The Race for Sponsored Links: Bidding Patterns for Search Advertising

Zsolt Katona and Miklos Sarvary*

July 9, 2008

*Zsolt Katona is a Ph.D. candidate and Miklos Sarvary is Professor of Marketing at INSEAD, Bd. de Constance, 77305, Fontainebleau, France. E-mail: zsolt.katona@insead.edu, miklos.sarvary@insead.edu. Phone: +33160729226, fax: +33160745500.

The Race for Sponsored Links: Bidding Patterns for Search Advertising

Abstract

Paid placements on search engines reached sales of over $10 billion in the U.S. last year and represent the most rapidly growing form of online advertising today. In its classic form, a search engine sets up an auction for each search word in which competing web sites bid for their sponsored links to be displayed next to the search results. We model this advertising market focusing on two of its key characteristics: (i) the interaction between the list of search results and the list of sponsored links on the search page and, (ii) the inherent differences in click-through rates between sites. We find that both of these special aspects of search advertising have a significant effect on sites' bidding behavior and the equilibrium prices of sponsored links. In three extensions, we also explore (i) heterogenous valuations across bidding sites, (ii) the endogenous choice of the number of sponsored links that the search engine sells, and (iii) a dynamic model where web sites' bidding behavior is a function of their previous positions on the sponsored list. Our results shed light on the seemingly random order of sites on search engines' list of sponsored links and their variation over time. They also provide normative insights for both buyers and sellers of search advertising.

Keywords: Internet Marketing, Position Auctions, Game Theory.

1 Introduction

Search advertising is becoming one of the dominant forms of online advertising. Potential advertisers bid for a place on the list of sponsored links that appears on a search engine's "results" page for a specific search word. In 2007, the revenues from such paid placements have doubled compared to 2006, reaching over $10 billion in the United States[1]. This fast growing market is increasingly dominated by Google, which today, controls some 70% of Internet searches.[2] How such advertising is priced and what purchase behavior will advertisers follow for this new form of advertising is the subject of the present paper.

Previous research studying search advertising has focused on the problem of multi-item (or position) auctions and examined the optimal bidding behavior of advertisers (Varian 2007, Edelman et al. 2007). However, a key characteristic of paid placements is that the consumer is facing two "competing" lists of sites that are both relevant in the context of the particular search: (i) the results list of the search (*organic links*) and (ii) the list of *sponsored links*. Furthermore, membership and position on the results list is generally exogenous and typically represents the site's popularity or inherent value. The search engine cannot use this list strategically without losing credibility from users. Thus, the existence of this search list cannot be ignored when one evaluates sites' bidding behavior for sponsored links appearing on the same page.

Another key characteristic of the problem is that the search engine can take into account advertisers' inherent traffic when awarding paid links. As the bids correspond to payments

[1]See "Inside Microsoft's War Against Google," *Business Week*, May 19, 2008, p.36, which reports total revenues from search advertising forecasted to reach $17.6 billion in the U.S. by 2012. Worldwide revenues from paid placements are expected to reach $45 billion by 2011 (see "Where is Microsoft Search?", *Business Week*, April 2, 2007, p. 30.)

[2]Ibid. Furthermore, other major search engines use similar methods to target searching consumers. AOL uses Google's search while Yahoo!'s search page is almost identical to Google's. Other popular sites like Amazon and eBay (also powered by Yahoo!) sell only a few sponsored links on their search pages and many times these are linked to their own content.

per-click, this information is important in determining the search engine's total revenue from a given sponsored link. Therefore, search engines take sites' click-through rates into account in addition to their per-click bids when awarding paid placements. Furthermore, the search engine can also determine how many sponsored links it offers for a particular search word. Again, advertisers' incentives for bidding and, in turn, the search engine's revenue will depend on this decision.

Finally, a third important characteristic of paid placements is that bidding for sponsored links happens frequently over time. This has two implications. First, it means that repeated bidding by the same players reveals their valuations for the different advertising links. Second, if the advertising effect of sponsored links has a lagged effect - as is often the case with advertising - then bidding strategies should be dynamic rather than optimized for a single time period.

We develop a model that takes into account these key aspects of search advertising. Specifically, in our base model, we explicitly describe consumers' clicking behavior on the search page as a function of sites' presence and order on the organic links list and/or among the sponsored links. Then, assuming (as is the case in practice) that click-through rates are known by the bidders, we derive sites' optimal bidding strategies for sponsored links and the search engine's optimal behavior, taking into account consumers' clicking patterns.

Our results shed light on the advertising patterns observed on different search pages. Specifically, search pages can be characterized by a variety of patterns in terms of the identity and position of sponsored links. In particular, there does not seem to be a clear relationship between the results list of a search and the list of sponsored links. Sometimes a site may appear in both or in only one (either one) of the lists. For example, at the time of writing this paper, on Google's search results page, for the word "travel", the two lists were entirely different. However, on the results page for the search word “airlines”, United Airlines appeared as the first search result and second on the sponsored links list[3]. One can also observe significant fluctuations in the sites’ order in the sponsored links list. Finally, the number of items listed in the sponsored list is also changing over time. Our model proposes a number of testable hypotheses that account for the variations described above. Furthermore, it also generates normative guidelines to both advertisers and the search engine on how to buy and sell sponsored links. For instance, our analysis suggests that a search engine can experimentally measure consumers’ clicking behavior on its search page and improve the weights to be used to correct sites’ bids for sponsored links.

In a second step, we provide three extensions to the base model. First, we explore the case when clicks are valued heterogeneously across sites. We find that the basic competitive dynamics do not change although the actual outcomes are influenced by sites’ specific valuations. In a second extension, we allow the search engine to choose the number of sponsored links to auction away. We show under what conditions it is worth for the search engine to increase or decrease the number of links. Finally, we also explore a dynamic model where sites bid repeatedly and consumer clicks have a lagged effect (e.g. due to a loyalty factor). Here, we again find conditions under which sites either alternate in winning the auction or their order remains relatively stable. In particular, we show that an alternating equilibrium is better for all the players.

The rest of the paper is organized as follows. The next section summarizes the relevant literature. This is followed by the basic model description in Section 3 and equilibrium analysis in Section 4. Section 5 explores the three extensions outlined above. We end with a summary of the main findings and model limitations in Section 6. All proofs and technical

[3]In a casual experiment, we have tested 50 randomly selected search words and found 6 occasions (12%) when there was an overlap between the search results and the list of sponsored links.

details appear in the Appendix.

2 Relevant Literature

Since search advertising is mostly responsible for the growth of the online advertising business, it has attracted significant interest in the economics literature[4]. Edelman et al. (2007) analyze the generalized second price auction that is used by most search engines to allocate sponsored links on search pages[5]. The paper focuses on equilibrium properties and compares these to other auction mechanisms. Varian (2007) studies a similar problem but assumes away uncertainty and shows that the equilibrium behavior matches empirical pricing patterns for sponsored links. More recent papers (Feng (2007), Feng et al. (2006) and Athey and Ellison (2008)) further elaborate on optimal auction design by considering reserve prices.

A separate set of papers explore the important issue of fraudulent behavior in the context of search advertising. Wilbur and Zhu (2008) study click fraud and its non-trivial effect on the distribution of surplus between advertisers and the search engine. In a related study, Bhargava et al. (2005) explore shill bidding in a consumer auction context where bidders can establish multiple identities.

While the previous streams add considerably to our understanding of how to efficiently allocate search advertising, it neglects the behavior of searching consumers. Chen and He (2006) also study competitive bidding for paid placements but assume differentiated advertisers and explicitly consider consumers who are initially uncertain about their valuations for products. They show how the auction mechanism improves the efficiency of consumer

[4]The other dominant advertising model - sites buying ads on each other's pages - is analyzed in Katona and Sarvary (2008). That paper studies equilibrium advertising prices and the endogenous network structure determined by the advertising links.

[5]This literature builds on an established stream of research on mechanism design represented by classic papers, such as Myerson (1981) and Maskin and Riley (2000).

search and results in possible price dispersions for advertising. Athey and Ellison (2008) extend this approach and further explore the implications of the results for optimal auction design.

Our work is different from these literature streams. We assume away fraud and are less focused on optimal auction design but are interested in capturing relevant behaviors from searching consumers. In particular, our focus is on the *interaction* between the search engine's basic service of finding relevant sites in a given search context and its private objective to sell sponsored links on search pages. We model the inherent competition between the output of these two processes and evaluate its effect on advertisers' behavior. In terms of modeling the allocation of sponsored links, our paper is closest to Varian (2007) but our focus is elsewhere. Rather than characterizing the optimal auction 'rules' for allocating multiple items, we are interested in the optimal weights that search engines should use to 'correct' advertisers' bids taking into account consumers' behavior on the search page.

