefer to start searching on the sponsored side. A substantial contribution of using a sophisticated model for consumers is that we are able to derive their optimal search behavior. Contrary to claims by search engines, we find that search engines fight SEO because of the trade-off advertisers face between investing in sponsored links and investing in influencing organic rankings. As a consequence, search engines may lose revenue if sites spend significant amounts on SEO activities instead of on paid links and content creation.

To approach the issue of diminished welfare from SEO, we first focus on the case where sponsored links are not available to advertisers and consumers. This base model serves as a benchmark and gives us a deeper understanding of the nature of the competition for organic links when using SEO activities. Our first result reveals that SEO can be advantageous by improving the organic ranking. In the absence of sponsored links, this only happens when advertiser quality and valuation are positively correlated. That is, if sites' valuations for consumers are correlated with their qualities, then consumers are better off with than without some positive level of SEO. By contrast, if there are sites that extract high value from visitors yet provide them with low quality, then SEO is generally detrimental to consumer welfare. The SEO process essentially allows sites with a high value for consumers to correct the search engine's imperfect ranking through a contest.

The second question we ask focuses on the full interaction between organic and sponsored links when SEO is possible. The institutional differences between the organic and sponsored lists are critical to the understanding of our model. First, advertisers usually pay for SEO services up front, and the effects can take months to materialize. On the other hand, bids for sponsored links can be frequently adjusted depending on the ordering of the organic list. Second, SEO typically involves a lump-sum payment for initial results, and the variable portion of the cost tends to be convex, whereas payment for sponsored links is on a per-click basis, with very little or no initial investment. Finally, there is substantial uncertainty in the outcome of the SEO process depending on the search engine algorithms, whereas sponsored links are allocated through a deterministic auction.

It is noteworthy that the presence of sponsored links accentuates the results of the base model and that SEO favors the high-quality advertiser regardless of the correlation between quality and valuation. The intuition is that sponsored links act as a backup for high-quality advertisers in case they do not possess the top organic link. When consumers have low search costs, they will eventually find the high-quality advertiser, reducing the value of the organic position for a low-quality player. In equilibrium, consumers will start searching on the organic side, and high-quality sites will have an increased chance of acquiring the organic link as SEO becomes more effective.

Although SEO clearly favors high-quality advertisers, we find that there is a strong tension between the interests of consumers and the search engine. As advertisers spend more on SEO and consumers are more likely to find what they are looking for on the organic side, they are less likely to click on revenue-generating sponsored links. This tension may explain why search engines take such a strong stance against SEO, even though they favor a similar mechanism on the sponsored side. Furthermore, we obtain an important normative result that could help search engines mitigate the revenue loss stemming from SEO: we find that there is an optimal minimum bid the search engine can set that is decreasing in the intensity of SEO. Setting the minimum bid too high, however, could drive more advertiser dollars away from the sponsored side toward SEO.

As common as the practice of SEO may be, research on the topic is scant. Many papers have focused on sponsored links and some on the interaction between the two lists. In all of these cases, however, the ranking of a website in the organic list is assumed exogenous, and the possibility of investing in SEO is ignored. On the topic of sponsored search, works such as those by Rutz and Bucklin (2007) and Ghose and Yang (2009) focus on consumer response to search advertising and the different characteristics that affect advertising efficiency. Other recent examples, such as those by Chen and He (2011), Athey and Ellison (2011), and Xu et al. (2011), analyze models that include both consumers and advertisers as active players.

A number of recent papers study the interplay between organic and sponsored lists. Katona and Sarvary (2010) show that the top organic sites may not have an incentive to bid for sponsored links. In an empirical piece, Yang and Ghose (2010) show

that organic links have a positive effect on the click-through rates of paid links, potentially increasing profits. Taylor (2013), White (2013), and Xu et al. (2012) study how the incentives of the search engine to provide high-quality organic results are affected by potential losses on sponsored links. The general notion is that search engines have an incentive to provide lower-quality results in order to maximize revenues.

The work of Xing and Lin (2006) is the closest antecedent to our paper. It defines "algorithm quality" and "algorithm robustness" to describe the search engine's ability to accurately identify relevant websites. Their paper shows that when advertisers' valuations for organic links are high enough, SEO is sustainable, and SEO service providers can then free ride on the search engine as a result of their "parasitic nature." The relationship between advertiser qualities and valuations and the strategic nature of consumer search is not taken into account. An earlier work by Sen (2005) develops a theoretical model that examines the optimal strategy of mixing between investing in SEO and buying ad placements. It is noteworthy that the model shows that SEO should not exist as part of an equilibrium strategy.

