View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Munich Personal RePEc Archive

MPRA

Munich Personal RePEc Archive

The Role of Search Engine Optimization in Search Rankings

Ron Berman and Zsolt Katona

University of California, Berkeley - Haas School of Business

12. January 2010

Online at http://mpra.ub.uni-muenchen.de/20780/

MPRA Paper No. 20780, posted 19. February 2010 23:38 UTC

The Role of Search Engine Optimization in Search Rankings *

Ron Berman and Zsolt Katona †

February 18, 2010

*We thank Pedro Gardete, Shachar Kariv, John Morgan, Miklos Sarvary, Dana Sisak, and seminar participants at Yale and UT Austin for their useful comments.

†Ron Berman is a Ph.D. student and Zsolt Katona is an Assistant Professor at the Haas School of Business, UC Berkeley, 94720-1900 CA. E-mail: ron_berman@haas.berkeley.edu, zskatona@haas.berkeley.edu.

The Role of Search Engine Optimization in Search Rankings

Abstract

Web sites invest significant resources in trying to influence their visibility among online search results. In addition to paying for sponsored links, they invest in methods known as search engine optimization (SEO). We study the economic incentives of Web sites to invest in SEO and its implications on visitor satisfaction and welfare. Our focus is on methods that improve rankings of sites among search results without improving their quality. We find that the process is equivalent to an all-pay auction with noise and headstarts. Our results show that in equilibrium, under certain conditions, some positive level of search engine optimization improves the search engine's ranking and thus the satisfaction of its visitors. In particular, if the quality of sites coincides with their valuation for visitors then search engine optimization serves as a mechanism that improves the ranking by correcting measurement errors. While this benefits consumers and search engines, sites participating in search engine optimization could be worse off unless their valuation for traffic is very high. We also investigate how search engine optimization affects investment by sites in content and find that it can lead to underinvestment as a result of wasteful spending on search engine optimization.

1 Introduction

Using an online search engine is among the most common navigational paths consumers use to discover information on the Web. As a result, search engine marketing is becoming a dominant form of online advertising. By utilizing search marketing, Web sites that wish to expose their content and merchandize to consumers can reach them when they search for specific keywords that convey information about their search goals. Such targeted advertising locations become valuable to advertisers who compete for appearing on the search results pages.

In order to accommodate advertisers, most search engines have divided their search results page into an *organic* results part and a *sponsored* results part. The left side of the screen is typically used to display organic results as a ranked list of site links ordered according to their relevance for the search query. The right side is often used to display sponsored links which are typically auctioned to advertisers using various mechanisms. Advertisers submit bids for having their ads placed among the sponsored links, and generally the highest bidders win the most visible links[1], usually on the top of the list. In this "official" way of search advertising, sites get access to the right side[2] of the search results page and it is clearly indicated to visitors that theses links are paid for.

Many advertisers, however, try to find their way to the top of the organic results' list instead of (or in addition to) competing for sponsored links. The collection of different actions that a site can take to improve its position on the organic list is called *search engine optimization* (SEO). Improving one's position can either be accomplished by making the site more relevant for consumers, or by investing in different techniques that affect the search engine's quality ranking process. These two types of SEO techniques are often referred to as *white hat SEO* and *black hat SEO* respectively. The important difference is that the latter type only improves the ranking of a site among search results without affecting its quality, whereas the former type changes the site's ranking by improving its content and by increasing visitor satisfaction. These activities include techniques of creating external links to the site

[1]The more sophisticated auction mechanisms also take into account the likelihood of a click on a given link, estimating it from historical click-through data.

[2]In some cases the search engine displays sponsored links on top of the organic results as well.

or changing the html source of the site's pages to influence the outcome of the automatic process that the search engine uses to evaluate each site's relevance. In Section 5 we discuss white hat SEO activities that consist of improving one's content and conclude that these are hard to separate from pure content investments. In the rest of the paper we focus on black hat SEO methods that do not improve content and only affect the ranking of a site. We use the term SEO to refer to such methods.

Search engines typically take a stance against black hat SEO and consider it cheating, but to a varying extent. In some cases, they entirely remove some sites from the organic list that are caught conducting such activities[3]. To set the rules, search engines sometimes publish guidelines describing undesired practices. Google, for example, prohibits buying incoming links to increase one's PageRank[4]. Yahoo, on the other hand, simply does not give weight to a paid link if they think it is not valuable to consumers[5]. In addition to simply stating what they consider allowable, search engines can also invest significant amounts in reducing the effectiveness of certain SEO activities[6]. To justify their position, search engines typically claim that manipulation of search engine results hurts consumer satisfaction and decreases the welfare of "honest" sites. One puzzling message that search engines convey is 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. Is not the case of SEO similar? If the most resourceful sites are the ones providing the best links, why not let them invest in improving their rankings? One reason why search engines may be unhappy with SEO is that if sites spend significant amounts on these activities they will spend less on paid links and content creation. To deal with this problem, some search engines allow payments for organic links and receive the money that sites would otherwise pay to third party search result optimizers. Baidu, the leading Chinese search engine and the world's third largest,

[3]BBC News reported that Google has blacklisted BMW.de for breaching its guidelines. See http://news.bbc.co.uk/2/hi/technology/4685750.stm

[4]Google Webmaster Central: http://www.google.com/support/webmasters/bin/answer.py?answer=66356

[5]Interview with Priyank Garg, director of product management for Yahoo! Search Technology: http://www.stonetemple.com/articles/interview-priyank-garg.shtml

[6]In response to Google's regular updates of its search algorithm, different sites shuffle up and down wildly in its search rankings. This phenomenon, which happens two or three times a year is called the "Google Dance" by search professionals who give names to these events as they do for hurricanes (see "Dancing with Google's spiders", *The Economist*, March 9, 2006).

does accept payments for organic links[7]. In some industries where rankings are largely based on consumer reviews it is very hard to determine which reviews are legitimate and which are not, so it is very costly to fight against SEO. For example Yelp, the leading local reviews search engine, manipulates results when a business subscribes to a Business Owner Account[8].

The above examples show that it is not clear what role black hat search engine optimization plays in the online advertising ecosystem and whether it is necessarily detrimental. Our goal is to explore the economics of the SEO process and its effects on consumers, advertisers and search engines. By doing so, we are able to uncover whether such manipulation of ranking results is really harmful to consumers, and if so when. We are also able to provide recommendations to search engines and advertisers on how to optimally invest in or against SEO. We assume that the search engine's goal is to maximize consumer satisfaction by displaying the most relevant links to visitors, who derive utility from finding relevant, high quality sites. We show that the search rankings process is a multi-item contest among web sites. In particular, our model is equivalent to an all-pay auction with headstarts in which the link slots on a search page are allocated to websites. The ranking mechanism computes a *score* for each site that depends on the site's measurement of quality. The search engine then displays a list of search results sorted according to site scores.

Our model diverges from traditional rent seeking analysis in two components. First, the quality of a site depends on consumers tastes for relevance. While one consumer might find a site very relevant to a search query, another might find it less relevant than other sites. The heterogeneity and unavoidable changes in consumer tastes make it literally impossible for a search engine to measure a site's relevance with absolute accuracy. The result is noise in the scoring process used to rank websites. The existence of this noise results in suboptimal slot allocations when SEO is not allowed. The second divergence of our model from traditional models results from differences in intrinsic qualities among sites, as well as from the different valuation each site places on visits by consumers. These asymmetries yield different scoring headstarts for each site during the ranking process. A site that is on average more relevant for

[7]Baidu scandal makes it to CCTV: http://shanghaiist.com/2008/11/23/baidu_scandal_makes_it_to_cctv.php

[8]East Bay Express. http://www.eastbayexpress.com/eastbay/yelp-and-the-business-of-extortion-20/Content?oid=1176635

consumers than other sites will have an initial advantage over other sites during the ranking process, making the use of SEO techniques cheaper for that site.

We find that under certain conditions even black hat SEO can be advantageous to the search engine and can increase consumer welfare in equilibrium. In particular, if sites' valuation for traffic is aligned with their relevance (quality) then the search engine is better off when allowing some positive level of SEO than when discouraging or not allowing SEO. If, on the other hand, there are sites with high valuation for visits, but low relevance (quality), then SEO is generally detrimental to the search engine and consumer welfare. In other words, if the sites that are valuable to consumers value visitors highly then a positive level of SEO is beneficial, but if some sites that do not provide value to visitors can extract high profits from them then SEO is detrimental. An example of such a "bad" site, which are often called *spam sites*, is a site that advertises products for a very low price to lure visitors, but later on uses the visitors' credit card details for fraudulent activities[9].

According to our results SEO can be beneficial under some conditions by improving the rankings and getting the highest quality site on the top. However, there are potential pitfalls of allowing too high levels of SEO. We discuss some of these in Section 5. For example, it is not clear how SEO affects investment in content quality. We extend our model to incorporate a content investment stage and find that high effectiveness of SEO might result in underinvestment in content when creating content is relatively expensive.

Despite the apparent importance of the topic, there has been very little research done on search engine optimization. At the same time, search engine optimization has grown to become a multi-billion dollar business[10]. Many papers have focused on the sponsored side of the search page 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 given as exogenous, and the possibility of investing in SEO is ignored, although marketers often face the problem whether and how much to invest in SEO. Our results provide useful recommendations useful to firms that are involved in search engine marketing and to search engines. Furthermore, given that search engines are viewed by many today as a major gateway for information discovery, there is an

[9]Researchers estimate (Benczur et al. 2008) that 10-20% of Web sites constitute spam.

[10]See the survey conducted by seomoz.com at http://www.seomoz.org/dp/seo-industry-survey-results.

emerging debate on the fairness of search results and ranking algorithms, and the possibility of regulating search engines. We hope to contribute to this topic by exploring the effects of search engine optimization on the different players in the field.