Beyond the explicit modeling of consumers' clicking behavior, our modeling approach is different in many other ways. We assume a concave response function to advertising that is well documented in marketing. As opposed to the existing literature, we also explore the endogenous choice of the number of sponsored links offered, which can be an important decision variable for the search engine. Furthermore, we study a dynamic model in which advertisers repeatedly bid for sponsored links and consumer visits have a lagged effect. This dynamic advertising model is related to previous work on the dynamic setting of marketing variables in a competitive context using a Markovian game. For an application on advertising see Villas-Boas (1993), while an application for dynamic R&D competition can be found in Ofek and Sarvary (2003). Our work uses a similar framework and relates to the results of both papers. The possibility of an alternating advertising pattern is similar to Villas-Boas (1993) and is largely driven by decreasing returns on advertising. However, in our model, as in Ofek and Sarvary (2003), we have a contest as advertisers' bid for each position on the list with only one winner. Our dynamic model is also somewhat related to the dynamic auction model of Zeithammer (2006). However, in our case this is a repeated auction for a per-period prize while his paper considers dynamic bidding for a single item.

Finally, recent empirical work on search advertising (Rutz and Bucklin 2007a,b) studies the effectiveness of paid placements with particular attention devoted to spillover and lagged effects as well as contexts when multiple search words are used. In another paper, Goldfarb and Tucker (2007) show that the auction mechanism allows search engines to discriminate between bidding firms with different inherent valuations for advertising. Our model extensions are largely motivated by these papers (see our dynamic model and our examination of heterogenous firm valuations as well as the discussion on multiple search words at the end of the paper) although the present paper admittedly has a more normative focus.

3 The Model

We assume n web sites that are indexed with respect to their exogenously given, inherent click-through rates (CTRs), $\gamma_1 > \gamma_2 > ... > \gamma_n$. These rates represent the value of the sites in the eyes of the consumers or can be thought of as their popularity in the context of a search word. The $(n+1)$th player is a search engine (SE), a special website[6]. The SE ranks the sites according to their popularity in a given search context, that is, the CTRs determine the ranking on the search list. This is consistent with the idea that the SE's basic service lies in finding sites, that consumers are most interested in. Notice that CTRs are typically known by sites and the SE because of regular bidding and also because such statistics are available on the Web. However, consumers typically do *not* know CTRs. In our model, the

[6]We assume that the SE is a monopolist. While this is not entirely true in practice, Google dominates the search industry with over 70% of all searches, a proportion that is growing (Ibid).

search engine returns the r highest ranked sites as the search results (Sites $1, 2, ..., r$). Next to these organic links, the SE also displays s number of sponsored links. The order of these links can be chosen by the SE and this choice is based on the bids submitted by the web sites. Let $l_1, l_2, ..., l_s$ denote the sites winning the sponsored links, in order of appearance. Thus, the output of the SE is modeled as a page with two lists: a search list and a sponsored ads list. Google's search page is exactly like this and other search engines have a similar format (see Figure 1).

3.1 Consumers' behavior on the search page

We assume, that the SE attracts a unit traffic of consumers, which is distributed in the following way. When a consumer arrives to the SE's page generated by the search, s/he either clicks on one of the regular results, one of the sponsored links or leaves the page without clicking. We assume that consumers' clicking behavior is affected by the following four factors.

1. The order in which sites are listed on the lists.
2. Differences in click probabilities between the sponsored list and the search result list.
3. Individual differences between sites in inherent CTRs or popularity.
4. Whether the site appears in both the organic and the sponsored lists or only one of the lists.

For the first factor, assume that $\alpha_1, \alpha_2, ... > 0$ denote the psychological order constants that determine how the possible clicks are distributed through an ordered list of items. That is, whenever someone sees an ordered list of *equally interesting* items s/he chooses the ith item with probability proportional to α_i. Generally, we can say that $\alpha_1 > \alpha_2 > ...$, but there might be exceptions. For example, the last item in a list may be more appealing than one in the middle. For the second factor, let $\beta > 0$ denote how many times more/less attractive a sponsored link is than an organic link. That is, how many times more/less consumers click on a sponsored link over an equally interesting link in the same position on the organic search list. Since consumers are likely to exhibit some level of aversion to advertising (see e.g. Lutz (1985) for a classic reference and Edwards et al. (2002) and Schlosser et al. (1999) for empirical evidence in the Internet context), we expect β to be less than 1, although we do not need to assume this. Combining the two factors, the distribution of consumers among the links, not taking into account individual differences between sites' popularity, is determined by the parameters: $\alpha_1, \alpha_2, ..., \alpha_r$ and $\beta\alpha_1, \beta\alpha_2, ..., \beta\alpha_s$. Specifically, $M = \sum_{i=1}^{r} \alpha_i + \beta \sum_{i=1}^{s} \alpha_i$ represents the maximum potential traffic that can flow through all the links on the search page. Since the SE has a unit traffic for each search word, we normalize M to 1. The real traffic that flows through the links is less than M however, because it also depends on the sites' popularity or CTR. This is taken into account in the third factor that we explore next.

For the third factor, that takes individual differences into account, we can multiply the α and β parameters with the inherent CTRs of the sites (γ_i). In any particular position, a site with a higher CTR is more likely to attract a click than another site in the same position having a lower CTR. For example, Site 1 will receive $\alpha_1\gamma_1$ clicks on the first organic link, whereas site j in the second position on the sponsored list will receive $\beta\alpha_2\gamma_j$ clicks on its sponsored link. That is, γ_i determines what proportion of the maximum clicks that are possible in a certain position a site will receive based on its popularity.

Finally, for the fourth factor, we assume that if a site is listed both among the regular search results and the sponsored links, the latter will have a lower click-probability than if the site were listed *only* on the sponsored links list. Specifically, let δ denote the strength of this effect, that is, the proportion of people who do not click on a sponsored link if it is

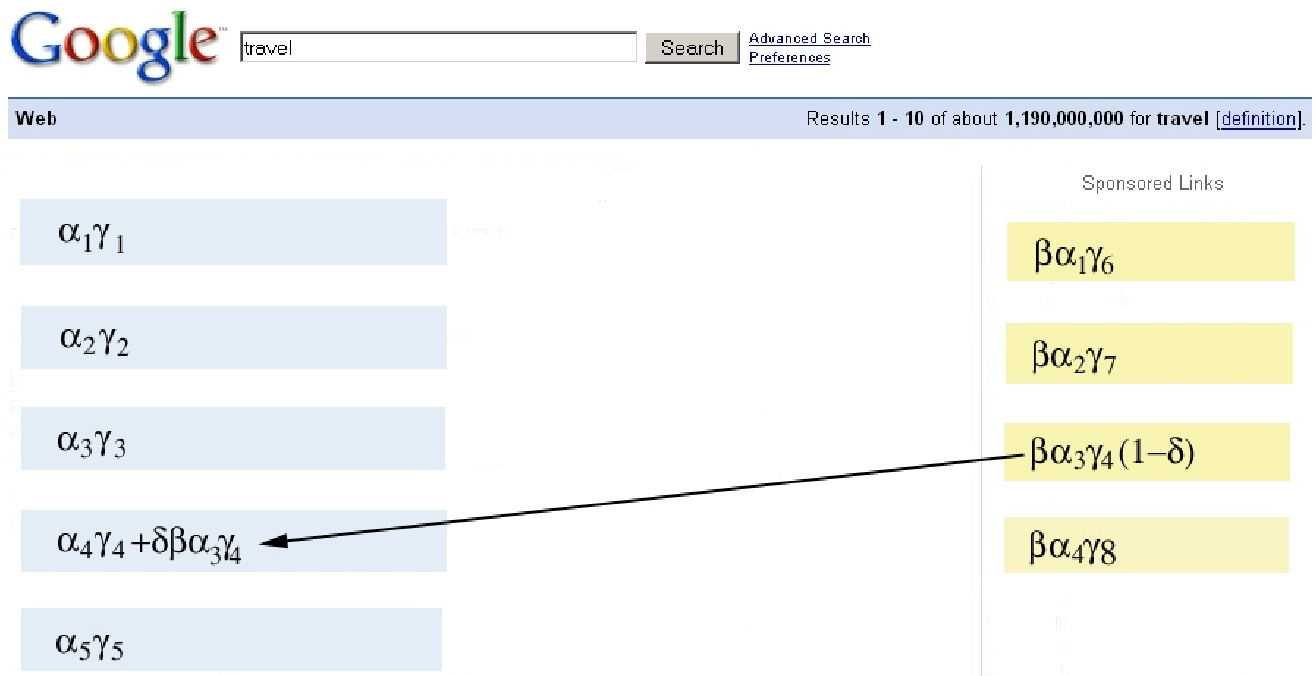


Figure 1: The distribution of clicks on a search page. The left list contains the organic search results, where sites appear in the order of CTRs, γ_i. The right list represents paid placements where the order depends on sites' bids. In this example, the third sponsored link also appears on the organic search list in position 4, therefore, $\delta\beta\alpha_3\gamma_4$ clicks are redirected to the organic link.

also displayed among the regular results but click on the regular link instead[7]. Note that the parameter δ does not have an effect on the total traffic that a site gets from the search engine because it simply changes the origin of this traffic. However, it affects the number of clicks on sponsored links, which will be important in the sites' bidding process and will also affect the SE's revenue. Figure 1 summarizes the four factors of our model describing the clicking behavior of consumers.