2. Model

We set up a static game in which consumers search for a phrase and advertisers compete for their visits. We assume there is a monopolistic search engine that provides search results to consumers by displaying links to one of two websites. These sites can also buy sponsored links from the search engine. Whenever a consumer enters the search phrase, the search engine ranks the sites according to a scoring mechanism and presents one organic link and one sponsored link according to the scores and bids of the sites. The incentives and characteristics of the search engine, advertisers, and consumers are described in the next section.

2.1. Websites and Consumers

Consumers in our model seek to consume one unit of a good that can have a quality $q_i \in \{q_L, q_H\}$ with $q_H > q_L$. The good is provided by websites and can either be information, content, or a physical product. Regardless of its nature, the good provides a utility of q_i to those who consume it (net of price). The two possible quality levels of q_H and q_L are common knowledge, but consumers need to search to discover the particular qualities provided by each website. When visiting the search engine, consumers see an organic link and possibly a sponsored link. To discover the quality provided by a site, consumers need to click onto the links. Upon visiting a site, consumers incur a search cost $c \geq 0$ and discover the quality of the good. They then decide whether to continue the search, abandon it, or consume the good they have found. The decision on which link to start with (organic or sponsored) and the decision to continue searching depends on the expected distribution of qualities behind each link. A rational consumer will continue searching only if the expected increase in utility from visiting the next link outweighs the search cost. Once the consumer has decided to stop searching, she will consume the good with the highest net utility, possibly returning to a previously visited link.

As an example, if the consumer started searching with the organic link and found a website providing quality q_H, she has no reason to continue searching. She will consume the good yielding utility of $q_H - c$. If, on the other hand, she started searching with the organic link and found a site providing quality q_L, she would prefer to continue searching when search costs are low. If she also found q_L behind the sponsored link when continuing, she would eventually receive utility $q_L - 2c$.

The website that provides the chosen good receives an exogenously determined revenue valued at $v_i \in \{v_L, v_H\}$ with $v_H > v_L$. The total revenue of site i (net of manufacturing costs) is thus the number of consumers who consume its good multiplied by v_i. For example, in the case when the good is a product sold by the advertiser, v_i can be thought of as the per-unit margin of the seller. The individual site qualities q_i and valuations v_i are known by the competing websites but are unknown to the consumers or the search engine a priori. However, the following distribution is common knowledge: $\Pr(q_i = q_L) = \Pr(q_i = q_H) = \frac{1}{2}$, $\Pr(v_i = v_L) = \Pr(v_i = v_H) = \frac{1}{2}$, and the correlation between q_i and v_i is ρ for each site i. Both qualities and valuations are independent across sites. The sign of the correlation between the quality and valuation of a particular site could be driven by several factors in a market. For example, in a vertically differentiated market, firms offering a higher-quality product can charge a premium and often make a higher margin, suggesting a possibly positive correlation. However, a negative correlation is also possible between qualities and valuations because of deceptive marketing practices or interaction with other channels.

To influence their organic ranking, websites can invest SEO effort e_i at a quadratic cost of $e_i^2/2$. To win the sponsored link, websites submit per-click bids, denoted by b_i. The total payment for the sponsored link is determined in a generalized second-price auction with minimum bid r, where bids are corrected for expected click-through rates (CTRs). The final payoff of site i is therefore its revenue minus the SEO investments costs and the sponsored payment.

2.2. The Search Engine

The search engine acts as an intermediary between consumers and websites. Its goal is to provide consumers with links to the highest-quality websites on

the organic side while making a profit through the auctioning of sponsored links. To rank websites, the search engine scores each website on its estimated quality using information gathered from the Internet using crawling algorithms and data mining methods. The search engine can therefore only measure quality with an error and cannot observe it directly. We model the score of each site as

$$s_i = q_i + \alpha e_i + \varepsilon_i, \tag{1}$$

where α is a parameter denoting the effectiveness of SEO and ε_i is the measurement noise, distributed according to a distribution with cumulative distribution function (c.d.f.) F_ε and mean 0. The parameter α measures how easy it is to change one's ranking using SEO methods. That is, $1/\alpha$ influences the cost of SEO, which can be controlled by several factors, including the search engine. Indeed, if the search engine ignores the possibility of SEO activities, α presumably increases.