The rest of the paper is organized as follows. Section 2 gives an overview of a small selection of this diverse literature and other research areas that are relevant to our work. In Section 3 we describe the main model and present the equilibrium outcomes of a simplified case in Section 4 with only one organic link. Next, we examine how the SEO game affects content investment in Section 5. Finally, Section 6 generalizes our model in several ways to show that our main results are robust, and introduces the notion of an ordered results list for comparison among search ranking mechanisms. All proofs and technical details appear in the Appendix.

2 Relevant Literature

The advent of online advertising technologies and the rapid growth of the industry led to an increase in the volume of research dedicated to this phenomenon. 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 impact advertising efficiency. Another major stream of research, including works by Edelman et al. (2007) and Varian (2007) focus mostly on the auction mechanism used by the different search engines to allocate their advertising slots. More recent examples, such as those by Chen and He (2006), Athey and Ellison (2009) and Aggarwal et al. (2008) analyze models that include both consumers and advertisers as active players. A full overview of the literature on sponsored search is out of the scope of this paper. Of special note, however, are papers that analyze different mechanisms to optimize allocation of sponsored links, either to maximize advertiser profits, search engine profits, or consumer welfare. Examples include Feng et al. (2007a), Feng et al. (2007b), Animesh et al. (2007), Chen et al. (2009) and many more.

A number of recent papers study the interplay between the organic list and the sponsored list. Katona and Sarvary (2010) show that the top organic sites may not have an incentive to bid for sponsored links. Xu et al. (2009) and White (2009) study how the search engine's advertising revenue from the sponsored links is affected by the organic listings.

Little attention was given to search engine optimization, although the use of SEO techniques is common practice. Works such as those by Pasquale (2006) Bracha and Pasquale (2007) and Mercadante (2008) consider the implications of search results manipulation using the traditional view that "cheating" has strictly negative results. Different options for regulation and the need (and legality) for it are examined. The economic implications of using bribes in contests is analyzed by Clark and Riis (2000). An important result is that allocative efficiency is not necessarily degraded by a bribery procedure, but might increase depending on the contest's parameters.

The work of Xing and Lin (2006) resembles ours the most by defining "algorithm quality" and "algorithm robustness" to describe the search engine's ability to identify relevant websites and eliminate non-relevant ones. Their paper shows that when advertisers' valuation for organic links is high enough, providers of SEO services are profitable, while search engines' profits suffer. Considering our result that using SEO can improve consumer welfare under noisy conditions, these results complement ours in explaining why search engines invest efforts in fighting SEO. 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. The model surprisingly shows that SEO should not exist as part of an equilibrium strategy.

A primary feature of our model is its equivalence to all-pay auctions when describing the game websites compete in for locations on search result pages. The process is a rent seeking process that is similar to lobbying and other processes described and analyzed in Hillman and Riley (1987) and other works. An extension of the all-pay model to multiple players and multiple items is analyzed in Barut and Kovenock (1998), Baye et al. (1996) and Clark and Riis (1998). For a survey of the literature on contests[11] under different information conditions and contest success functions, see Konrad (2007) and Sisak (2009).

Our use of all-pay auctions takes into account initial asymmetries among sites resulting from measurement error and different website qualities. The different qualities, measured as a relevance measure for consumers, translate into a headstart in the initial score calculated

[11]Contests are all-pay competitions among bidders where the bids do not reach the auctioneer

by the search engine to determine the auction winners. The existence of such a headstart, which in many cases is analogous to differences in abilities of the players, results in different equilibria as described in Kirkegaard (2009) and analyzed under more general conditions in Siegel (2009). Our application is unique in that it considers the cases where the initial headstart is biased by noise inherent in the quality measurement process. Krishna (2007) is one of the few examples taking noise into consideration in an auction setting. This noise is the main reason for the initially inefficient allocation of organic link slots, which can be corrected by allowing for SEO. We believe that by exploring a topic that has essentially not been studied before and by enhancing the above mentioned all-pay auction models, our paper fills a gap on the boundary of the marketing and information systems literatures.

3 Model

A search engine (SE) is a website that provides the following service to its visitors: they enter queries (search phrases) into a search form and the SE returns a number of results for this query displaying them in an ordered list. This list contains a number of links to other websites in the order of the relevance of their content for the given search phrase. In our model, we focus on a single keyword and we assume that the *relevance*, or *quality* of a search result is essentially the probability that a consumer is satisfied with the site once clicking on the link[12]. We further assume that, for the purpose of ordering the search results, the SE's objective is to maximize the expected consumer satisfaction[13] hence its goal is to present the most relevant results to its visitors, and its utility is equal to the expected satisfaction level of consumers.

In order to rank websites, the search engine uses information gathered from crawling algorithms and data mining methods on the Internet. Let q_i denote the relevance of site i in the context of a given keyword. It is reasonable to assume that the search engine can only measure quality with an error, and cannot observe it directly. The initial quality score that

[12]Modeling consumer satisfaction as a 0-1 variable is relatively simple, but captures the essence. In Section 5 we will discuss alternative formulations where q_i is the average utility that a consumer gains after clicking on link i.

[13]Since providing these results is typically the search engines core service to consumers its reputation and long term profits strongly depend on the quality of this service.

the SE assigns to a site i is thus $s_i^S = q_i + \sigma\varepsilon_i$, where ε_i are assumed to be independent and are drawn from the same distribution. If the Web sites do not take any action the results will be ordered according to the s_i^S's as assigned by the search engine. If, however, Web sites can invest in SEO[14], they have the option to influence their position after observing the initial scores. The effectiveness of SEO is measured by the parameter α in the following way. If site i invests b_i in SEO, its final score becomes $s_i^F = s_i^S + \alpha b_i$. That is, depending on the effectiveness of SEO, sites can influence their score to a varying extent which in turn determines their final location in the organic list of search results. We assume that there are n websites providing informational content or products to consumers and that those sites derive some utility from the visiting consumers. We index the sites in decreasing order of their quality: $q_1 \geq q_2 \geq \ldots q_n$. The sites' profits primarily depend on their traffic. We assume that site i derives utility $v_i(t)$ of having t customers click its link in a given time period.

The behavior of the unit mass of consumers in our model is relatively simple. If a consumer is presented with one link to site i, s/he clicks on it, and is satisfied with probability q_i, receiving a utility normalized to 1 if satisfied. If there are $k > 1$ links, the consumers traverse the list of links sequentially in the order presented to them by the SE. When a consumer is satisfied with a site s/he visits following a link, the searching ends. When a consumer is dissatisfied with the j^{th} site visited, s/he continues to the next link with probability $c_j < 1$. This latter parameter captures the fact that consumers do not search forever, but they quit if they do not find what they are looking for after a while. An alternative formulation that would lead to a similar process is if we assume that consumers incur a small cost every time they click on an additional link. We provide full details of the search sequence by consumers in Section 6.2 dealing with multiple search results.

The sequence of the game is as following. First the search engine measures the relevance of each website and publishes $s_i^S = q_i + \sigma\varepsilon_i$. Next, the websites, after observing s_i^S, simultaneously decide how much they want to invest in SEO, changing the scores to $s_i^F = s_i^S + \alpha \cdot b_i$. The search engine then recalculates the scores and displays an ordered list of search results sorted in a decreasing order of the final site scores s_i^F. Finally, visitors click on the results

[14] We do not explicitly model who conducts the SEO activities. It could be the site itself, a third party, or in the extreme case, the search engine itself.

according to their order until being satisfied, and payoffs are realized at the end. Our assumption on the timing of the above events is somewhat simplistic, but it is the most plausible way of capturing Web sites' reactions to their ranking results and their subsequent investment in SEO.

We start our analysis by examining a simple case that illustrates the main forces governing SEO. In this case, we assume that there is one organic link displayed on the SE ($k = 1$) and that there are two bidders ($n = 2$) with $q_1 > q_2$. We generalize to the case of $n > 2$ sites, and multiple $k > 1$ links in Section 6.

4 SEO Equilibrium - One link

We assume that there is one organic link, and that the utility sites derive from incoming traffic is linear in traffic, such that $v_i(t) = v_i t$. Since there is a unit mass of consumers that click on the link displayed in the search result, the valuation that sites have for the appearing on the list is v_1 and v_2, respectively. We set the distribution of ε_i to take the value of either 1 or -1 with equal probabilities. We assume $\sigma > |q_1 - q_2|/2$ to ensure that the error can affect the ordering of sites [15].

First, as a benchmark, let us examine the case in which search engine optimization is *not* possible, i.e., when $\alpha = 0$. In this case sites cannot influence their position among the search results. Since $q_1 \geq q_2$, the SE's expected utility is $\frac{3}{4}q_1 + \frac{1}{4}q_2$. The effect of the error in the SE's measurement process is clear. With a certain probability (1/4 in this case), the order will be suboptimal leading to a drop in expected utility compared to the first best case of q_1.

When search engine optimization is effective, i.e., when $\alpha > 0$, websites have a tool to influence the order of results. The ability to influence, however, is typically asymmetric, since sites have different starting scores s_i^S. A site that is in the first position in the SE's initial ranking has a headstart and hence can remain the first even if it invests less in SEO than its competitor. Another characteristic of the game sites play is that their SEO investment is sunk no matter what the outcome of the game is. Our first result states that sites essentially participate in an all-pay auction with headstarts

[15] Otherwise the error never changes the order of results and the setup is equivalent to one with no error

Lemma 1 *The game that sites play after observing their starting scores is equivalent to an all-pay auction with headstarts.*

These games are generalizations of basic all-pay auctions without headstarts. In these auctions players submit bids for an object that they have different valuations for. The player with the highest bid wins the object, but all players have to pay their bid to the auctioneer (hence the term "all-pay auction"). If players have headstarts then the winner is the player with the highest score - the sum of bid and headstart (see the Appendix and Kirkegaard (2009) for details).