[7]Similarly to the case of β, we speculate that, due to aversion to advertising, consumers prefer organic results, that is, $\delta > 0$. However, our results also hold for negative values of δ.

Given these factors, we now determine how the traffic of the search engine is distributed through the web sites. Let $A(i)$ denote the function that takes a value of 0 if Site i does not win a sponsored link, that is, $i \notin \{l_1, l_2, ..., l_s\}$ and α_j if Site i wins the jth sponsored link. With these, the total traffic that Site i gets from the search engine if it appears on the search list (i.e. if $i \leq r$) is:

$$t_i = t_i^R + t_i^S = \gamma_i \alpha_i + \gamma_i \beta A(i).$$

If i does not make it on the organic search list, i.e. if $i > r$ then the traffic is:

$$t_i = t_i^S = \gamma_i \beta A(i).$$

Note that these quantities largely depend on γ_i, the site's inherent CTR or popularity, which determines how many clicks a site's link receives given its position.

3.2 Web sites

Web sites make profits from the traffic that arrives to their sites from the search engine[8]. Let us assume, that there is a common $R(t)$ function for all sites that determines the revenue associated with t amount of traffic for a given search word. As such, we assume that for each word there exist a common function determining how clicks can be converted into revenues. In Section 5.1, we relax this assumption and allow for individual differences in sites' valuations. Here, we naturally assume that $R(t)$ is increasing and concave[9]. In order to obtain sponsored links, sites have to submit bids to the search engine. The bid that Site i submits, b_i is the maximum amount that it is willing to pay for unit traffic (per-click). If the search engine decides to include Site i among the sponsored links, Site i has to pay an

[8]Thus, we ignore the fact, that sites could already have different amounts of incoming traffic from other sources. If we naturally assume that sites with a higher CTR also have higher outside traffic, then the results still hold.

[9]See Rutz and Bucklin (2007b) for a detailed analysis on how $R(t)$ could be estimated in practice.

advertising fee of $p_i t_i^S$, where $p_i \leq b_i$ is set by the search engine. Therefore, Site i's utility is

$$u_i = R(t_i^R + t_i^S) - p_i t_i^S$$

if it wins a sponsored link and $u_i = R(t_i^R)$ otherwise, where t_i^S depends on which sites win the sponsored links.

Thus, in our model, the SE uses an auction to allocate the sponsored links. This is consistent with what search engines do in reality. However, an auction may not be necessary for such allocation since all players have common knowledge about all valuations, i.e. the game is one of complete information. As we mentioned before, we assume this because CTRs are common knowledge and repeated bidding gives ample time and data for all players to discover the valuations of other parties. A similar argument is advanced in Edelman et al. (2007) and Varian (2007). As such, the auction mechanism is used as an efficient pricing mechanism. While in theory, the SE could calculate and offer the optimal price for each sponsored link, allowing sites to self-select for each position, pricing links through an auction is easier and more robust to variations of participants over time.[10] There are also costs associated with setting prices individually, which might be overwhelming for the enormous number of possible keywords (see Zeithammer and Liu (2008) for a study of the tradeoffs between using fixed prices or auctions).

At this point, the SE is completely free to determine the order of winners and advertising fee it charges for a click, $p_i \leq b_i$. First, we will show that in a one-period game the SE sets $p_i = b_i$ corresponding to a first price auction, then we will discuss the different types of auctions that search engines use in practice. Based on this discussion, in Section 3.3, we will restrict the SE's strategies and define the types of equilibria we use in the subsequent

[10] The auction may also resolve problems related to some level of information asymmetry about valuations. Our perspective follows Varian (2007) in that asymmetric information is not the key issue in the pricing of sponsored links. We would like to thank the Area Editor for drawing our attention to this issue.

analysis.

The timing of the game is the following. First, web sites simultaneously submit their b_i bids, knowing all the click-through rates and $R(t)$. Then, the search engine decides which sites it will include among the sponsored links and in what order. Finally, sites pay the advertising fee to the search engine and realize profits from the traffic they receive.

3.3 The Search Engine

First, we determine the SE's best response to given bids $b_1, b_2, ..., b_n$ in the second stage of the game. Although it would seem so, the best strategy is not to simply assign the sponsored links to web sites in the order of their bids. The SE has to consider the sites' CTRs, since the total traffic it sells to them and thus its revenue depends on these rates. Therefore, a site with a high CTR may pay a higher total fee even if its bid is low. An opposite effect is that sites with the highest inherent CTRs will also appear on the regular search list. As a result, they will attain fewer clicks on the sponsored link because a δ proportion of the consumers will click on the regular search results link instead[11]. Formally, the SE maximizes its profit,

$$\Pi_{SE} = \sum_{i=1}^{s} t_i^S p_i.$$

The following claim summarizes the SE's best response to the bids. Let $I(i)$ denote the function that takes the value 1 if $i \leq r$ and 0 otherwise. The SE's decision can be described by the series $l_1, l_2, ..., l_n$, where Site l_i will get sponsored link i. Sites $l_{s+1}, l_{s+2}, ..., l_n$ will not get a sponsored link.

Claim 1 *In equilibrium,*

$$\gamma_{l_i} b_{l_i}(1 - \delta I(l_i)) \geq \gamma_{l_j} b_{l_j}(1 - \delta I(l_j))$$

[11]In the exceptional case of $\delta < 0$, it is the sponsored link that receives more clicks.

holds for $i < j$, where $i \leq s$ and the SE sets $p_i = b_i$.

In other words, the search engine ranks the sites according to their $\gamma_i b_i(1 - \delta I(i))$ and charges each site's bid. That is, for sites that are not in the top r among the search results, their position among the sponsored links is determined by their inherent CTR multiplied by their bid. For top sites, this value is multiplied by $(1 - \delta)$, accounting for consumers who choose to click on the results link instead of the sponsored link.

As a result of Claim 1, in a non-repeated game, the search engine's best strategy is to charge the highest bid (corrected with the CTR). The reason is that, in this simple case in which sites only bid once, the search engine does not have to worry about influencing sites' subsequent bidding strategies. This corresponds to a first price auction. However, in reality search engines use second price auctions (some of them correcting for differences in CTRs, some of them not) to avoid the problem that when multiple items with different values are auctioned, then the first price auction typically does not have an equilibrium. This is because bids in a first price auction always converge towards each other, which makes it impossible to reflect the differences in valuations for the different items[12]. Thus, for our analysis, it is important to discuss the different types of auctions and equilibria that can be used in our models.

In our analysis, we assume that web sites have full information about each others' bids, valuations and CTRs. This is consistent with reality: quite well known valuations across sites are typical characteristics of auctions of sponsored links. When competitors' valuations are known, a first price auction for a single item typically has an infinity of equilibria. For example, let $v_1 > v_2 > ... > v_n$ be the valuations of n bidders for a single item. If a first price

[12]The existence of an equilibrium may not be important to the SE, although it guarantees a certain level of price stability as sellers tend to converge to it over time. An additional reason to use a second price auction is that, if valuations are uncertain, then the second price auction is a mechanism that leads to truth-telling in a single-item auction.

auction is applied then the winner pays its bid. In equilibrium, the winner is always player 1 and the winning bid, b_1 can take any value in the $(v_2, v_1]$ interval. Thus, the auctioneer's revenue is between v_2 and v_1. We denote this type of equilibrium by FNE (first price Nash equilibrium).