Sponsored links are awarded by the search engine in a standard click-through rate-corrected, second-price auction with a reserve minimum bid of r. If website i has an expected click-through rate of ctr_i, the search engine awards the links in order of the ranking of the scores $ctr_i \cdot b_i$, as long as they are higher than the minimum bid. When a consumer clicks on a sponsored link, the website who owns it pays the bid of the next highest bidder corrected for the click-through rate differences. The click-through rates are a result of the endogenous consumer search process in equilibrium. They determine the payoff of the search engine as well as influence the incentives of the advertisers to invest in SEO. Our model takes these click-through rates into account when considering the bids of advertisers for sponsored links.

2.3. Timing

At the beginning of the game, the search engine publishes the minimum bid for sponsored links r. In parallel, nature determines the quality q_i and the valuation v_i for each website, given the correlation parameter ρ, but independently across sites. Websites then decide on the amount of effort e_i to invest in SEO. The search engine then determines the scores s_i of each site and publishes their score ranking. Following the organic ranking, sites bid for the sponsored links that are then awarded according to a CTR-corrected generalized second-price auction with a minimum bid r. Once both rankings have been finalized, consumers initiate a search process.

Before visiting the search engine, consumers decide which link gives them the highest expected utility and start their search with that link.[2] The consumers then decide whether to consume the good encountered or continue their search. Once the consumer has searched through all of the links, has decided to stop searching, and consumes, payoffs are realized.

3. SEO Equilibrium

3.1. Organic Links Only

When the minimum bid is higher than what the profit websites expect from a visitor, advertisers cannot afford sponsored links. This scenario is very common when sites provide free content to consumers and make a profit by selling advertising. It also serves as a benchmark case before analyzing the impact of sponsored links on the SEO process. The expected payoff of site i is then

$$\pi_i = v_i \cdot \Pr(s_i > s_j) - \frac{e_i^2}{2}. \tag{2}$$

To illustrate our results, we assume that the measurement error has a uniform distribution $\varepsilon_i \sim \mathrm{U}[-\sigma/2, \sigma/2]$ with a large enough support.[3] To show the impact of SEO on consumers and the overall ranking, we use $P(\alpha) = P(\alpha; \sigma, v_1, v_2, q_1, q_2)$ to denote the efficiency of the ranking process, which is the probability of the website with the highest-quality winning the organic link. Because the utility of the consumer is the quality of the consumed good, consumer welfare increases with efficiency.

Simple analysis shows that when search engine optimization is *not* possible, i.e., when $\alpha = 0$, we get $P(0) < 1$ as long as $q_1 \neq q_2$ because of the noise in the ranking process. Furthermore, $P(0, \sigma)$ is decreasing in σ as higher levels of noise make the ranking less efficient. When search engine optimization becomes effective, i.e., when $\alpha > 0$, websites can actively influence the order of results. The following proposition summarizes how SEO affects the ranking, consumer welfare, and firm profits.

Proposition 1. (a) *When $\rho = 1$, any $\alpha > 0$ that is not too large improves the efficiency of the ranking and consumer satisfaction. However, when $\rho = -1$, SEO is detrimental to consumer satisfaction. For intermediate $-1 < \rho < 1$ values, SEO can improve consumer satisfaction for some α values.*

(b) *Suppose α is small. When $\rho = -1$, both sites' profits are decreasing in α. When $\rho = 1$, sites' profits are decreasing in α, with the exception of the higher-quality site, whose profits are increasing iff $v_H > 2v_L$.*

[2] Because there might be a case with no sponsored links, we assume that consumers incur the cost c of the first search even if their favorite link does not exist. This is a technical assumption that makes the analysis cleaner. Alternative, and perhaps more realistic, assumptions lead to similar results.

[3] We need to assume $\sigma > q_1 - q_2$ for the error to have any effect. The online appendix (available at http://dx.doi.org/10.1287/mksc.2013.0783) illustrates equivalent results for the general distribution of the errors.