The level of headstart in our model depends on the starting scores and hence on the error. For example, if $q_1 > q_2$ and $\varepsilon_1 = \varepsilon_2 = 1$, the error does not affect the order (which is $q_1 \geq q_2$) nor the difference between the starting scores $(q_1 - q_2)$. Since SEO effectiveness is α, an investment of b only changes the scores by αb, therefore the headstart of site 1 is $\frac{q_1 - q_2}{\alpha}$. As the size of the headstart decreases with α, the more effective SEO is, the less the initial difference in scores matters. Even if site 1 is more relevant than site 2, it is not always the case that it has a headstart. If $\varepsilon_1 = -1$ and $\varepsilon_2 = 1$ then $s_1^S = q_1 - \sigma < s_2^S = q_2 + \sigma$ given our assumption on the lower bound on σ. Thus, player 2 has a headstart of $\frac{q_2 + 2\sigma - q_1}{\alpha}$. By analyzing the outcome of the all-pay auction given the starting scores, we can determine the expected utility of the SE and the websites.

All-pay auctions with complete information typically do not have pure-strategy Nash-equilibria, but the unique mixed strategy equilibrium is very intuitive. In a simple auction with two players with valuations $v_1 > v_2$, both players mix between bidding 0 and v_2 with different distributions[16]. Player 1, having the higher valuation, wins with the higher probability of $v_1/2v_2$ and player 2's surplus is 0. Thus, only the player with the highest valuation makes a positive profit in expectation, but the chance of winning gives an incentive to the other player to submit positive bids. In the case of an all-pay auction with headstarts the equilibrium is very similar and the player with the highest potential score (valuation plus headstart) wins with higher probability and the other player's expected surplus is 0. The winner's expected surplus is equal to the sum of differences in valuations and headstarts.

[16]See the Appendix for detailed bidding distributions.

4.1 The effect of SEO on efficiency and consumer welfare

To examine the outcomes of the SEO game, we use $E(\alpha) = E(\alpha; \sigma, v_1, v_2, q_1, q_2)$ to denote the efficiency of the auction. In this case it is simply the probability that the player with the more relevant link wins the auction (that is, player 1). Note that the efficiency coincides with the search engine's objective function as it wants the more relevant link to come up first. The payoff of the search engine is a linear function of the efficiency:

$$\pi_{SE} = q_2 + (q_1 - q_2)E(\alpha)$$

If there is no SEO, that is when $\alpha = 0$ (and $\sigma > |q_1 - q_2|/2$), we have $E(0) = 3/4$. Our goal is therefore to determine whether the efficiency exceeds this value for positive α SEO effectiveness levels. It is useful, however, to begin with analyzing how the efficiency depends on valuations and qualities for given α and σ values. The following Lemma summarizes our initial results.

Lemma 2 *For any fixed α and σ, $E(\alpha; \sigma, v_1, v_2, q_1, q_2)$ is increasing in v_1 and q_1 and is decreasing in v_2 and q_2.*

Thus, the efficiency of the ranking increases when the most relevant site becomes even more relevant and also when its valuation for clicks increases. When there is no SEO, i.e., when $\alpha = 0$, the Lemma holds because the efficiency simply does not change with v_1, v_2, q_1, q_2, but when $\alpha > 0$ the efficiency strictly increases and decreases in the respective variables. In essence, the Lemma tells us that no matter how effective SEO is, the less sites valuations are aligned with their relevance levels, the less efficient the rankings are.

The following proposition summarizes the main result of our paper, showing how SEO affects the efficiency of the ranking.

Proposition 1

1. *For any $\sigma > |q_1 - q_2|/2$, there exists a positive $\hat{\alpha} = \hat{\alpha}(\sigma, v_1, v_2, q_1, q_2)$ SEO effectiveness level such that $E(\hat{\alpha}) \geq E(0)$.*

2. *If* $v_1/v_2 > 3/2$ *then for any* $\sigma > |q_1 - q_2|/2$*, there exists a positive* $\hat{\alpha} = \hat{\alpha}(\sigma, v_1, v_2, q_1, q_2)$ *such that* $E(\hat{\alpha}) > E(0)$*.*

3. *If* $v_1 < v_2$ *and* $\sigma \geq \frac{v_2+v_1}{v_2-v_1}\frac{q_1-q_2}{2}$ *then for any* $\alpha > 0$ *we have* $E(\alpha) \leq E(0)$*.*

The first part of the proposition tells us that for any level of error there is a positive level of SEO that does not reduce the efficiency of the ranking. Practically, if the level of SEO effectiveness is very low (that is, its cost is high) then firms will not invest and the ranking will not be altered.

The second point yields more interesting results. Essentially, it demonstrates that positive levels of search engine optimization do improve the efficiency of the ranking in some cases. When high quality sites value visitors relatively high compared to lower quality sites, SEO is beneficial to both the search engine and consumers regardless of the level of error. The third part states the opposite: when the player with the lower quality has a higher valuation then all positive levels of SEO are detrimental to the efficiency (unless the error is very small). Although our model assumes given deterministic valuations and qualities, one could imagine a setting where these parameters are randomly drawn from a given distribution. In that case, our results yield that the higher the correlation between a site's quality and valuation, the more likely that SEO improves efficiency [17].

The intuition for the results is as follows. The SEO mechanism favors bidders with high valuations. Since the SE cannot perfectly measure site qualities, this mechanism corrects some of the error when valuations increase monotonically in quality. When lower quality sites have high valuation for traffic, however, SEO creates incentives that are not compatible with the utilities of consumers or the search engine. In this latter case, the high valuation sites which are not relevant can get ahead by investing in SEO unless the error is relatively small[18]. Examples are cases of "spammer" sites that mislead consumers. In these cases consumers do not gain any utility from visiting such sites, but the sites may profit from

[17] If the joint distribution of v_i and q_i is lognormal then one can show that the condition for part 2 is more likely to hold when the correlation between v_i and q_i is higher.

[18] When σ is close to $\frac{q_1-q_2}{2}$ the efficiency can slightly exceed 3/4 for some small $alpha > 0$ values. This happens because ineffective SEO gives player 1 the chance to win in cases when the error favors player 2 while not affecting the other cases.

consumer visits.

To better illustrate our results we consider two examples. Using SEO is very common in the online news industry where newspapers, magazines and blogs compete for visitors. As many content sites derive their income from displaying advertisements to their visitors, they compete for catching a larger share of visitors for their content. In many cases, the content of news items is not highly differentiated among those sites, and other techniques are used to lure visitors. One technique commonly used by blogs is a mutual exchange of site links through “trackbacks”. Whenever a blog post quotes an item from another blog by linking to it, the quoted blog reciprocates by adding an outgoing link to the quoting blog, called a “trackback”. The technique exploits the fact that search engine ranking processes use incoming links as a measure of site popularity and quality. The use of trackbacks is so prevalent that all major blogging software automatically handle their creation. Despite the wide use of these SEO techniques news search is one of the emerging areas in the search market. This is potentially a result of the fact that sites that offer “fake” or unoriginal news that is not valuable to consumers are not able to attract advertisers resulting in low profitability and hence a low valuation for links. Only sites that provide useful content are profitable enough to be able to afford the SEO investments needed to get to one of the top positions.

Another industry where SEO is commonly used is the e-commerce industry. Many online shops compete for selling the same standard products, with minimal differentiation in costs, prices and quality of service. In order to gain more visits by consumers, shops create multiple product pages for the same product called *landing pages*. Each landing page contains different wordings for describing the same product, hoping to match more visitor search queries. As opposed to the news industry, e-commerce produces many spamming sites. In some product categories, sites may be able to mislead consumers and even if a small proportion of them buys a low quality product, profits and thus valuation for traffic can be quite high. This phenomenon naturally depends on the product category, but in the ones where repeat purchases are rare and where consumers are inexperienced in judging product quality, SEO can be detrimental to consumers and high quality outlets.

As these two examples illustrate, the effects of SEO highly depend on the ecosystem in the industry. Our result describe the outcomes in different cases and show that in many cases the SE is better off allowing some positive level of SEO. It is not clear, however, what the optimal level of SEO is in these cases. In particular, how does it depend on the variance of the measurement error? To answer this, let $\hat{A}(\sigma)$ denote the set of α SEO effectiveness levels that maximize the search engine's utility function. For two sets $A_1 \subseteq \mathbb{R}$ and $A_2 \subseteq \mathbb{R}$, we say that $A_1 \succeq A_2$ if and only if for any $\alpha_1 \in A_1$ there is an $\alpha_2 \in A_2$ such that $\alpha_2 \leq \alpha_1$ and for any $\alpha_2' \in A_2$ there is an $\alpha_1' \in A_1$ such that $\alpha_1' \geq \alpha_2'$.

Corollary 1 *If $v_1/v_2 > 3/2$, then the optimal SEO effectiveness is increasing as the variance of the measurement error increases. In particular, for any $\sigma_1 > \sigma_2 > 0$, we have $\hat{A}(\sigma_1) \succeq \hat{A}(\sigma_2)$.*

We have already shown that SEO can be beneficial because it can serve as a mechanism that corrects the search engine's error when measuring how relevant sites are. The above corollary tells us that if the error is higher more effective SEO is required to correct the error. This suggests, somewhat counter intuitively, that investments against SEO on the SE's part complement investment in better search algorithms, and do not substitute them. That is, only search engines that are already very good at estimating true qualities should fight hard against black hat SEO. Nevertheless, measurement error is not always under the control of the search engine, but can depend on external factors and vary from keyword to keyword. Therefore, it may make sense to allow higher levels of SEO in areas where the quality measurement is very noisy. Another factor to consider is that although higher levels of SEO may improve the ranking, the investments sites make in it mostly end up at firms offering SEO services and not the SE. One solution to this problem is to allow sites to pay for placing links among organic results, similar to what sites such as *Baidu.com* or *Yelp.com* do. There are arguably many potential pitfalls to including paid links as part of the organic results and we do not suggest that this is always optimal, but our model does offer a plausible explanation as to why some search engines may do so.