In the case of a second price single-item auction, anyone can win the auction in a Nash equilibrium (SNE). If every player bids zero except player i, who bids $v_0 > v_1$, then the winner is player i, who has to pay nothing. In general, the second highest bid is always below v_1, so the auctioneer's revenue is somewhere between 0 and v_1. To restrict the possible outcomes of a second price auction, Varian (2007) introduced the notion of symmetric equilibria for multi-item second price auctions, (SSNE), which is a subset of the pure-strategy Nash-equilibria. In such an equilibrium, the player in position k is better off paying the bid of the player in position $k+1$, then would be in position l paying the bid of player $l+1$. This is a stronger restriction than in an SNE for moving up in the ranking because in an SNE a player is only supposed to be better off paying bid $k+1$ for position k than paying bid l for position l. Since bid l is higher than bid $l+1$, an SSNE is always an SNE but the opposite is not true. According to Varian (2007), in an SSNE, the order of winners is always $1, 2, 3, ...$, that is, in case of a single item the winner is always player 1. Furthermore, the auctioneer's maximum SSNE revenue is the same as the maximum SNE revenue and is equal to v_1 in case of a single item. Since the equilibria in a first price single-item auction (FNE) and symmetric equilibria in a second price single-item auction (SSNE) give the same results for the bid orders and *maximum* revenues of the seller, we can use the two concepts interchangeably for our analysis, if there is only one sponsored link. For multiple links, the FNE usually does not exist, so in this case, we will always use the SSNE as the equilibrium concept. That is, we will restrict the search engine's strategy space to running second price auctions.

We always correct for CTRs as it is established in Claim 1. Player i's bid is multiplied by $\gamma_i(1-\delta I(i))$ and the search engine ranks the

$$F_i = \gamma_i b_i (1 - \delta I(i)),$$

values when determining the order of sites and the prices. In a first price auction, Site i has to pay $p_i = b_i$ for a click, corresponding to a total fee of $\beta A(i) F_i$, where $A(i)$ reflects its position. In a second price auction, if Site i is followed by Site j in the order then Site i has to pay

$$p_i = \frac{F_j}{\gamma_i(1-\delta I(i))} = b_j \frac{\gamma_j(1-\delta I(j))}{\gamma_i(1-\delta I(i))}$$

for a click, totaling to a fee of $\beta A(i) F_j$. The next section determines the equilibrium bids.

4 Equilibrium analysis

4.1 Bidding strategies for one sponsored link

To illustrate the primary forces that work in the game, we first consider the case in which there is only one sponsored link offered, that is, $s = 1$. Let

$$G(i) = R(I(i)\gamma_i\alpha_i + \gamma_i\beta\alpha_1) - R(I(i)\gamma_i\alpha_i)$$

denote the revenue gain for Site i of winning the sponsored link. Clearly, the total fee Site i will pay for the sponsored link cannot exceed $G(i)$. Let $w_1, w_2, ..., w_n$ be a permutation of sites such that $G(w_1) > G(w_2) \geq ... \geq G(w_n)$ holds[13]. Furthermore, let P_1 denote the total fee that the winner pays for the sponsored link,[14] which is equal to the seller's revenue.

[13]The assumption that there is a single highest value eases the presentation of results, but does not change them qualitatively.

[14]In case of a first price auction, this is calculated from its own bid. In case of a second price auction, it is calculated from the second highest bid, corrected for CTRs.

Proposition 1 *In any FNE and SSNE, the winner of the sponsored link is Site w_1 and the total fee it pays is $G(w_1) \geq P_1 \geq G(w_2)$.*

Given the assumption that $R()$ is increasing and concave, the winner can be any site from 1 to $r+1$, depending on the parameters. For example, if $R()$ were linear then the site with the highest $\gamma_i \beta \alpha_1$, that is, Site 1 would be the winner. However, if R() is very concave or the γ_i's are not too far from each other, that is $\gamma_1 - \gamma_{r+1} \to 0$, then the winner is Site $r+1$. These two cases illustrate the two forces that work against each other in determining the outcome. On one hand, since $R()$ is concave, sites who already receive traffic from the search engine through regular results have a lower benefit from winning the link[15]. On the other hand, sites with a higher γ_i obtain more traffic from a sponsored link, therefore, they are willing to pay more for such a link. If the latter effect is stronger, then a top site wins, otherwise a regularly lower ranked site wins the sponsored link. In reality, these two cases translate to the distinct, observed scenarios we mentioned above. For the word "travel", the sponsored links and search result are distinct. In contrast, for the word "airlines", a site appearing among the top search results also obtains a (top) sponsored link.

The following corollary describes the equilibrium bids.

Corollary 1 *The winning bid in an FNE is*

$$\frac{G(w_1)}{\beta \alpha_1 \gamma_{w_1}(1-\delta I(w_1))} \geq b_1 > \frac{G(w_2)}{\beta \alpha_1 \gamma_{w_1}(1-\delta I(w_1))}.$$

In an SSNE, the winning bid can be arbitrarily high, but the second highest bid is

$$\frac{G(w_1)}{\beta \alpha_2 \gamma_{w_2}(1-\delta I(w_2))} \geq b_2 > \frac{G(w_2)}{\beta \alpha_2 \gamma_{w_2}(1-\delta I(w_2))}.$$

[15] This force is even stronger if we assume that sites with a high CTR have a larger traffic independent from the SE.

Note that the bids largely depend on the parameters. Sites with similar valuations might submit significantly different bids based on their CTR's or their position among the regular search results.

4.2 Bidding strategies for multiple sponsored links

We will now discuss the general case, with multiple sponsored links ($s > 1$). As mentioned before, the first price auction does not work in this case, thus we analyze the SSNE only. Let

$$G_j(i) = R(I(i)\gamma_i\alpha_i + \gamma_i\beta\alpha_j)) - R(I(i)\gamma_i\alpha_i)$$

denote the revenue gain for Site i of winning sponsored link j ($j = 1, ..., s$). Let $w_1, w_2, ..., w_n$ denote the sites in the order of their CTR-corrected bids (F_i's). Furthermore, let P_i denote the total fee that Site i pays for the advertising:

$$P_i = b_{w_{i+1}}\alpha_{w_i}\beta\gamma_{w_i}(1 - \delta I(w_i)).$$

The search engine ranks the sites according to their CTR-corrected bids, that is, if the order is $w_1, w_2, ...,$ then the following have to hold for $2 \geq i \geq s$:

$$\frac{P_{i-1}}{\alpha_{i-1}} > \frac{P_i}{\alpha_i}. \tag{1}$$

In any equilibrium, Site w_k does not have an incentive to bid less and get to a lower position. Therefore,

$$G_k(w_k) - P_k \geq G_l(w_k) - P_l. \tag{2}$$

Furthermore, according to the definition of a symmetric equilibrium, Site w_l does not want to get into position k even if it has to pay P_k (and not P_{k-1}). That is,

$$G_l(w_l) - P_l \geq G_k(w_l) - P_k. \tag{3}$$

Combining (1), (2) and (3), we get the following inequalities, describing the equilibria of the auction:

$$G_k(w_k) - G_l(w_k) \geq P_k - P_l \geq G_k(w_l) - G_l(w_l). \tag{4}$$

The complexity of the problem does not allow us to characterize all the SSNEs in this general case. Multiple equilibria may exist, where the order of winners is different. The following example illustrates the complexity of the problem even in a simple case.

Example 1 Assume $s = 2$ and $n = 3$, with the following valuations:

$$G_1(1) = 10,\ \ G_2(1) = 8,\ \ G_1(2) = 9,\ \ G_2(6) = 6,\ \ G_1(3) = 8,\ \ G_2(3) = 7.$$

These gains can be derived from a suitable $R()$ function, γ-s and α-s. Note that with prices $P_1 = 9$ and $P_2 = 7$, the equilibrium order of sites can be either $(w_1 = 1, w_2 = 3, w_3 = 2)$ or $(w_1 = 2, w_2 = 1, w_3 = 3)$.

To solve for the maximum and minimum revenue equilibria in the general problem, we would have to solve the linear program defined by (1) and (4) for every i, k and l. While this problem is still very complex, with a minor restriction, we can easily solve it.

Definition 1 *We say that the preferences of sites i and j are aligned, if $G_1(i) > G_1(j)$ implies $G_k(i) - G_l(i) > G_k(j) - G_l(j)$ for every $1 \leq k, l \leq s + 1$.*

The assumption of aligned preferences is rather natural. It means that there is a consensus between players about the value of different positions. With this, we can determine the equilibrium ranking of sites.

Lemma 1 *In any SSNE, $G_k(w_1) \geq G_k(w_2) \geq ... \geq G_k(w_{s+1})$ for any $1 \geq k \geq s + 1$.*

In order to fully describe the equilibria we also have to assume that sites' valuation for the position they are in is high enough relative to the next site's valuation of the next position. Specifically, we assume that

$$G_j(w_j) - G_{j+1}(w_j) > \frac{\alpha_j - \alpha_{j+1}}{\alpha_{j+1} - \alpha_{j+2}}(G_{j+1}(w_{j+1}) - G_{j+2}(w_{j+1})) \tag{5}$$

holds for every $1 \geq j \geq s-1$ (see the Appendix for more details on this assumption). With these assumptions, we can describe the SSNE, following the path proposed by Varian (2007).