Part (a) of Proposition 1 demonstrates the main effect of equilibrium SEO investments on the ranking. The SEO mechanism gives both sites incentives to invest in trying to improve their ranking but favors bidders with high valuations. Because the search engine cannot measure site qualities perfectly, this mechanism corrects some of the error when valuations are positively correlated with qualities. On the flip side, when lower-quality sites have high valuations for traffic, SEO creates incentives that are not compatible with the utilities of consumers. In this latter case, the high valuation sites that are not relevant can get ahead by investing in SEO. Examples are cases of "spammer" sites that intentionally mislead consumers. Consumers gain little utility from visiting such sites, but these sites may profit from consumer visits.

Closer examination of the proof suggests that $\partial P(\alpha, \sigma)/\partial\alpha\partial\sigma$ is positive for small α's. This suggests, somewhat counterintuitively, that investments against SEO on the search engine's part complement investments in better search algorithms rather than substitute them. That is, only search engines that are already very good at estimating true qualities should fight hard against SEO. Nevertheless, as measurement error can depend on exogenous factors and can vary from keyword to keyword, it may make sense to allow higher levels of SEO in areas where the quality measurement is very noisy.

To analyze the relationship between α and advertiser profits, we focus on small levels[4] of α. As part (b) of Proposition 1 shows, the player with the lower valuation is always worse off with higher SEO effectiveness regardless of its quality. The only site that benefits from SEO is the one with a quality advantage—and only if its valuation is substantially higher than its competitor's. The intuition follows from the fact that higher levels of SEO emphasize the differences in valuations; the higher the difference, the more likely that the higher valuation will win. It is noteworthy that an advantage in valuation only helps when the site also has a higher quality; that is, spammer sites with low quality and high valuation will not benefit from SEO because of the intense competition with better sites.

3.2. The Role of Sponsored Links

We now examine how the availability of sponsored advertising changes the incentive of investing in SEO and the resulting link order. Because the search engine's main source of revenue comes from sponsored links, this analysis is crucial to understanding how SEO affects the search engine's revenue. We solve the model outlined in §2 with $r < v_H$. That is, at the minimum, sites with a high valuation will be able to pay for sponsored links. When describing the intuition, we focus on the case of $r < v_L$ so that any site can afford sponsored links.

To determine advertisers' SEO efforts and sponsored bids, we also need to uncover where consumers start their search process. We assume that consumers always incur a small but positive search cost. They have rational expectations and start with the link that gives them the highest probability of finding a high-quality result without searching further. The following proposition summarizes our main results.

Proposition 2. *There exists a $\bar{c} > 0$, such that if $c < \bar{c}$, then*

(a) *in the unique equilibrium, consumers begin their search on the organic side;*

(b) *if $r < v_L$, the likelihood of a high-quality organic link is increasing in α for any $-1 \le \rho \le 1$;*

(c) *if $v_L \le r$, the likelihood of a high-quality organic link is increasing in α iff ρ is high enough; and*

(d) *the search engine's revenue increases in α iff the likelihood of a high-quality organic link decreases.*

In short, we prove that the presence of sponsored links accentuates the potential benefits of SEO on increasing the quality of the organic link. As α increases and SEO becomes more effective, the probability that the higher-quality site acquires the organic link increases even if advertisers' qualities and their valuations for consumers are negatively correlated. Contrary to the commonly held view that SEO often helps low-quality sites climb to the top of the organic list if they have enough resources, we find that in the presence of sponsored links, low-quality sites cannot take advantage of SEO. The intuition relies on the notion that sponsored links serve as a second chance to acquire clicks from the search engine for the site that does not possess the organic link. However, as a result of exhaustive consumer search, high-quality sites enjoy a distinct advantage as they are likely to be found regardless of their position. Low-quality advertisers, on the other hand, suffer if a higher-quality competitor is also on the search page. Thus, a low-quality site's incentive to obtain the organic link will be reduced, whereas high-quality sites will face less competition in the SEO game and will be more likely to win it. For high-quality sites, the main value of acquiring the top organic link is not merely the access to consumers. Instead, the high-quality site benefits from the organic link because it does not have to pay for the access to consumers, as it would have to on the sponsored side.

In the ensuing equilibrium, high-quality advertisers always spend more on SEO than their low-quality competitors. Because this increases the chances of

[4] This relationship can be quite complex in the general case.

high-quality organic links, we find that rational consumers start their search on the organic side. Consumers benefit from finding a high-quality link as early as possible, and thus, more effective SEO increases their welfare by increasing the likelihood of a high-quality organic link. This fact, however, hurts the search engine whose revenues decrease when the high-quality advertiser competes less for the sponsored link. The misalignment between consumer welfare and search engine profits has already been recognized by White (2013) and Taylor (2013). Our results reconfirm this tension and shed light on an interesting fact: the main danger of SEO for search engines is not the disruption of the organic list, which has long-term impact on reputation and visitors, but rather decreased revenues on the sponsored side, which are of a short-term nature. Often advertisers pay third parties to conduct SEO services instead of paying the search engine for sponsored links. The result from the advertiser's perspective is not much different, but the search engine is stripped of significant revenues.