4.2 The effect of SEO on advertiser profits

When there is a positive level of SEO effectiveness, sites have a natural incentive to invest in SEO, which they do not have when $\alpha = 0$. In the extreme case of $\alpha \to \infty$ the difference in initial scores dissipates and the game becomes a regular all-pay auction. If, for example, $v_1 > v_2$ then player 1's expected payoff is $v_1 - v_2$, whereas player 2 makes nothing in expectation. Comparing this to the case in which there is no SEO - player 2 making $v_2/4$ and player 1 making $3v_1/4$ - reveals that player 2 is worse off with SEO whereas player 1 is better off iff $v_1 > 4v_2$. This implies that high levels of SEO only increase profits for sites with outstanding valuations. The following corollary provides detailed results on the sites' payoffs.

Corollary 2

1. *If $v_1 > v_2$ then Player 2's payoff is decreasing in α.*

2. *If $v_1 > v_2$ there always exists an $\alpha^* > 0$ such that Player 1 is better off with an SEO effectiveness level of $\alpha = \alpha^*$ than with $\alpha = 0$. If $v_1 > 4v_2$ or $\sigma < \frac{v_1}{v_2}\frac{q_1-q_2}{2}$ then Player 1 is strictly better off.*

The player with the lower valuation is therefore worse off with higher SEO. Player 1, on the other hand, is better off with a certain positive level of SEO, especially if its valuation is much higher than its competitor's and if the measurement error is small. The intuition from the former follows from the fact that higher levels of SEO emphasize the differences in valuations, and the higher the difference the more likely that the higher valuation wins. For the latter condition, smaller measurement errors make it easier for the player with the higher starting score to win and to take advantage of SEO. The corollary shows that the player with the higher valuation is generally happy with some positive level of SEO, but further analysis (see the proof) suggests that it is not clear whether Player 1 or the search engine prefer a higher level of SEO.

5 Investing in Quality of Content and SEO

So far we have focused on the investment that sites can make to improve their ranking without affecting their relevance. We did not, however, consider potential investments in content quality. In reality, sites have several different tools to change their content and to affect their rankings. Some of these tools, such as the aforementioned black hat methods, clearly do not change content quality while other tools purely improve content and change rankings indirectly through real quality improvement. In general, it is difficult to draw a clear line between the two types since many activities, such as white hat methods, fall between the two. In this section, we separate content investment from SEO and investigate how it changes our main results. We acknowledge that since the common use of the term SEO incorporates many different types of activities, some of them (especially those referred to as white hat) would fall under the content investment category in our model. In our definition, an investment that purely changes content quality and *only* affects ranking through content quality is a content investment.

To incorporate content investments into the model, we extend the game and add a content investment stage before the SEO stage. In this first round, sites can decide how much they want to spend on improving their quality of content and given these quality levels they decide how much to invest in SEO as in Section 4. All other assumptions are the same: two sites are competing for one organic link. Let c denote the marginal cost of increasing quality[19].

There are two ways how an investment in content quality could affect sites in the SEO stage. First, an increase in q_i increases the chance of the link being displayed on the top of the organic list by the search engine, with or without SEO. Second, it can change the valuation sites place on visits by consumers. In our basic setup v_i denotes the valuation for the link, which is the revenue accumulated across all of the unit mass of consumers visiting the site, whether satisfied or dissatisfied with the content presented to them. It is reasonable to assume, however, that the revenues are derived from the satisfied users, and that an increase in satisfaction levels of users also increases the total value derived by the site per unit mass

[19] We assume that content costs are linearly increasing, however, a convex cost function would yield similar results

of consumers. Therefore, investment in quality can also increase the valuation of the site getting the top link. We first ignore this effect and focus on the case in which valuation is not affected by the quality investment.

5.1 Fixed Valuations

Here, we solve a game in which sites' valuations for getting to the top position are fixed and not affected by the quality investment. As a benchmark, let us consider the case in which there is no SEO, i.e., $\alpha = 0$. For simplicity we also assume that $\sigma = 0$, that is, there is no measurement error in this benchmark case. Since there is no SEO, the game becomes a one-shot game, essentially an all-pay auction, where the winner gets all the benefits. As we mentioned before, these games do not have pure-strategy equilibria, but the mixed strategy equilibria are very intuitive. The site with the higher valuation (e.g. player 1) wins the auction with a higher probability $(1 - v_2/2v_1)$ and has an expected payoff of $v_1 - v_2$, whereas the other player has an expected payoff 0. To make the analysis simple, we examine the case when $v_1 = v_2 = v$. In this case the players win with equal probability, make 0 payoff and each invest $v/2c$ in expectation.

Now let us examine the case in which sites first have the option of investing in content, then in SEO. The exhaustive description of the mixed strategies would be too complex, therefore we focus on the symmetric sub-game perfect equilibrium with pure strategies in the first stage. This allows us to point out the differences that the content investment stage makes without determining all the mixed strategy equilibria.

Proposition 2 *The only possible equilibrium with pure strategies in the first stage is the one in which sites do not invest in content but will invest in SEO and earn an expected payoff of* $\min(\sigma/2\alpha, v/4)$*. The above mentioned is an equilibrium if and only if (*$c > \frac{v\alpha - 5\sigma}{2v\alpha - 4\sigma}\frac{1}{\alpha}$ *and* $\sigma < v\alpha/4$*) or (*$c > \frac{3}{4}\frac{1}{\alpha}$ *and* $v\alpha/4 \leq \sigma < v\alpha/2$*) or (*$c > \frac{1}{2}\frac{1}{\alpha}$ *and* $v\alpha/2 \leq \sigma$*)*

Generally, if the cost of investing in content is high relative to the effectiveness of SEO then sites will give up their investments in quality and instead will focus on search engine optimization. Note that this critical cost level is decreasing in SEO effectiveness, therefore the more effective search engine optimization is the more likely that sites will not improve their content. In all the cases, when $c > 1/\alpha$ we get the no investment equilibrium. Since the inverse of SEO effectiveness is essentially its cost we obtain the result that if content is more expensive than SEO then sites will not invest in it[20]. In particular, to compare with the benchmark case when $\sigma = 0$, we get that sites do not invest in content iff $c > \frac{1}{2}\frac{1}{\alpha}$, that is, the critical cost above which sites substitute SEO for content is half of the SEO cost. If content costs are low then sites will naturally improve their content, but there are generally no equilibria with pure strategies in the first stage, which is natural given the all-pay auction structure.

The intuition behind the results is straightforward. If the cost of content is high compared to SEO then sites will not attempt to get a headstart for the SEO game, they will instead compete with their SEO investments directly. If, on the other hand, content costs are low, they may engage in investing in content to get a headstart for the SEO game, where they do not have to invest heavily. We do not explore in detail what happens if c is low because of the complexity of the game. There are certainly no equilibria with pure strategies in the first case, but the game becomes similar to an all-pay auction where, presumably, an equilibrium in mixed strategies exist with positive expected content investment. The results thus show that having a highly effective SEO level can result in underinvestment in quality which hurts consumers and the search engine in the long run.

In the above setup we restricted the benefits of content investment to improving the position in the contest for the organic link and did not consider the additional effects of making more profit per visitor. In what follows, we investigate how the possibility of SEO affects content investment when such investment improves profitability in addition to attracting more visitors.

5.2 Valuations Increasing with Quality

We proceed with analyzing the case when an investment in content also increases sites' valuation for a link. There are several possible functional forms to capture this relationship. We employ the most parsimonious form by assuming that $v_i = \underline{v} + q_i m$, where $\underline{v}$ and m

[20]Note that this is not a binding condition. The Proposition shows that in many cases even if content is cheaper than SEO sites will not invest in it.

are site-independent positive parameters. Here, $\underline{v}$ can be interpreted as the sites' baseline valuation of customers, whereas m corresponds to the margin gained on a satisfied customers. With this formulation, it is obvious that a monopolist would invest in content iff $c < m$. However, if there are two players they may compete even if $c \geq m$, incentivized by the fixed portion of the reward. In the benchmark case of no SEO (and $\sigma = 0$), the second stage of the game becomes completely determined and the first stage is an all-pay auction with variable rewards. One can determine that in equilibrium players mix between 0 and $\frac{v}{c-m}$ with a mean investment of $\frac{v}{m}\left(-\frac{c}{m}\ln(1-m/c)-1\right) > 0$. That is, no matter how high c is compared to m sites invest in content. Now let us examine the situation with positive levels of SEO in the second stage.

Proposition 3 *When $v_i = \underline{v} + q_i m$, the only possible equilibrium with pure strategies in the first stage is the one in which sites do not invest in content. The above mentioned is an equilibrium if $c > \frac{1}{\alpha} + m$.*

These results are very similar to that of the case with fixed valuations in that if content quality is expensive enough then sites will forfeit the opportunity to invest in it. An important difference is that since the quality investment will also change valuations, sites will take that into account when comparing its costs to SEO. It is useful to note that the critical value for c above which this happens is generally decreasing and often lower than $\frac{1}{\alpha} + m$. Therefore the more effective SEO sites expect in the second stage the more reluctant they will be to invest in content because of anticipating the wasteful spending on SEO.

In summary, the results in this section highlight the fact that search engines should be careful when allowing some level of SEO activity to take advantage of its forces that improve the ranking, because at the same time it could discourage sites from investing in content, thus hurting consumer satisfaction in the long run.