Proposition 2 *If all the sites' preferences are aligned and (5) holds, then an SSNE exists. Furthermore,*

1. *The maximum SSNE income of the seller is*
$$M(s) = \sum_{j=1}^{s-1} [j(G_j(w_j) - G_{j+1}(w_j))] + sG_s(w_s).$$
2. *The maximum SSNE income is equal to the maximum SNE income.*

The results are similar to the case in which there is only one sponsored link to bid for. The set and order of winners is determined by two factors. Sites with higher traffic from other sources, such as regular search results, have a lower marginal valuation for traffic, however sites with higher CTRs value sponsored links higher. It is clear that the order among those sites that do not receive regular search results will be decreasing in the CTR, that is, $r+1, r+2, ..., n$. However, the top r sites may end up in any position depending on their parameters.

Example 2 Let us consider an example of twenty sites competing for five sponsored links with the following parameters: *$n = 20$, $r = 10$, $s = 5$, $\gamma_i = 0.5 - 0.025(i-1)$, $\alpha_i = (20-(i-1))/232.5$, $\beta = 0.5$, $\delta = 0.6$, and $R(x) = \log(1+30x)$.* Then site 11 gets the top sponsored link, followed by sites 12, 3, 4, and 2.

Figure 2 shows the valuations of the twenty sites for the five sponsored links. The parameters are such that sites 11 and 12 have the highest valuations for the sponsored links because they are the sites with the highest CTRs that are *not* listed among the regular search results. Since the advertising response function is concave, these sites have a higher marginal valuation for a click. As a result, the winner of the first sponsored link is Site 11, followed by Sites 12, 3, 4, and 2. Figure 3 shows the equilibrium prices the sites pay and the bids they submit. Here, the sites are listed in their order of appearance. It is not surprising, that the total fee they pay is decreasing with the position they are in. However, it is interesting to see that higher per-click bids do not automatically lead to a better position. Generally, sites with higher inherent CTRs do not need to bid too high, however, top sites (such as 3, 4 and 2) still have to bid higher than others for the same position because their higher CTRs guarantees them a position on the SE's search results list which, in turn, directs traffic away from the sponsored link. In our example on Figure 3 the 6th site's bid is higher than that of the 5th site but this site did not manage to fetch a sponsored link.

In summary, the present model explains why sponsored links may exhibit peculiar and seemingly unpredictable patterns on SEs' search pages. Top sites in terms of CTR will rank high on the SE's search results list, therefore are likely to benefit less from advertising links. Furthermore, from the SE's perspective, even if they bid high for a sponsored link, consumers may actually click on the search result link instead. These two effects may cause sites with lower CTRs to win the auction on the sponsored list. However, if the popularity of a site is large enough compared to secondary sites then these effects are not enough to compensate for the inherent advantage of a site in directing traffic to itself and top sites may still end-up high on the list of sponsored links. Thus, the presence and order of sites on the sponsored links list is a result of many interacting factors, including the sites' inherent popularity (CTR) and - more importantly - consumers' clicking behavior on the search page.

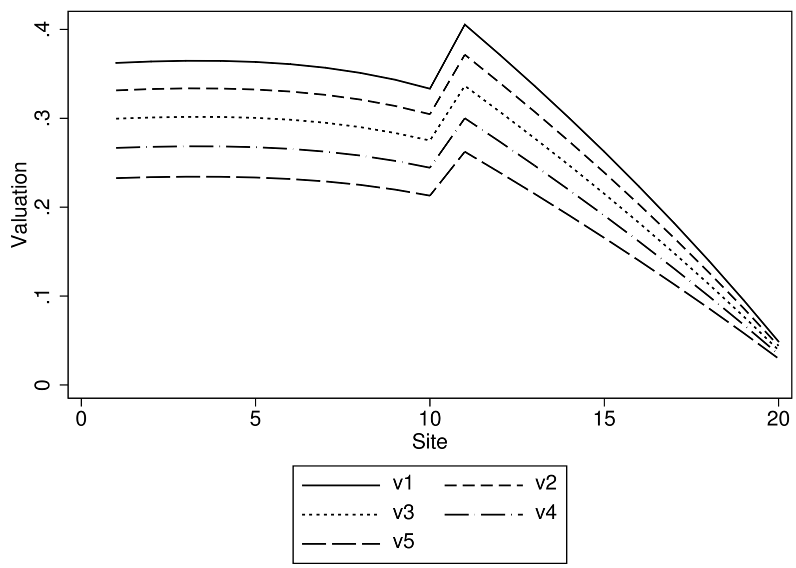
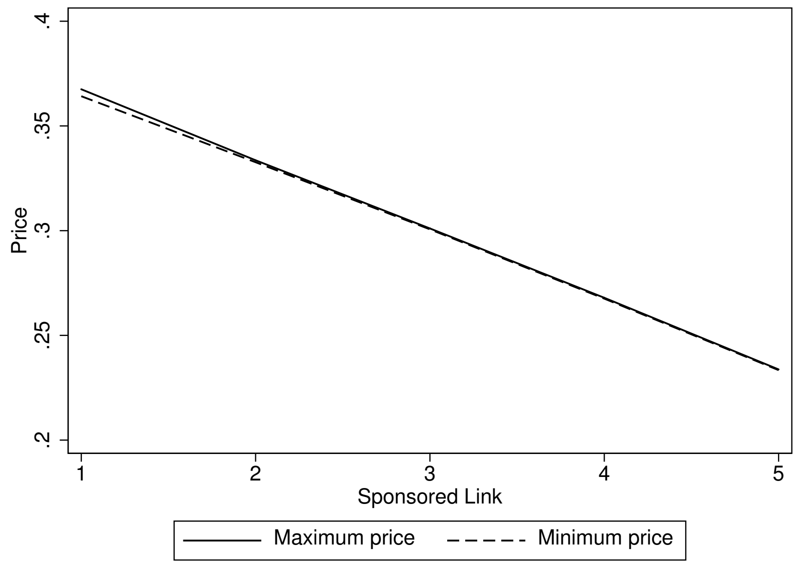


Figure 2: Sites' valuation of the five sponsored links. The parameters are: $n = 20$, $r = 10$, $s = 5$, $\gamma_i = 0.5 - 0.025(i-1)$, $\alpha_i = (20-(i-1))/232.5$, $\beta = 0.5$, $\delta = 0.6$, and $R(x) = \log(1+30x)$.



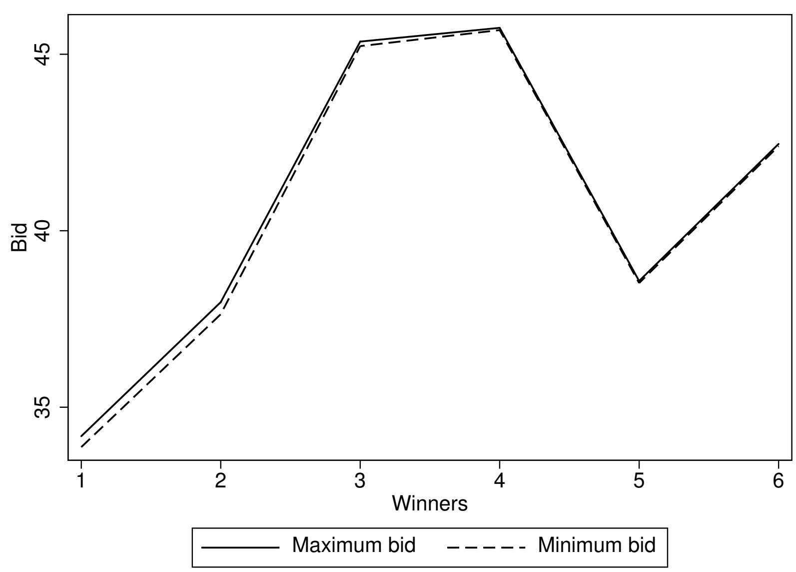


Figure 3: Fees payed by and bids submitted by the five winners in their order of appearance: (11,12,3,4,2). The parameters are: $n = 20$, $r = 10$, $s = 5$, $\gamma_i = 0.5 - 0.025(i-1)$, $\alpha_i = (20-(i-1))/232.5$, $\beta = 0.5$, $\delta = 0.6$, and $R(x) = \log(1+30x)$.