The search engine has an important tool on the sponsored side, setting the minimum bid that affects what the winning advertiser pays. In the absence of SEO, an increased minimum bid directly increases the revenue from advertisers who have a valuation above the minimum bid. When SEO is possible, the situation is different.

Corollary 1. *There exists an* $\hat{r}(\alpha) > 0$ *such that the search engine's revenue is increasing in* r *for* $r < \hat{r}(\alpha)$ *and decreasing for* $\hat{r}(\alpha) < r < v_L$. *When* v_L *is high enough,* $\hat{r}(\alpha)$ *is the unique optimal minimum bid, which is decreasing in* α.

The inverse U shape of the effect is a result of two opposing forces. An increasing minimum bid increases revenue directly. However, in the presence of SEO, a higher minimum bid makes sites invest more in SEO, which makes the high-quality site more likely to acquire the organic link. This, in turn, will lower sponsored revenues as most of these revenues come from the case when the low-quality site possesses the organic link. The combination of these two forces will make the search engine's revenue initially increase with an increased minimum bid but begin to decrease when sites invest more in SEO. The maximal profit is reached at a lower minimum bid as SEO becomes more effective (α increases). Finally, we examine how a site's revenues are affected by SEO.

Corollary 2. *If* $r < v_L$ *and the two sites have different qualities, the profit of the higher-quality site increases and the profit of the lower-quality site decreases in* α.

As we explained above, the possibility of using sponsored links as a backup gives an advantage to the higher-quality site. The more effective search engine optimization is, the less the site has to spend to secure the top organic link. The lower-quality site faces the exact opposite situation. When the two sites have the same qualities, SEO only makes a difference when those qualities are low. In this case, a higher α benefits the site with the higher valuation.

4. Conclusion

The options facing consumers when using an online search engine are highly affected by search engine marketing decisions made by website owners and the policy of the search engine. Site owners can choose to invest in SEO effort to promote their site in organic listings as well as bid for sponsored links. Search engines can choose to handicap SEO activities or impose a minimum bid requirement. We find that, contrary to popular belief, SEO can sometimes be beneficial to consumers by giving an advantage to high-quality sites, especially when the search engine's crawling algorithms do not provide an accurate ranking. Such improvement in the quality of search results will attract more consumers yet will hurt the revenues of search engines.

Our results also provide important recommendations to advertisers. When organic links are the only option, SEO is an important tool to increase a site's visibility for advertisers who can afford to pay more. The majority of online advertisers invest in both SEO and sponsored links and face an important dilemma as to how to allocate their budget between the two activities. Our results imply that high-quality sites have an advantage as they can always use sponsored links as a backup option if their organic link does not place well. As a consequence, the main value of SEO for them is to avoid the potentially hefty payments for sponsored clicks.

We believe that the economics of search engine optimization is a topic of high importance for both academics and practitioners. In this paper we examine the basic forces of this intriguing, complex ecosystem. Given the complexity of the problem, our model has a number of limitations that could be explored by future research. First, we model SEO as a static game, whereas in reality, sites invest in SEO dynamically, reacting to each other's and the search engine's actions. Our static approach limits our ability to explore how the search engine's reputation is affected in the long run. Second, we focus our attention on a single keyword with one organic and one sponsored link throughout the paper. In reality, advertisers bid for millions of keywords to obtain sponsored links. Conducting SEO is less a fine-grained activity, and it may affect the ranking of a site for several different keywords. Third, we use the term "SEO" exclusively for the black hat type of optimization and do

not model white hat methods that directly increase quality. Finally, we assume that consumers search rationally and stick to their objectives. In reality, consumers might make mistakes or get distracted by different types of links, leading to clicks that our model does not predict. Despite these limitations, we believe that our paper is an important step in the direction of understanding the role of search engine optimization in marketing.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2013.0783.