6 Model Generalizations

6.1 Single link, multiple sites and arbitrary error distribution

In this section we show that our main results are robust under more general assumptions. First, we relax the assumption on the distribution of the search engine's measurement error and allow more than two sites to compete for one organic link. We assume that there are n websites that are considered by the search engine for inclusion in the organic list (consisting of one link) with $q_1 > q_2 > \ldots > q_n$. All the other assumptions are identical to those in Section 4. Regarding the error ε_i , we allow its distribution to be arbitrary with a mean of zero and a finite variance normalized to 1. Similarly to Section 4, we assume that the error is large enough that it makes a difference, that is, we assume that $\varepsilon_1 - \varepsilon_2$ can take a value of less than $q_2 - q_1$ with positive probability. The following proposition shows that even in this case SEO can improve the efficiency of the auction if the valuations of the most relevant sites are high enough.

Proposition 4 *For any σ there exist a $\hat{v}(\sigma) > v_2$ and an $\hat{\alpha}(\sigma) > 0$ such that if $v_1 > \hat{v}$ and $v_2 > v_i$ for every $i \geq 3$ then $E(\hat{\alpha}) > E(0)$.*

This result generalizes our results in Proposition 1. If the valuations of the two most relevant sites are high enough then the rest of the sites are in double disadvantage due to the low starting scores and their lower valuations. This will lead to only the first two sites investing in SEO for high enough effectiveness levels. The competition of these two sites is similar to that in Section 4: If the error does not reverse their starting scores compared to their true relevance, then site 1 has a high chance to win keeping the right order. If, however, the ranking is reversed due to the error then although site 2 might have a good chance to get into the first position, for high SEO effectiveness levels the higher valuation of site 1 limits this probability.

6.2 Multiple links

In this section we analyze the case of search results comprised of multiple links. We begin by considering how consumers behave when presented with an ordered list of results. The reason is that our assumption that consumers will certainly click any link on the list does not hold anymore. It is also not obvious what order of search consumers will choose, and what type of utility they experience when being dissatisfied with a visited site.

Suppose the search engine assigns n websites with qualities $q_1 > q_2 > \ldots > q_n$ to n links, and presents an ordered list to consumers. If the qualities of sites were fully observable, either directly or through a truth-telling mechanism, the search engine could sort the list in descending order of site qualities. Previous research by Athey and Ellison (2009) and Chen and He (2006) shows that consumer surplus increases when presented with a fully sorted list compared to a randomly ordered list, while Aggarwal et al. (2008) show that under a Markovian consumer search model, similar monotonic click probabilities arise in the optimal assignment.

In the organic results case, however, the ranked list produced by the search engine is different in two important manners. First, the fact that the search engine has errors in its calculation of scores inhibits its ability to fully sort sites according to their qualities. Second, when SEO is not effective ($\alpha = 0$), the resulting list is not completely random. We therefore require a structure that allows the comparison of welfare derived from differently sorted list.

Formally, we define a structure on ordered lists that allows a measurement of their distance from being optimally sorted. An *ordered list* of n items is a collection of n random variables $Q = (f_1, \ldots, f_n)$. The variable f_j contains the distribution of qualities of sites appearing in location j in the list. In a completely random list, for example, $f_j(q_i) = 1/n$ for every j and i, while in a fully sorted list, which we denote Q^S, $f_j^S(q_i) = 1$ for $j = i$ and zero otherwise.

When SEO is not effective in our game, the distribution of scores for each site has a different mean, where sites with higher qualities have higher score means, and thus higher probability of appearing higher on the ranked list. The result is that the initial ordered list displayed by the search engine in not completely random. When SEO is effective, however, more relevant sites have an even higher probability of appearing higher on the ranked list, as is shown below. The result is an ordered list that is, in a sense, closer to the first best fully sorted list, but can never match it because of the noisy process. A common claim in the literature is that sponsored link auctions improve social and consumer welfare tremendously.

We believe the benchmark case for such auctions should not be randomly ordered lists, but rather lists resulting from organic search rankings. Given these lists are not random, the change in welfare decreases. Nevertheless, we leave exploration of the magnitude of this welfare change for further research.

For our purposes, the structural representation of ordered lists would allow for a comparison of two lists according to their distance from the fully sorted list in terms of the expected consumer surplus generated by a search on each list. The actual calculation and comparison of the surpluses, however, is complex for general distributions of errors, values and qualities. A more natural measure of distance of ordered lists is to sum the sum-square distances of each f_j from the fully sorted f_j for all j, yielding

$$\delta = \sum_{j=1}^{j=n} \sum_{i=1}^{i=n} [f_j(q_i) - f_j^S(q_i)]^2.$$

We conjecture that lists with a smaller distance measure yield higher expected consumer surplus. In what follows, we use several simplifications to show that consumer surplus increases when SEO is effective.

We assume consumers traverse ordered lists in a sequential manner, starting with the link at location 1. When encountering a link in position j, consumers click it and visit the page it leads to, being satisfied with probability $\bar{q}_j = \mathbb{E}(f_j)$.[21] If the consumers are not satisfied with the result, they move on to the next link with probability $0 < c_j < 1$, or stop otherwise. The next link is then clicked, and the consumer is satisfied with probability $\bar{q}_{j+1}$ and so forth. The consumer's expected utility from traversing the first k links is thus

$$E_k(Q) = \sum_{j=1}^{k} [c_j \cdot \mathbb{E}(f_j) \cdot Pr(z_1 = \ldots = z_{j-1} = 0)]$$

when z_j is the probability that link j was searched and was found unsatisfying.

The ranking process, followed by the search process, is analogous to the search engine performing a sequential all-pay auction of the links as described in Clark and Riis (1998), where in each stage a single link is auctioned using the game described in section 3, and is

[21]The qualities of sites are probabilities, thus the expected value of f_j is the expected satisfaction from visiting sites in location j on a list

presented to the visitor. If the visitor is not satisfied with the result, the winning website of the previous auction is removed from the list of competitors, and the auction of another link commences again. We impose two simplifications to analyze the game. First, we assume that the search engine picks a website to display for link j from f_j independently. This would mean that the case of a website being displayed twice in a ranked search process might happen with some probability. This probability decreases with the SEO effectiveness parameter. Second, we assume that in the auction process for link j, only the weights on q_j and q_{j+1} are updated, both for the distribution f_j and the distributions of lower links on the list, $f_{j+1} \ldots f_n$.

Proposition 5 extends our previous results to the multiple links case, and shows that when noise exists in the ranking mechanism, SEO can improve consumer's welfare:

Proposition 5 *Let $Q(\alpha) = (f_1(\alpha), \ldots, f_n(\alpha))$ be the ranked list resulting from n sequential all-pay auctions with effectiveness level α, where in auction j the sites with quality scores q_j and q_{j+1} are the sole competitors, then for any σ and for any $\varepsilon_i \sim i.i.d(0, \sigma)$, $1 \leq i \leq n$ there exists a $\Delta_v(\sigma) > 0$ and a positive $\hat{\alpha}(\sigma)$ such that if $v_i > v_{i+1} + \Delta_v(\sigma)$ for all $1 \leq i \leq n$ then $E_k(Q(\hat{\alpha})) > E_k(Q(0))$ for every $1 \leq k \leq n$.*

Proposition 5 says that if the valuation sites place on visitors are spread enough, then a sequential game of presenting one link at a time to the consumer while allowing for SEO can increase the consumer's satisfaction. The intuition behind the result is that each time a consumer is dissatisfied with a link but chooses to continue, the search engine can just run an auction for one link with the remaining competitors, and present it to the consumer. The game is then reduced to multiple repetitions of our single link game.

An interesting result of our formulation is that consumer welfare increases no matter how many links are searched before stopping. In a fully sorted list, however, assuming that a consumer would *always* search through the first k links means that any order of the first k links will yield the same consumer welfare, as these are k independent Bernoulli trials. There is significant evidence, however, that a consumer's probability of visiting a site (also called a click-through-rate, or CTR) is dependent on the site's location on the ranked list, as well as on other results displayed on the list. Several theoretical models explain this phenomenon, but typically show that the probability of visiting any site is monotonically decreasing in its distance from the top of the list. Using empirical data, Ghose and Yang (2009), Jeziorski and Segal (2009) and others have shown that this dependence is not straightforward, and that CTRs do not necessarily decrease in a monotonic way across a ranked list. Our assumption that consumers flip a coin after unsuccessful searches to decide whether to continue or not consolidates the majority of the evidence with our improved efficiency result.

7 Conclusion

We model search engine optimization as an all-pay auction where sites can invest in improving their search ranking without changing their link's relevance. We find that some level of SEO can be useful to the search engine and its customers because it acts as a mechanism that improves the rankings by placing sites with high valuations for the links higher on the results list. In general, if sites' valuations for consumers are aligned with how much utility consumers gain when visiting them then SEO is beneficial to both the search engine and consumers. Participating sites, on the other hand, might be worse off as they carry the extra burden of having to invest in SEO, whereas if search engine optimization does not exist they do not have to make additional effort. In the case when a low quality site has high valuation for traffic, it can benefit from the presence of SEO and get to an undeserved, better position, hindering search engine payoffs and consumer benefits.

In general, we find that when the search algorithms of the SE are less accurate, higher levels of SEO are more likely to be beneficial. Nevertheless, money that web sites spend on SEO is, to some extent, wasteful spending and could affect sites behavior in different ways. For example, they could spend less on sponsored links, hurting SE profits, or could invest less in content. We thoroughly investigate how the presence of SEO affects sites incentives to invest in quality improvement and find that if the marginal cost of content is high then sites may underinvest as a consequence of the presence of SEO. This phenomenon of underinvestment is more likely with higher SEO effectiveness levels.