The model shows how behavioral measures of α, β and δ can help SEs as well as web sites to better optimize their strategies.

5 Extensions

5.1 Heterogeneity in sites' valuations

In the model, we assume that sites value incoming traffic similarly. The rationale behind this assumption is that for a given word, there is a standard rate of converting traffic to revenues and most sites have the same $R(t)$ function. However, there might be cases in which sites are heterogeneous with respect to their valuation of traffic. For example, as a result of its branding strategy, a company may have an incentive to attract more traffic to increase its brand recognition resulting in higher long-term profits. Here, we examine the implications of heterogeneity in sites' valuation for traffic. Let us assume that site i has the return function

$$R_i(t) = \vartheta_i R(t),$$

where ϑ_i denotes Site i's traffic conversion parameter. That is, every site has a similarly shaped traffic return function, but there are individual differences in how sites can make revenues from one visitor. Then, using previous notation, the gain for Site i of winning the sponsored link j is

$$G_j(i) = \vartheta_i[R(I(i)\gamma_i\alpha_i + \gamma_i\beta\alpha_j)) - R(I(i)\gamma_i\alpha_i)].$$

With these modified gain functions, we can apply Proposition 2 (the conditions do not change). The results are similar, but we can observe a simple effect of higher valuation for traffic. It simply boosts sites' willingness to pay for sponsored links, thus sites with a higher valuation get a better sponsored link. In the extreme case, when sites have similar inherent CTRs ($\gamma_1 - \gamma_n \to 0$) and an extra visitor results in constant extra revenue ($R(t)$ is linear), this effect dominates and sites' valuation for traffic (ϑ_i) determines the order of sponsored links. In a typical case, however, like Example 2, it is combined with the other previously discussed factors. Let us consider Example 2 again and assume that $\vartheta_i = 1$ for all i, except for Sites 2 and 3, for which, $\vartheta_2 = \vartheta_3 = 1.1$. Then, the order of the five sponsored links changes to $11, 3, 2, 12, 4$ from $11, 12, 3, 4, 2$. That is, Sites 2 and 3 improved their position because they value an extra visitor relatively higher, but still could not get in front of Site 11 which values visitors even higher because it does not receive traffic from organic search results. Also, the relative order of Sites 2 and 3 did not change because their valuations were increased to the same extent.

In summary, heterogeneity in sites' valuation for traffic does have an effect on the order of sponsored links and the bids. A higher valuation leads to higher bids resulting in a better position in the sponsored links list. An interesting aspect of these valuations is that, presumably, this is where sites may have some private information in the sense that sites do not perfectly know each other's ϑ_i's. Although it is out of the scope of the present paper to solve an incomplete information scenario, we can speculate that a second price auction would lead to sites' revealing their valuations. However, it is not simply the case that a site with higher valuation for traffic bids higher, one has to combine the different effects described before and carefully examine all the factors.

5.2 Endogenizing the number of sponsored links

So far we have considered the number of sponsored links displayed by the search engine given. In this section, we compare the search engine's revenue in cases of offering different numbers of links. For the sake of simplicity we assume a linear revenue function, that is, $R(t) = at$. Then $G_k(i) = \beta\gamma_i\alpha_k$. We assume that the search engine makes a decision about the number of sponsored links and announces it prior to the auction. When it makes the decision it has to take into consideration two forces. First, if it offers more links for sale, it will receive payments from more sites. However, when the number of links is increased, the traffic flowing through each one goes down. Let us compare the cases when the search engine offers s sponsored links and when it offers $t < s$ instead. If $\beta\alpha_j$ is the traffic going to sponsored link j in the first case, then in the second case, it increases to

$$\beta\alpha_j' = \beta\alpha_j(1 + \beta\alpha_{t+1} + ... + \beta\alpha_s).$$

As we saw in the previous section, there are usually many equilibria and the revenue of the SE cannot be determined. Here, we will only compare the maximum revenues the SE can attain by selling different number of sponsored links.

Proposition 3 *The SE can attain a higher maximum revenue by offering $t < s$ sponsored links instead of s, if and only if,*

$$\beta(\alpha_{t+1} + ... + \alpha_s)\left(\sum_{j=1}^{t} j\gamma_j\alpha_j - \sum_{j=1}^{t-1} j\gamma_j\alpha_{j+1}\right) > \sum_{j=t+1}^{s} j\gamma_j\alpha_j - \sum_{j=t}^{s-1} j\gamma_j\alpha_{j+1}.$$

Decreasing the number of sponsored links increases the traffic on the remaining ones. Thus, the sites are willing to pay more for them. The LHS of the inequality is equal to this benefit. However, by forgoing sponsored links $t+1$ to s, the SE loses $s-t$ advertisers. The resulting loss is the RHS of the inequality. Note that the RHS is sometimes negative, that is, even without the increased traffic on the remaining links the SE may have an incentive to decrease the number of links. This is a result of the fact that the value of sponsored links increases in the advertisers' eyes and they are willing to pay more for them.

Example 3 Assume that $s = 2$ and $t = 1$. The SE is better off offering one link, iff,

$$\beta\alpha_1 > \frac{2\gamma_2 - \gamma_1}{\gamma_1}.$$

In essence, the SE should offer only one sponsored link when the second highest CTR is relatively low. In particular, if $\gamma_2 < \gamma_1/2$, then the SE is better off selling one link even if the second link still drains the traffic. More generally, the SE should only add additional links as long as the CTR of an additional site getting that link is relatively high. In other words, if there is a sharp drop in the top CTRs after the i-th site then selling more than i sponsored links may not make sense.

5.3 Dynamic bidding for sponsored links

In the previous models, we assumed that the process through which the sponsored links are assigned is a one-shot game. However, in reality, the auctions for the links take place repeatedly. We cannot always ignore the effects that previous bids and results have on the current auction. An important effect is, for example, that when a site wins a sponsored link, the traffic that it receives through the link may have a lagged effect. Such lagged effects have been documented in Rutz and Bucklin (2007a). Some consumers who get to a web site through advertising may become regular customers of the site. If they want to return to the site they do not need the sponsored link again, they may remember or "bookmark" the site's address. This effect however, decreases with time. For the sake of simplicity, we assume that it lasts only for one time period and that there is only one sponsored link. Precisely, if a consumer arrives from the SE to the site in a given time period, then with probability q s/he will return in the next period without the use of the search engine. Then, if Site i receives traffic t_i from the search engine in a given period, then the lagged effect of this traffic is qt_i in the next period.

Now let us examine how this effect changes sites' valuations of the sponsored link. If a site did not win the sponsored link in the previous period then the gain associated with winning it is

$$G_l(i) = R((1+q)I(i)\gamma_i\alpha_i + \gamma_i\beta\alpha_1)) - R((1+q)I(i)\gamma_i\alpha_i),$$

where we also deal with the lagged effect of regular search results. On the other hand, if the site did win the sponsored link in the previous period, then its gain is

$$G_w(i) = R((1+q)I(i)\gamma_i\alpha_i + (1+q)\gamma_i\beta\alpha_1)) - R((1+q)I(i)\gamma_i\alpha_i + q\gamma_i\beta\alpha_1).$$

Therefore, if $R()$ is strictly concave, then $G_w(i) < G_l(i)$. The intuition is that due to decreasing marginal returns, the site values the sponsored link less, if it has already won the link in the previous period. To solve the repeated game we use the concept of Markov-perfect equilibrium, where players' actions only depend on the states of the world. In this case, the states represent the possible winners of the auction and when a site wins the auction, the world moves to that state. In such an equilibrium, forward looking players choose their strategies to maximize their profits over time using the discount factor δ. Let $V_i^{(j)}$ denote Site i's discounted equilibrium profits counted from a period, when the previous winner is Site j. Sites' payoffs in the current period will be determined by their bids. If Site i does not win the auction, it does not make any profit in the current period, that is, its overall discounted profit will be

$$\delta V_i^{(w)},$$

where w is the winner of the current auction. On the other hand, if Site i wins the auction then it will make a profit of $v_i = G_w(i) - P$ if $i = j$ and $v_i = G_l(i) - P'$ if $i \neq j$, where P and P' are the prices the winner has to pay (these depend on the bids). Therefore, its overall discounted profit will be

$$v_i + \delta V_i^{(i)}.$$

In equilibrium, player i chooses its bid to maximize this quantity, that is,

$$V_i^{(j)} = \max_{b_i}(\delta V_i^{(w)}, v_i + \delta V_i^{(i)}),$$

where w and v_i both depend on b_i.