Acknowledgments

The authors thank Pedro Gardete, Ernan Haruvy, Ganesh Iyer, Shachar Kariv, John Morgan, Miklos Sarvary, Dana Sisak, Felix Vardy, Kenneth Wilbur, Yi Zhu, and seminar participants at the Interdisciplinary Center (IDC) Herzliya, Hong Kong University of Science and Technology, University of Florida, University of Houston, University of Texas at Austin, and Yale University for their useful comments. They are also grateful to the editor, the associate editor, and the two anonymous reviewers for their helpful suggestions.

Appendix

Proof of Proposition 1. Let $F_{\varepsilon_i-\varepsilon_j}$ be the c.d.f. of a triangle distribution $\varepsilon_i - \varepsilon_j \sim \mathrm{T}[-\sigma, \sigma]$ with mean zero and $f_{\varepsilon_i-\varepsilon_j}$ being its probability density function. Each website faces the following first-order condition (FOC) with respect to their scores resulting from the profit function:

$$v_i \cdot f_{\varepsilon_i-\varepsilon_j}(\bar{s}_i - \bar{s}_j) = \frac{\bar{s}_i - q_i}{\alpha^2}, \tag{3}$$

where $\bar{s}_i = \mathbb{E}_{\varepsilon_i}[s_i]$. Let $x = \bar{s}_i - \bar{s}_j$ and $\mu = q_i - q_j$. By subtracting both FOCs and using the fact that $f_{\varepsilon_i-\varepsilon_j}$ is symmetric around zero, we can rewrite the condition as

$$f_{\varepsilon_i-\varepsilon_j}(x) = \frac{x-\mu}{\alpha^2(v_i - v_j)}. \tag{4}$$

An interior solution x^* would require both FOCs and second-order conditions (SOCs) to hold as well as $-\sigma \le x^* \le \sigma$. When $v_i > v_j$ and $\alpha^2 \ge -\sigma(\mu/(v_i - v_j))$, or when $v_i < v_j$ and $\alpha^2 < \sigma(\mu/(v_j - v_i))$, the equilibrium solution is

$$s_i^* - s_j^* = x_R^* = \frac{\sigma^2\mu + \sigma\alpha^2(v_i - v_j)}{\sigma^2 + \alpha^2(v_i - v_j)}.$$

When $v_i < v_j$ and $\alpha^2 \ge \sigma(\mu/(v_j - v_i))$, or when $v_i > v_j$ and $\alpha^2 < -\sigma(\mu/(v_i - v_j))$, the equilibrium solution is

$$s_i^* - s_j^* = x_L^* = \frac{\sigma^2\mu + \sigma\alpha^2(v_i - v_j)}{\sigma^2 - \alpha^2(v_i - v_j)}.$$

We can immediately verify that the condition $\sigma > \mu$ ensures that $-\sigma \le x \le \sigma$, whereas $\alpha^2 < \sigma^2/v_H$ ensures that both the FOCs and the SOCs hold. Under the condition on α, the equilibrium point is a unique extremum and thus a global maximum.

To examine the effects of the equilibrium SEO investment on the ranking efficiency and consumer satisfaction, we let $P(\alpha)$ denote the probability that the player with the highest quality wins the organic link. Assume that $q_H = q_1 > q_2 = q_L$. In the perfectly correlated case, $x^* = x_R^*$. We then have $P(\alpha) = F_{\varepsilon_1-\varepsilon_2}(x_R^*)$ and $P'(\alpha) = f_{\varepsilon_1-\varepsilon_2}(x_R^*)(\partial x/\partial\alpha)|_{x=x_R^*} > 0$. In the perfectly negatively correlated case, when $\rho = -1$, we have $x^* = x_L^*$; thus $P'(\alpha) = f_{\varepsilon_1-\varepsilon_2}(x_L^*)(\partial x/\partial\alpha)|_{x=x_L^*} < 0$. Building on the two extreme cases, one can show for intermediate correlation values that $P(\alpha) > P(0)$ for certain $\alpha > 0$ and $0 < \rho < 1$.