Our paper has important practical implications. Contrary to the popular belief, allowing sites to invest in improving their ranking without improving their relevance can be beneficial to the search engine and the consumers, but can hurt the top sites even if they end up high eventually. Our result explain why some search engines seem to work very hard to reduce the possibility and effectiveness of SEO, while others like Baidu or Yelp even offer such services themselves. Our results suggest that when the search algorithms are not very accurate or when SEO methods are hard to identify, allowing some level of black hat SEO does not necessarily compromise the results. Despite the potential advantages, search engines should be careful not to allow SEO to be too effective, so that sites do not invest in SEO instead of content or sponsored links. If sites invest too much in wasteful SEO spending they might cut back on content quality hurting consumer and search engine benefits in the long run as the funds sites spend on SEO are not transferred to consumers in matters of content improvement. Our results also provide important recommendations to websites that are competing for top organic links. Contrary to popular belief, sites engaging in SEO are not only sites that wish to achieve a better position than their true quality merits, but also top quality sites that need to defend their position.

We believe that the economics of search engine optimization is a topic of high importance to both academics and practitioners. In this paper we examine the basic forces of this interesting and unusual ecosystem. Given the complexity of the problem, our model has a number of limitations that could be explored by future research. First of all, 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. We also assume that the number of consumers that visit the search engine is fixed and does not depend on the quality of the results. This might not be true in a competitive environment where search engines with inferior links are likely to receive fewer visitors. However, this effect would reinforce our results in the following way. Lower quality sites would realize that, on the long run, acquiring top links via SEO would reduce the overall quality of the search engine, resulting in fewer visitors. This would essentially decrease low quality sites' valuation for top links. Finally, we do not explicitly examine how search engines can invest in reducing SEO. Our results yield that under certain conditions having some SEO is beneficial, but we do not determine how much it is worth investing against SEO when it is too effective and is detrimental. Depending on the costs of reducing SEO, it might be unprofitable for the search engine to do so.

Appendix

PROOF OF LEMMA 1: We decompose the final scores of both sites into a headstart h and a bid as follows: $\tilde{s}_1^F = h + b_1$ and $\tilde{s}_2^F = b_2$ where $h = \frac{s_1^S - s_2^S}{\alpha}$. The decomposed scores have the property that for every b_1, b_2 $\tilde{s}_1^F \geq \tilde{s}_2^F \iff s_1^F \geq s_2^F$ and thus preserve the outcome of the SEO game. Since the investments are sunk and only the winner receives the benefits (with the exception of a draw) the SEO game is equivalent to an all-pay auction with a headstart of $h = \frac{s_1^S - s_2^S}{\alpha}$. In the following, we present the solution of such a game to facilitate the presentation of the remaining proofs. $\square$

SOLUTION OF ALL-PAY AUCTIONS WITH HEADSTARTS: As derived by Kirkegaard (2009), the generic two player all-pay auction with headstarts has a unique mixed strategy equilibrium. When players valuations are $v_1 \geq v_2$ and player 1 has a headstart of h then s/he wins the auction with the following probabilities:

$$W_1(h) = Pr(1 \text{ wins}|h \geq 0) = \begin{cases} 1 & h > v_2 \\ 1 - \frac{v_2}{2v_1} + \frac{h^2}{2v_1 v_2} & h \leq v_2 \end{cases}$$

$$W_1(h) = Pr(1 \text{ wins}|h < 0) = \begin{cases} 1 - \frac{v_2}{2v_1} & h \geq v_2 - v_1 \\ \frac{v_1^2 - h^2}{2v_1 v_2} & -v_1 \leq h < v_2 - v_1 \\ 0 & \text{otherwise} \end{cases}$$

For completeness, we specify the players' cumulative bidding distributions. When h is positive,

$$F_1(b) = \begin{cases} 0 & b \leq 0 \\ \frac{h+b}{v_2} & b \in (0, v_2 - h] \\ 1 & b > v_2 - h \end{cases} \qquad F_2(b) = \begin{cases} 0 & b \leq 0 \\ 1 - \frac{v_2 - h}{v_1} & b \in (0, h] \\ 1 - \frac{v_2 - b}{v_1} & b \in (h, v_2] \\ 1 & b > v_2 \end{cases} \tag{1}$$

When h is negative,

$$F_1(b) = \begin{cases} 0 & b \leq h \\ \frac{b-h}{v_2} & b \in (h, v_2 + h] \\ 1 & b > v_2 + h \end{cases} \qquad F_2(b) = \begin{cases} 0 & b \leq 0 \\ 1 - \frac{v_2 - b}{v_1} & b \in (0, v_2] \\ 1 & b > v_2 \end{cases} \tag{2}$$

In our model, the value of the headstart is determined by the different realizations of the errors $\varepsilon_1, \varepsilon_2$. There are four possible realizations with equal probability: $h_1 = h_2 = \frac{q_1 - q_2}{\alpha}$, $h_3 = \frac{q_1 - q_2 + 2\sigma}{\alpha}$ and $h_4 = \frac{q_1 - q_2 - 2\sigma}{\alpha}$.

Recall, that we use $E(\alpha) = E(\alpha;\sigma) = \mathbb{E}[Pr(1 \text{ wins})]$ to denote the efficiency of the ranking, which matches the search engine's utility function, with $E(0) = 3/4$ in the benchmark case.

PROOF OF LEMMA 2: Since $E(\alpha) = \frac{1}{2}W_1(h_1)+\frac{1}{4}W_1(h_3)+\frac{1}{4}W_1(h_4)$ and the headstart does not depend on v_1 and v_2, it is enough to show that $W_1(\cdot)$ is increasing in v_1 and decreasing in v_2. These easily follow from the definition of $W_1(\cdot)$. The results on q_1 and q_2 follow from the fact that h_1, h_3, h_4 are all increasing in q_1 and decreasing in q_2, and $W_1(\cdot)$ depend on them only through h in which it is increasing. □

PROOF OF PROPOSITION 1: We use the notation $P_i = Pr(1 \text{ wins}|h_i)$. Given the above described equilibrium of the two-player all-pay auctions we have $P_i = W_1(h_i)$. We further define $\alpha_1 = \frac{q_1-q_2}{v_2}$, $\alpha_3 = \frac{q_1-q_2+2\sigma}{v_2}$, $\alpha_4 = \frac{q_2-q_1+2\sigma}{v_1}$, $\alpha_4' = \frac{q_2-q_1+2\sigma}{v_1-v_2}$. Note that $P_1 = P_2$, since the headstarts in the first two case are equal. Thus $E(\alpha) = \frac{1}{2}P_1 + \frac{1}{4}P_3 + \frac{1}{4}P_4$, and $P_1 = 1$ iff $\alpha \leq \alpha_1$, $P_3 = 1$ iff $\alpha \leq \alpha_3$, $P_4 = 1 - \frac{v_2}{2v_1}$ iff $\alpha \geq \alpha_4'$. Furthermore, it is easy to check that $\alpha_1 \leq \alpha_3$, $\alpha_4 \leq \alpha_3$, and $\alpha_4 \leq \alpha_4'$.

We proceed by separating the three parts of the proposition:

- Part 1: By setting $\alpha = \alpha_1$, we have $P_1 = P_3 = 1$, and thus $E(\alpha) \geq 3/4$ for any σ.

- Part 2: In order to prove this part, we determine the α value that yields the highest efficiency level for a given σ if $v_1/v_2 > 3/2$. As noted above, $E(\alpha)$ is a linear combination of $W_1(h_1)$, $W_1(h_3)$, $W_1(h_4)$. Since $W_1(\cdot)$ is continuous and h_1, h_3, h_4 are all continuous in α, it follows that $E(\alpha)$ is continuous in α. However, $E(\alpha)$ is not differentiable everywhere, but there are only a finite number of points where it is not. Therefore it suffices to examine the sign of $E'(\alpha)$ to determine whether it is increasing or not. This requires tedious analysis, since depending on the value of σ the formula describing $E(\alpha)$ is different in up to five intervals. We identify five different formulas that $E(\alpha)$

can take in different intervals and take their derivatives:

$$
\begin{aligned}
E'(\alpha) = E_1'(\alpha) &= \frac{(q_1 - q_2 - 2\sigma)^2}{4\alpha^3 v_1 v_2} \text{ if } \alpha_4 \le \alpha \le \alpha_1 \& \alpha_4', \\
E'(\alpha) = E_2'(\alpha) &= -\frac{(q_1 - q_2)^2}{2\alpha^3 v_1 v_2} \text{ if } \alpha_1 \le \alpha \le \alpha_4, \\
E'(\alpha) = E_3'(\alpha) &= -\frac{2(q_1 - q_2)^2 + (q_1 - q_2 + 2\sigma)^2}{4\alpha^3 v_1 v_2} \text{ if } \alpha_3 \& \alpha_4' \le \alpha, \\
E'(\alpha) = E_4'(\alpha) &= \frac{4\sigma^2 - (q_1 - q_2)(4\sigma + q_1 - q_2)}{4\alpha^3 v_1 v_2} \text{ if } \alpha_1 \& \alpha_4 \le \alpha \le \alpha_4', \\
E'(\alpha) = E_5'(\alpha) &= -\frac{(q_1 - q_2)(4\sigma + q_1 - q_2)}{2\alpha^3 v_1 v_2} \text{ if } \alpha_3 \& \alpha_4' \le \alpha.
\end{aligned}
$$

In any other range the derivative of $E(\alpha)$ is 0. It is clear from the above formulas that $E_1'(\alpha)$ is always positive and that $E_2'(\alpha)$, $E_3'(\alpha)$, and $E_5'(\alpha)$ are always negative. Furthermore, one can show that

$$
E_4'(\alpha) > 0 \text{ iff } \sigma > \frac{1 + \sqrt{2}}{2}(q_1 - q_2).
$$

This allows us to determine the maximal $E(\alpha)$ for different values of σ in four different cases.