Since there is only one sponsored link, we can use the first price auction's equilibrium and the second price auction's symmetric equilibrium concepts interchangeably. We will determine the Markov-perfect first price Nash-equilibria (MFNE) and Markov-perfect second price symmetric Nash-equilibria (MSSNE) of the game. Regarding the valuations, let us assume that only the first two sites have a high enough valuation to win the auction, that is $G_l(j) < \min(G_w(1), G_w(2))$ for $j \geq 3$. Then, we only have to examine the auction where Sites 1 and 2 bid for the link.

Proposition 4

1. *If $G_l(2) < G_w(1)$, then Site 1 is the winner in every period and*

$$G_w(1) \geq P_1 \geq G_l(2).$$

2. *If $G_l(1) < G_w(2)$, then Site 2 is the winner in every period and*

$$G_w(2) \geq P_1 \geq G_l(1).$$

3. *In every other case, the two sites alternate winning.*

In essence, if a site values winning the link for a second time higher than the other site does for the first time, then that bidder is the winner always. Otherwise, the two sites alternately win and lose the auction. The intuition is that when a site wins the link in one period, then its valuation goes down in the next period and the other site is willing to pay more for the link. Now that this other site wins the auction, the valuations will again cross each other leading to the alternation. Therefore, the only way one site can win the auction in every period is if its valuation dominates the other site's valuation in the sense that, even after winning it, the site is willing to bid more than its losing competitor.

Interestingly, and consistently with our focus, the search results play an important role here. In addition to the natural results when the best site wins always (the one which is the first in the organic lists), there are scenarios under which the site in the second position wins all the time. This can happen for example, if the first hits in the lists are much better than the second positions. In this case, the site which is in lower position on the organic list, competes very aggressively for the sponsored link. Also, for words for which the traffic return function ($R(t)$) is very steep for the first few visitors and then becomes flat quickly (a very negative second derivative), the site which already has many visitors from organic search has a low incentive to compete. The typical situation however, is that neither site's valuations dominate the other's and consequently, they alternately win the sponsored link. Next, we examine how the strength of the lagged effect (the value of q) affects the outcome.

Corollary 2 *There is a $q^* > 0$, such that*

- *if $0 < q < q^*$ then the winner is always the same site,*
- *if $q^* < q$ then the two sites alternate as winners.*

In other words, as the ratio of returning customers increases, at one point the type of equilibrium changes and the two sites start winning alternately. This critical value is smaller if the marginal return on traffic decreases quickly.

Finally, let us compare the search engine's income in the different cases. Let us assume that $G_l(1) > G_l(2)$, that is, either Site 1 wins the link always or sites alternate winning.

Then the search engine's maximum discounted income (in the two cases, respectively) is[16]:

$$M_1 = \frac{G_w(1)}{1-\delta},$$

$$M_2 = \frac{G_l(1) + \delta G_l(2)}{1-\delta^2}.$$

It is worth noting, that M_1 and M_2 not only represent the SE's maximum income in the two cases, but also the total surplus of all players (SE and sites) in all the equilibria of the given type[17]. It is an interesting question what happens to this surplus, when the type of the outcome changes. Are the sites and the search engine better off under an alternating winning scenario or with a fixed winner? We compare these values around the boundary of the two regions, which separates the alternating and non-alternating equilibria, that is, where $G_w(1) = G_l(2)$.

Corollary 3

$$\lim_{G_w(1)-G_l(2)\to 0^+} M_1 < \lim_{G_w(1)-G_l(2)\to 0^-} M_2,$$

and the difference increases in q and δ.

We find a discontinuity in the total income at the boundary of the two regions, because the SE and the sites are strictly better off in the case of an alternating equilibrium. The intuition is that the alternating assignment of the SE's traffic is a more efficient allocation than when one site is the winner in every period. This extra revenue is higher if the ratio of returning consumers is higher and if the discount rate is higher. Whether the SE or the sites appropriate this extra revenue depends on the actual bids. The key insight however, is that *all* players are better off in an alternating equilibrium. Again, knowing consumers' behavior on the search page, the SE can influence the design of the auction to increase the likelihood of such an outcome.

[16]The proof can be found in the proof of Proposition 4.
[17]Individual incomes depend on how this surplus is divided in a given equilibrium.

6 Conclusion

In this paper, we have modeled the race for sponsored advertising links on a search engine's page between web sites endowed with different click-through rates. We argue that the SE's problem can not simply be described as a multi-item auction. The existence of the search results list on the SE's page represents an important externality for both types of players. In addition to exploring the effect of this externality on the allocation outcomes we also study two other issues: the endogenous choice of the number of sponsored links, and the dynamics of the bidding behavior.

Our key result is that we explain the mechanism that may lead to wildly different patterns observed in the behavior of sponsored links. In particular, top sites who rank high on the SE's search results list are likely to benefit less from advertising links. Furthermore, from the SE's perspective, even if they bid high for a sponsored link, consumers may actually not click on this link but rather click on the organic link instead. These two effects may cause secondary sites to end up winning the auction on the sponsored list. On the other hand, if the popularity of a site is high enough compared to secondary sites then the above effects are not enough to compensate for the inherent advantage of a site in directing traffic to itself and top sites may still end up high on the list of sponsored links.

We also explore three extensions. First, we relax the assumption that sites value traffic uniformly. We find that while sites valuations matter in terms of the actual bids, the basic competitive mechanisms remain the same. Second, endogenizing the number of sponsored links allocated by the SE, we show that the SE can increase traffic flowing through sponsored links by decreasing the number of these links. A decrease in this number increases the value of the links and may result in compensating for the loss associated with a smaller number of links. Finally, we examine a dynamic model in which online advertising has a lagged effect on the site that wins the sponsored link. We identify dynamic bidding patterns that lead to alternating or constant allocations of the sponsored links, depending on the strength of the lagged effect. Interestingly, we find that the search engine and the sites together are strictly better off under an alternating equilibrium.

Our analytic results have interesting normative implications. Our core result may help search engines refine the weights attributed to sites' bids for sponsored links. By explicitly measuring the parameters describing consumers' behavior on a search page, the weights attributed to bids can be corrected beyond the sites' CTRs. We also provide insights with respect to when a SE should add/substract a sponsored link from the page. In particular, we find that this decision primarily depends on the distribution of CTRs across sites. When a sharp drop occurs in this distribution, then the SE should stop adding sponsored links to the page. Finally, our analysis of the dynamic game suggests that the SE should try to promote an alternating bidding pattern between sites. Again, understanding consumers' behavior on the search page and maybe influencing it might help the SE to do so.

6.1 Limitations

Our model also has a number limitations. First, when modeling consumer's behavior on the results page, we assumed that a person either clicks on one link or leaves the page without clicking. In reality, someone can click on one link then return to the search page if s/he is not satisfied with that particular link and click on another link. To account for multiple visits, we could apply our assumptions "per visit" and not "per person". This way, we model each visit separately and one person can make several visits. These visits, however, have to be independent is this setting. Clearly, it is a limitation of our study that we ignore the possible connection between these visits.

Second, in reality sites not only place bids according to what they are willing to pay for a click, but they can also set daily or monthly budgets. Then, there is an automatic system that submits the site's bid continuously until the budget is reached, after which the system automatically withdraws their bids. We do not model this feature since, in our model, sites can perfectly estimate how much traffic they get through a sponsored link. However, it is a limitation of our model that it does not consider uncertainty regarding the number of clicks on a link. Furthermore, in the last extension, we only model repeated bidding in a discrete setting in which sites submit bids for each consecutive time period, but not continuously.

Third, while we have explored the case when sites have different valuations, it is likely that part of these valuations are private information and do not get perfectly revealed through repeated bidding. In this case, one would need to deal with a game of incomplete information. While we do not believe that our qualitative insights with respect to the interaction of the search list and sponsored links list would change, one would need to carefully compare the different auction mechanisms to solve the SE's problem. We have left this to future research.

Finally, throughout the paper we assumed that every consumer is interested in the same topic and the results include the same pages for every query. Obviously, this is rather unrealistic and the allocation of sponsored links in relation to a given search word changes when multiple *interacting* search words are considered. As reported in Rutz and Bucklin (2007b), most advertisers manage/bid for a bundle of key words. Web sites may offer content in every topic, although their relevance may vary from topic to topic. In other words, the inherent CTRs may be different for the same site in different topics. For example travelocity.com may have a high CTR in the context of travel but most likely has a lower one when consumers are searching for home appliances.

To overcome this limitation, we have conducted additional analysis (available on request) with the following results. If the search words are unrelated, then the site with highest valuation wins the auction for each word. On the other hand, if the words are related and presumably the same site has the highest valuation for different words, then either that site wins all the auctions or it wins only a few of them. In this latter case, the intuition is that winning one auction boosts the winner's traffic, therefore, it does not value the traffic in other auctions that high, leaving the opportunity to the site with the second highest valuation to win there. In the extreme case, if there are nearly as many related words as bidders, then even the site with the lowest valuation can end up winning a sponsored link.