To prove part (b), when $\rho = -1$, taking the derivative with respect to α of the profit functions of both players shows that at the limit of $\alpha \to 0$, the profit never increases for any of the equilibrium conditions. It should be noted that at some conditions the profit might increase for higher values of α. When $\rho = 1$, solving directly for player 1,

$$\pi_i(\alpha) - \pi_i(\alpha = 0) = \pi_i(\alpha)|_{x^*} - v_1 F_{\varepsilon_1-\varepsilon_2}(q_H - q_L) > 0, \tag{5}$$

yields the conditions $v_H > 2v_L$ or $\alpha^2 > \sigma^2(2v_L - v_H)/(v_H - v_L)^2$, where the latter condition is ruled out if α is small enough. For player 2, the same exercise shows that there is no solution for $\alpha > 0$ that increases player 2's profit. □

Proof of Proposition 2. We use backward induction and first determine the sponsored bids given the allocation of the organic link and then the SEO investments in three different cases with respect to the site qualities. Initially, we assume that consumers start with the organic link. Later, we will show that this is an equilibrium strategy and that starting with the sponsored link cannot be (part (a)). We will also determine the threshold $\bar{c}$. We start with the $r < v_L$ case and then show how the analysis changes for $v_L \le r < v_H$. Let w_O denote the organic and w_S the sponsored winner. The main technique we use is to compare the profits in equilibrium when the player does and does not occupy the organic link. The difference between these profits is the value of the organic link for that player.

Case 1. When $q_i = q_j = q_H$, consumers stop searching at the organic link and do not search further. This renders the sponsored link useless for both players leading to no valid bids above the reserve price, r. The SEO game is therefore equivalent to the case with no sponsored links.

Case 2. When $q_i = q_j = q_L$, consumers will not be satisfied with the organic link and continue to the sponsored link as long as it does not lead to the same site. If site i is the organic winner, then $ctr_i = 0$ for the sponsored link, leaving the sponsored link for site $j \ne i$ to win at a price per click equal to r. Since $q_i = q_j$, consumers do not go back to the organic link, leaving 0 profits for site i. The organic link is worthless; therefore no site will invest in SEO.

Case 3. When $q_i = q_H$ and $q_j = q_L$, consumers will stop at the organic link if $w_O = i$. Just as in Case 1, no site will submit a valid bid higher than r. If $w_O = j$, consumers will not be satisfied with a low-quality organic link and will continue searching, as long as the sponsored link is different from the organic. As in Case 2, $ctr_i = 1$ and $ctr_j = 0$, leading to $w_S = i$ at a price per click of r. Hence site i, with the high quality, will capture all the demand regardless of which position it is in. When $w_O = i$, this will lead to $\pi_i^O = v_i$, but

when $w_O = j$ and $w_S = i$, site i has to pay for the sponsored link and $\pi_i^S = v_i - r$. The value of winning the organic link will therefore be $\pi_i^O - \pi_i^S = r$ for site i and $\pi_j^O - \pi_j^S = 0$ for site j. Applying the results of Proposition 1 with $v_i' = r$, $v_j' = 0$, $q_i' = q_H$, $q_j' = q_L$, we get the optimal SEO efforts and the probability of a high-quality organic link:

$$e_i^* = \frac{\alpha r(\sigma - q_H + q_L)}{\alpha^2 r + \sigma^2}, \quad e_j^* = 0,$$
$$P^* = P(\alpha \mid q_i = q_H, q_j = q_L) = 1 - \frac{1}{2}\left(\frac{\sigma(\sigma - q_H + q_L)}{\alpha^2 r + \sigma^2}\right)^2. \tag{6}$$

Here, P^* is increasing in α; that is, $w_O = i$ becomes more likely as α increases regardless of ρ, proving part (b) of the proposition.

In part (c), when $v_L \le r < v_H$, the analysis is identical to the above except in Case 3, when $w_O = j$ and $v_i = v_L < r$. In this case, site i with $q_i = q_H$ cannot afford the sponsored link and will profit $\pi_i^O - \pi_i^S = v_L - 0 = v_L$ from getting the organic link, whereas site j will profit $\pi_j^O - \pi_j^S = v_H - 0 = v_H$. According to Proposition 1, a higher α decreases $\Pr(w_O = i)$, but the probability of this case is $\Pr(q_i = q_H, q_j = q_L, v_i = v_L, v = v_H) = ((1-\rho)/4)^2$, which decreases with ρ and reaches 0 when $\rho = 1$. Thus, SEO will only increase the probability of the high-quality site acquiring the organic link if ρ is high enough, proving part (c).