1. If $\frac{q_1 - q_2}{2} \le \sigma \le \frac{v_1}{v_2} \frac{q_1 - q_2}{2}$ then $\alpha_4 \le \alpha_4' \le \alpha_1 \le \alpha_3$ and the derivative of $E(\alpha)$ takes the following values in the five intervals respectively: $0, E_1'(\alpha), 0, E_2'(\alpha), E_3'(\alpha)$. Therefore $E(\alpha)$ is first constant, then increasing, then constant again and then strictly decreasing. Thus, any value between α_4' and α_1 maximizes $E(\alpha)$. Using the notation of Corollary 1, $\hat{A}(\sigma) = [\alpha_4', \alpha_1]$.

2. If $\frac{v_1}{v_2} \frac{q_1 - q_2}{2} \le \sigma \le \frac{v_1 + v_2}{v_2} \frac{q_1 - q_2}{2}$ then $\alpha_4 \le \alpha_1 \le \alpha_4' \le \alpha_3$ and the derivative of $E(\alpha)$ takes the following values in the five intervals respectively: 0, $E_1'(\alpha)$, $E_4'(\alpha)$, $E_2'(\alpha)$, $E_3'(\alpha)$. Therefore $E(\alpha)$ is first constant, then decreasing, then strictly increasing, then depending on the sign of $E_4'(\alpha)$ increasing or decreasing, and finally strictly decreasing. Therefore if $\sigma < \frac{1+\sqrt{2}}{2}(q_1 - q_2)$ then α_1 maximizes $E(\alpha)$, that is $\hat{A}(\sigma) = \{\alpha_1\}$. If $\sigma = \frac{1+\sqrt{2}}{2}(q_1 - q_2)$ then $E(\alpha)$ is constant between α_1 and α_4', that is $\hat{A}(\sigma) = [\alpha_1, \alpha_4']$. Finally, if $\sigma = \frac{1+\sqrt{2}}{2}(q_1 - q_2)$ then $\hat{A}(\sigma) = \{\alpha_4'\}$.

3. If $\frac{v_1+v_2}{v_2}\frac{q_1-q_2}{2} \le \sigma \le \frac{v_1}{2v_2-v_1}\frac{q_1-q_2}{2}$ then $\alpha_1 \le \alpha_4 \le \alpha_4' \le \alpha_3$ and the derivative of $E(\alpha)$ takes the following values in the five intervals respectively: 0, $E_2'(\alpha)$, $E_4'(\alpha)$, $E_2'(\alpha)$, $E_3'(\alpha)$. In this case $E_4'(\alpha) > 0$ since $\sigma \ge \frac{v_1+v_2}{v_2}\frac{q_1-q_2}{2} \ge (1+\frac{3}{2})\frac{q_1-q_2}{2} > (1+\sqrt{2})\frac{q_1-q_2}{2}$. Therefore $E(\alpha)$ is first constant, then decreasing, then strictly increasing again and finally strictly decreasing. Thus, there are two candidates for the argmax: α_1 and α_4'. One can show that $E_4(\alpha_4') > E_2(\alpha_1)$ iff $v_1 > \sqrt{2}v_2$, therefore α_4' maximizes $E(\alpha)$ in this case.

4. If $\frac{v_1}{2v_2-v_1}\frac{q_1-q_2}{2} \le \sigma$ then $\alpha_1 \le \alpha_4 \le \alpha_3 \le \alpha_4'$ and the derivative of $E(\alpha)$ takes the following values in the five intervals respectively: 0, $E_2'(\alpha)$, $E_4'(\alpha)$, $E_5'(\alpha)$, $E_3'(\alpha)$. Similarly to the previous case $E_4'(\alpha) > 0$, therefore $E(\alpha)$ is first constant, then decreasing, then strictly increasing again and finally strictly decreasing. Comparing the two candidates for the argmax yields that $E_4(\alpha_3) > E_2(\alpha_1)$ iff $v_1 > (3/2)v_2$, that is α_3 maximizes $E(\alpha)$ in this case.

In each of the cases above, it is clear that the maximum is higher than $E(0) = 3/4$. In cases 1 and 2, $E(\alpha)$ is strictly increasing after a constant value of 3/4 and in cases 3 and 4 we directly compared to $E_2(\alpha_1) = 3/4$. This completes the proof of Part 2.

- Part 3: One can derive the efficiency function for different cases as in Part 2. It follows that if $\sigma \ge \frac{v_2+v_1}{v_2-v_1}\frac{q_1-q_2}{2}$ then $E'(\alpha)$ is first 0 then negative and finally positive. Therefore $E(\alpha)$ either has a maximum in $\alpha = 0$ or as it approaches infinity. However,

$$E(\alpha) \underset{\alpha\to\infty}{\longrightarrow} \frac{v_1}{2v_2} \le \frac{1}{2} < \frac{3}{4} = E(0).$$

$\square$

Proof of Corollary 1: In the proof of Proposition 1, we determined the values of α that maximize $E(\alpha)$ for different σ's. In summary:

$$\hat{A}(\sigma) = \begin{cases} [\alpha_4', \alpha_1] & \text{if } \frac{q_1-q_2}{2} \le \sigma \le \frac{v_1}{v_2}\frac{q_1-q_2}{2} \\ \alpha_1 & \text{if } \frac{v_1}{v_2}\frac{q_1-q_2}{2} < \sigma \le (1+\sqrt{2})\frac{q_1-q_2}{2} \\ [\alpha_1, \alpha_4'] & \text{if } \sigma = (1+\sqrt{2})\frac{q_1-q_2}{2} \\ \alpha_4' & \text{if } (1+\sqrt{2})\frac{q_1-q_2}{2} < \sigma \le \frac{v_1+v_2}{v_2}\frac{q_1-q_2}{2} \\ \alpha_4' & \text{if } \frac{v_1+v_2}{v_2}\frac{q_1-q_2}{2} \le \sigma \le \frac{v_1}{2v_2-v_1}\frac{q_1-q_2}{2} \\ \alpha_3 & \text{if } \frac{v_1}{2v_2-v_1}\frac{q_1-q_2}{2} \le \sigma \end{cases}$$

It is straightforward to check that all of α_1, α_3, and α'_4 are increasing in σ and that the $\hat{A}(\sigma)$ is increasing over the entire range. $\square$

Proof of Corollary 2: First, we describe the payoffs of the two players in an all-pay auction with headstarts. When players follow the mixed strategies described in (1) and (2), player 1's payoff is:

$$\pi_1(h) = \begin{cases} 0 & h \leq v_2 - v_1 \\ v_1 - v_2 + h & v_2 - v_1 < h < v_2 \\ 1 & h \geq v_2 \end{cases}$$

where h is the headstart of player 1. The payoff of player 2 can be obtained from the same formula by changing the roles. Then, we get player i's total payoff by linearly combining the above quantities:

$$\pi_1 = \frac{1}{2}\pi_1(h_1) + \frac{1}{4}\pi_1(h_3) + \frac{1}{4}\pi_1(h_4)$$

Then following the same steps as in the proof of Proposition 1, we can determine the values of α that maximize Player 1's payoff for different σ's. We get the following results. If $v_1 \leq 3v_2$ then

$$\arg\max_{\alpha} \pi_1 = \begin{cases} \alpha_1 & \text{if } \sigma \leq \frac{v_1}{v_2}\frac{q_1-q_2}{2} \\ [0, \alpha_1] & \text{if } \frac{v_1}{v_2}\frac{q_1-q_2}{2} < \sigma \end{cases}$$

In case of $3v_2 < v_1 \leq 4v_2$

$$\arg\max_{\alpha} \pi_1 = \begin{cases} \alpha_1 & \text{if } \sigma \leq \frac{v_1}{v_2}\frac{q_1-q_2}{2} \\ \alpha_3 & \text{if } \frac{v_1}{v_2}\frac{q_1-q_2}{2} < \sigma \leq \frac{v_1}{4v_2-v_1}\frac{q_1-q_2}{2} \\ [0, \alpha_1] & \text{if } \frac{v_1}{4v_2-v_1}\frac{q_1-q_2}{2} < \sigma \end{cases}$$

Finally, when $v_1 \leq 4v_2$ we have

$$\arg\max_{\alpha} \pi_1 = \begin{cases} \alpha_1 & \text{if } \sigma \leq \frac{v_1}{v_2}\frac{q_1-q_2}{2} \\ s_3 & \text{if } \frac{v_1}{v_2}\frac{q_1-q_2}{2} < \sigma \end{cases}$$

It is easy to see that with exception of the two cases when the optimal α is anywhere between 0 and α_1, Player 1 is strictly better off with a particular positive level of SEO than without it. $\square$

Proof of Proposition 2: Recall that we can determine Player 1's payoff as

$$\pi_1 = \frac{1}{2}\pi_1(h_1) + \frac{1}{4}\pi_1(h_3) + \frac{1}{4}\pi_1(h_4),$$

whereas Player 2's payoff is given by reversing the roles. Examining the above formula shows that each players payoff only depends on $q_1 - q_2$ and not the individual qualities and is a continuous function which is linear with different slopes in different intervals. The function takes 3 different forms in the cases of $v\alpha/2 \geq \sigma$, $v\alpha/4 \geq \sigma < v\alpha/2$, $\sigma < v\alpha/4$. In all cases the slope is less than or equal to $\frac{1}{4}\frac{1}{s}$ for $q_1 - q_2 < 0$. In the first two cases the slope is decreasing for $q_1 - q_2 > 0$ with a slope of $\frac{1}{2}\frac{1}{s}$ and $\frac{3}{4}\frac{1}{s}$ above zero, respectively in the two cases. In the third case the slope above zero is first $\frac{3}{4}\frac{1}{s}$ then $\frac{1}{s}$. First we show that only $(0, 0)$ can be a pure strategy equilibrium in the first stage. Since the slope below zero is quite shallow, if the marginal cost of content $c > \frac{1}{4}\frac{1}{s}$ then sites have an incentive to decrease the investment. If, on the other hand $c \leq \frac{1}{4}\frac{1}{s}$ then sites have an incentive to invest more since the slope above zero is always higher than this. To show when $(0, 0)$ is an equilibrium, we fix $q_2 = 0$ and look at whether Player 1 has an incentive invest more than 0. In the first two cases the critical value of c will just be the slope of Player 1's payoff function above zero: $\frac{1}{2}\frac{1}{s}$ and $\frac{3}{4}\frac{1}{s}$, respectively. Player 1 does not have an incentive to deviate from zero iff c is higher than the respective slope. In the third case determining the critical value requires more detailed analysis since the slope increases. The analysis reveals that the best player 1 can do is achieve a payoff of $v - 2\sigma/s - \sigma/(2s)$ with an investment of $vs - 2\sigma$ yielding the critical value for c and completing the proof. $\square$

Proof of Proposition 3: Following the same steps as in the proof of Proposition 2 we can determine the payoffs as a function of each player's quality investment. The same logic shows that the only possible pure strategy equilibrium is the one in which players do not invest in content. One can then separate three cases in which the payoff functions take different forms and show that the slope is always at most $\frac{1}{\alpha} + m$. $\square$

Proof of Proposition 4: First, we derive the efficiency of the ranking process given the error term. Let $f_\Delta(.)$ be the density function of the distribution of $\varepsilon_1 - \varepsilon_2$. If α is high enough then the headstarts diminish and the sum of the valuation and the headstart for sites 3 and below will be lower than for 1 and 2. In an all-pay auction this leads to only the first two site's participation. Then using the notation of the proof of Proposition 1 we have

$$Pr(1 \text{ wins}) = E(\alpha) = \mathbb{E}W_1\left(\frac{q_1 - q_2 + \sigma\varepsilon_1 - \sigma\varepsilon_2}{\alpha}\right) = \int_{-\infty}^{+\infty} W_1(h) f_\Delta\left(\left(\frac{h - q_1 + q_2}{\sigma}\right)\alpha\right) dh.$$

Note that for positive h the value of $W_1(h)$ is strictly higher than $1-v_1/2v_2$ and for $h > v_2 - v_1$ it is exactly $1 - v_2/2v_1$. Furthermore if α is high enough then for an arbitrary small δ we have

$$\int_{-\infty}^{v_2-v_1} f_\Delta\left(\left(\frac{h - q_1 + q_2}{\sigma}\right)\alpha\right) dh < \delta.$$

Let $\hat{\alpha}$ denote such a high α and let δ be smaller than $1 - E(0) > 0$. Then $E(\hat{\alpha}) > (1 - v_2/2v_1)(1 - \delta)$. Finally, let $\hat{v} = \frac{v_2(1-\delta)}{2(1-\delta-E(0))}$. Then $v_1 > \hat{v}$ yields $E(\hat{\alpha}) > E(0)$.

□

Proof of Proposition 5: Our proof uses an induction on k, the number of links searched by the consumer. When $k = 1$, proposition 4 gives the necessary conditions and result. Given the independence assumption on choice of links to display in each auction the expected utility of the consumer from searching k links is

$$E_k = E_{k-1} + (1 - E_{k-1})c_k\bar{q}_k > E_{k-1}.$$

By the induction hypothesis, there exists an $\hat{\alpha}(\sigma)$ and a $\Delta_v(\sigma)$ such that if $v_i > v_{i+1} + \Delta_v(\sigma)$ for $1 \le i \le k-1$ then $\Delta E_{k-1} = E_{k-1}(Q(\hat{\alpha})) - E_{k-1}(Q(0)) > 0$.

Suppose $v_k > v_{k+1} + \Delta_v(\sigma)$. Since $E_k > E_{k-1}$ for every k and every α, it must be that $\Delta E_k \ge \Delta E_{k-1} > 0$ which concludes the proof.

□

References

Aggarwal, Gagan, Jon Feldman, S. Muthukrishnan, Martin Pál. 2008. Sponsored search auctions with markovian users. *WINE '08: Proceedings of the 4th International Workshop on Internet and Network Economics*. Springer-Verlag, Berlin, Heidelberg, 621–628.

Animesh, A., V. Ramachandran, S. Viswanathan. 2007. An empirical investigation of the performance of online sponsored search markets. *Proceedings of the ninth international conference on Electronic commerce*. ACM, 160.

Athey, Susan, Glenn Ellison. 2009. Position auctions with consumer search. NBER Working Papers 15253, National Bureau of Economic Research, Inc. URL `http://ideas.repec.org/p/nbr/nberwo/15253.html`.

Barut, Yasar, Dan Kovenock. 1998. The symmetric multiple prize all-pay auction with complete information. *European Journal of Political Economy* **14** 627–644.

Baye, Michael R., Dan Kovenock, Casper G. de Vries. 1996. The all-pay auction with complete information. *Economic Theory* **8** 291–305.

Benczur, Andras A., DAvid Siklosi, Jacint Szabo, Istvan Biro, Zsolt Fekete, Miklos Kurucz, Attila Pereszlenyi, Simon Racz, Adrienn Szabo. 2008. Web spam: a survey with vision for the archivist. IWAW '08 International Web Archiving Workshop. Available at: http://iwaw.europarchive.org/08/IWAW2008-Benczur.pdf.

Bracha, O., F. Pasquale. 2007. Federal Search Commission-Access, Fairness, and Accountability in the Law of Search. *Cornell L. Rev.* **93** 1149.

Chen, J., D. Liu, A.B. Whinston. 2009. Auctioning keywords in online search. *Journal of Marketing* **73**(4) 125–141.

Chen, Yongmin, Chuan He. 2006. Paid Placement: Advertising and Search on the Internet. *SSRN eLibrary* .

Clark, Derek J, Christian Riis. 1998. Competition over more than one prize. *American Economic Review* **88**(1) 276–89. URL `http://ideas.repec.org/a/aea/aecrev/v88y1998i1p276-89.html`.

Clark, D.J., C. Riis. 2000. Allocation efficiency in a competitive bribery game. *Journal of Economic Behavior and Organization* **42**(1) 109–124.

Edelman, Benjamin, Michael Ostrovsky, Michael Schwarz. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review* **97**(1) 242–259. URL `http://ideas.repec.org/a/aea/aecrev/v97y2007i1p242-259.html`.

Feng, J., H.K. Bhargava, D.M. Pennock. 2007a. Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *INFORMS Journal on Computing* **19**(1) 137.

Feng, J., Z.J.M. Shen, R.L. Zhan. 2007b. Ranked items auctions and online advertisement. *Production and Operations Management* **16**(4) 510–522.

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

Hillman, Arye L., John G. Riley. 1987. Politically contestable rents and transfers. UCLA Economics Working Papers 452, UCLA Department of Economics. URL `http://ideas.repec.org/p/cla/uclawp/452.html`.

Jeziorski, Przemyslaw, Ilya Segal. 2009. What Makes Them Click: Empirical Analysis of Consumer Demand for Search Advertising. *SSRN eLibrary* .

Katona, Zsolt, Miklos Sarvary. 2010. The race for sponsored links: Bidding patterns for search advertising. *Marketing Science* forthcoming.

Kirkegaard, René. 2009. Favoritism in asymmetric contests: Head starts and handicaps. *working paper* .

Konrad, Kai A. 2007. Strategy in contests – an introduction. *WZB-Markets and Politics Working Paper No. SP II 2007-01* .

Krishna, Vijay John Morgan. 2007. The winner-take-all principle in small tournaments. *Advances in Applied Microeconomics* **28** 849–862.

Mercadante, J.M. 2008. What is Google's Reputation Score-A Method for Modified Self-Regulation of Search. *Seton Hall Legis. J.* **33** 327.

Pasquale, F. 2006. Rankings, Reductionism, and Responsibility. *Clev. St. L. Rev.* **54** 115.

Rutz, Oliver J., Randolph E. Bucklin. 2007. A Model of Individual Keyword Performance in Paid Search Advertising. *SSRN eLibrary* .

Sen, Ravi. 2005. Optimal search engine marketing strategy. *Int. J. Electron. Commerce* **10**(1) 9–25.

Siegel, Ron. 2009. All-pay contests. *Econometrica* **77**(1) 71–92.

Sisak, Dana. 2009. Multiple-prize contests the optimal allocation of prizes. *Journal of Economic Surveys* **23** 82–114(33). URL `http://www.ingentaconnect.com/content/bpl/joes/2009/00000023/00000001/art00003`.

Varian, Hal R. 2007. Position auctions. *International Journal of Industrial Organization* **25**(6) 1163–1178. URL `http://ideas.repec.org/a/eee/indorg/v25y2007i6p1163-1178.html`.

White, Alexander. 2009. Search engines: Left side quality versus right side pro ts. *working paper, Touluse School of Economics* .

Xing, Bo, Zhangxi Lin. 2006. The impact of search engine optimization on online advertising market. *ICEC '06: Proceedings of the 8th international conference on Electronic commerce*. ACM, New York, NY, USA, 519–529. doi:http://doi.acm.org/10.1145/1151454.1151531.

Xu, Lizhen, Jianqing Chen, Andrew B. Whinston. 2009. Too organic for organic listing? interplay between organic and sponsored listing in search advertising. *working paper, Available at SSRN: http://ssrn.com/abstract=1409450* .