Returning to part (a), combining the three cases, it is clear that the organic link is more likely to be of high quality than the sponsored link. It is therefore rational for consumers to start their search with the organic link. On the other hand, assuming that consumers start with the sponsored link, redoing the same analysis shows that even then the organic link is more likely to be high quality. Starting with the sponsored link is therefore never an equilibrium strategy. Furthermore, to determine $\bar{c}$, we need to calculate the expected benefit of continuing the search when finding q_L. This is simply

$$(q_H - q_L)\Pr(q_{w_S} = q_H \mid q_{w_O} = q_L) = (q_H - q_L)\frac{(1/2)(1-P^*)}{1/4 + (1/2)(1-P^*)},$$

where P^* is defined in (6). For a consumer to even start searching, it is sufficient to assume $c < q_L$. Therefore,

$$\bar{c} = \min\left(q_L, (q_H - q_L)\frac{1-P^*}{3/2 - P^*}\right). \tag{7}$$

To prove part (d), we only need to examine Case 3, because neither consumer welfare nor search engine revenue is affected by SEO in Cases 1 and 2. In Case 3, consumers always find q_H eventually, but they are better off finding it right away, when $w_O = i$. Therefore, consumer welfare increases iff $P(\alpha)$ increases. On the other hand, search engine revenues are higher when the low-quality site acquires the organic link—that is, the revenue increases iff $P(\alpha)$ decreases—proving part (d). □

Proof of Corollary 1. Consumers only click the sponsored link if the organic link is of low quality. Thus, the search engine's revenue is $R^{SE} = (1 - P(\alpha)) \cdot r$, since the search engine makes exactly r when the low-quality site gets the organic link. From the proof of Proposition 1, we can derive

$$P(\alpha) = P(\alpha, r) = \frac{3}{4} - \frac{\sigma^2(\sigma - q_H + q_L)^2}{4(\sigma^2 + r\alpha^2)^2},$$

which is clearly increasing in r. Differentiating the revenue with respect to r yields

$$\partial R^{SE}/\partial r = 1 - P(\alpha, r) - r \cdot \frac{\partial P(\alpha, r)}{\partial r} = \frac{1}{4} + \frac{\sigma^2(\sigma - q_H + q_L)^2(\sigma^2 - r\alpha^2)}{4(\sigma^2 + r\alpha^2)^3}.$$

The above derivative is positive if r is below a suitable $\hat{r}(\alpha)$, leading to an inverse U-shaped revenue function below v_L. The implicit function theorem yields that $\hat{r}(\alpha)$ is decreasing. □

Proof of Corollary 2. When $r < v_L$, the higher-quality site has an effective valuation of r for the organic link, whereas the low-quality site has an effective valuation of 0. From the proof of Proposition 1, it is clear that the high-quality site has an increasing chance of acquiring the organic link, and its profit increases as α increases. □

References

Athey S, Ellison G (2011) Position auctions with consumer search. *Quart. J. Econom.* 126(3):1213–1270.

Chen Y, He C (2011) Paid placement: Advertising and search on the Internet. *Econom. J.* 121(556):309–328.

Econsultancy (2010) State of search engine marketing report 2010 (in association with SEMPO). Report, Econsultancy, New York.

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

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

Rutz OJ, Bucklin RE (2007) A model of individual keyword performance in paid search advertising. Working paper, Yale School of Management, New Haven, CT. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1024765.

Segal D (2011) The dirty little secrets of search. *New York Times* (February 12) http://www.nytimes.com/2011/02/13/business/13search.html.

Sen R (2005) Optimal search engine marketing strategy. *Internat. J. Electron. Commerce* 10(1):9–25.

Taylor G (2013) Search quality and revenue cannibalisation by competing search engines. *J. Econom. Management Strategy* 22(3). Forthcoming.

VanBoskirk S (2009) US interactive marketing forecast, 2009 to 2014. Report, Forrester Research, Cambridge, MA.

White A (2013) Search engines: Left side quality versus right side profits. Working paper, School of Economics and Management, Tsinghua University, Beijing.

Xing B, Lin Z (2006) The impact of search engine optimization on online advertising market. *ICEC '06: Proc. 8th Internat. Conf. Electronic Commerce* (ACM, New York), 519–529.

Xu L, Chen J, Whinston AB (2011) Price competition and endogenous valuation in search advertising. *J. Marketing Res.* 48(3):566–586.

Xu L, Chen J, Whinston A (2012) Effects of the presence of organic listing in search advertising. *Inform. Systems Res.* 23(4):1284–1302.

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.