The pricing of search advertising is a dynamic field that provides a fertile area for future research. Rather than focusing on various auction mechanisms, our goal was to concentrate on the interaction (conflict) between the SE's core business as a reliable source of information and its business as an advertiser. Our results provide insights on how to minimize the conflict between these business objectives. Clearly, there are many possible ways in which the present analysis can be extended, including empirical work to test some of the analytic results.

Appendix

Proof of Claim 1:

The search engine wishes to maximize the income from the s winners of the sponsored links. Given the order of sites it is obviously optimal to set the p_i's to the maximum, that is, $p_i = b_i$, because it does not affect sites' bidding strategies as sites only place bids once, in the first stage of the game. Regarding the order of sites, if Site i acquires a sponsored link, the search engine will receive a total payment of $\beta A(i) F_i$ from that site, where $F_i = \gamma_i b_i (1 - \delta I(i))$. The F_i values are site specific and only depend on the site's parameters, whereas the $A(i)$ values are determined by the search engine, when it assigns the sponsored links. In order to maximize $\beta \sum_{i=1}^{n} A(i) F_i$, the SE has to assign the α's in a decreasing order of the F_i values. $\square$

Proof of Proposition 1:

As we have discussed before, the winner – both in an FNE and an SSNE – is the site with highest valuation, The payment of the winner is between the first and second valuations. $\square$

Proof of Lemma 1:

If sites' preferences are aligned, then (4) yields $G_1(w_l) \geq G_1(w_m)$ for every $l < m$, proving the lemma. $\square$

Proof of Proposition 2:

In order to prove the existence of an SSNE, we have to show that there exist $P_1 \geq P_2 \geq \ldots \geq P_s$, such that, they satisfy inequalities (2) and (3) for every $1 \leq k < l \leq s$. We will show that if the sites' preferences are aligned, then it is enough to check that $P_1 \geq P_2 \geq \ldots \geq P_s$

satisfy a subset of them, namely the following inequalities, for every j:

$$G_j(w_j) - G_{j+1}(w_j) \geq P_j - P_{j+1} \geq G_j(w_{j+1}) - G_{j+1}(w_{j+1}). \tag{6}$$

We have to show, that all the inequalities in (2) and (3) follow from those in (6). Let $1 \leq k < l \leq s$ be arbitrary indices. Summing (6) for $j = k$ to l, we get

$$\sum_{j=k}^{l-1} [G_j(w_j) - G_{j+1}(w_j) \geq P_k - P_l \geq \sum_{j=k}^{l-1} [G_j(w_{j+1}) - G_{j+1}(w_{j+1})].$$

Since the preferences are aligned, $G_j(w_k) - G_{j+1}(w_k) \geq G_j(w_j) - G_{j+1}(w_j)$ for $j > k$, therefore, we obtain

$$G_k(w_k) - G_l(w_k) \geq P_k - P_l,$$

and similarly

$$P_k - P_l \geq G_k(w_l) - G_l(w_l).$$

We have shown, that the system given by (2) and (3) is equivalent to that defined by (6). That is, it is always enough to check whether a site wants to get to a position which is one higher or lower. Therefore, given that (1) holds, the values of $P_j - P_{j+1}$ can be chosen arbitrarily from the intervals given in (6), fixing $P_{s+1} = 0$. In (5), we basically assume that selecting the maximum values does not violate (1). Thus, we get the second part of proposition by summing the left hand sides of (6) in the following way.

$$\sum_{i=1}^{s} P_i = sP_s + \sum_{j=1}^{s-1} j(P_j - P_{j+1}).$$

For the fourth part, let us note that every SSNE is an SNE, therefore the maximum SNE income is at least as high as the maximum SSNE income. For the other direction, let P_i^N denote the expenditure of Site i in an SNE with maximum revenue and let P_i^S denote the same expenditure in a maximum revenue SSNE. From the previous part, we know that

$$P_j^S = P_{j+1}^S + G_j(w_j) - G_{j+1}(w_j),$$

However, according to the definition of an SNE,

$$P_j^N \leq P_{j+1}^N + G_j(w_j) - G_{j+1}(w_j).$$

Since $G_{s+1}(w_s) = 0$,

$$P_s^N \leq G_s(w_s) = P_s^S.$$

Then, it is easy to show recursively that $P_i^N \leq P_i^S$, completing the proof. □

Proof of Proposition 3:

According to Proposition 2, the maximum equilibrium revenue of the SE, in case of selling s links, is

$$M(s) = \beta \left(\sum_{j=1}^{s} j\gamma_j\alpha_j - \sum_{j=1}^{s-1} j\gamma_j\alpha_{j+1} \right).$$

If the SE decides to instead sell only t links, the traffic on the remaining links will increase by a factor of $(1 + \beta(\alpha_{t+1} + ... + \alpha_s))$. Therefore, the maximum equilibrium revenue will be

$$(1 + \beta(\alpha_{t+1} + ... + \alpha_s))\beta \left(\sum_{j=1}^{t} j\gamma_j\alpha_j - \sum_{j=1}^{t-1} j\gamma_j\alpha_{j+1} \right)$$

in this case. Comparing the two quantities, we get the expression in the proposition. □

Proof of Proposition 4:

We will assume without loss of generality that $G_l(1) > G_l(2)$. The proof of the opposite case is straightforward. First, we prove the third part of the proposition, that is, identify the conditions necessary for an alternating equilibrium. In such an equilibrium, bidding strategies are such, that if Site i has won the previous auction then Site $j = 3 - i$ is the current winner. Let $P^{(i)}$ denote the fee that Site $j = 3 - i$ has to pay in the auction when Site i is the previous winner. Let $V_i^{(j)}$ denote the discounted equilibrium profits of Site i

from a given period when Site j is the previous winner. In an alternating equilibrium,

$$\begin{aligned} V_1^{(1)} &= \delta V_1^{(2)} \\ V_1^{(2)} &= G_l(1) - P^{(2)} + \delta V_1^{(1)} \\ V_2^{(1)} &= G_l(2) - P^{(1)} + \delta V_2^{(2)} \\ V_2^{(2)} &= \delta V_2^{(1)}. \end{aligned}$$

Therefore,

$$\begin{aligned} V_1^{(2)} &= \frac{G_l(1) - P^{(2)}}{1-\delta^2} \\ V_2^{(1)} &= \frac{G_l(2) - P^{(1)}}{1-\delta^2}. \end{aligned}$$

The sufficient and necessary conditions these valuations and prices have to satisfy are that in a given auction, the winner has to have a higher valuation and the fee payed by the winner must fall between the two players' valuations (both in an MFNE and MSSNE). For example, if the previous winner is Site 1, then the current winner must be Site 2, therefore,

$$G_w(1) + \delta(V_1^{(1)} - V_1^{(2)}) \leq P^{(1)} \leq G_l(2) + \delta(V_1^{(1)} - V_1^{(2)})$$

must hold. Plugging the corresponding formulas, we obtain

$$G_w(1) - \frac{1-\delta}{1-\delta^2}(G_l(1) - P^{(2)}) \leq P^{(1)} \leq G_l(2). \tag{7}$$

Comparing the valuations in a period when Site 2 is the previous winner, we get a similar inequality,

$$G_w(2) - \frac{1-\delta}{1-\delta^2}(G_l(2) - P^{(1)}) \leq P^{(2)} \leq G_l(1). \tag{8}$$

The set defined by (7) and (8) is a two-dimensional simplex. It is easy to see that it is non-empty iff $G_l(2) \geq G_w(1)$ (given the other restrictions on the parameters).

The maximum discounted income of the seller depends on the first period of the game. Let P_t denote its income in period t. If Site 1 is the first winner, then it would be

$$\sum_{t=1}^{\infty} \delta^{t-1} P_t = \frac{P^{(2)} + \delta P^{(1)}}{1 - \delta^2}.$$

If Site 2 is the first winner, then it is

$$\frac{P^{(1)} + \delta P^{(2)}}{1 - \delta^2}.$$

We determine the maximum for both and consider the higher value. Clearly, since Site 1 has higher valuations, the SE's income will be higher if Site 1 is the first winner. Maximizing $P^{(2)} + \delta P^{(1)}$ on the simplex defined by (7) and (8), we get

$$M_2 = \frac{G_l(1) + \delta G_l(2)}{1 - \delta^2}.$$

The first part of the proposition can be proven by following the same steps. However, it is obvious, that since in both states site 1 has a higher valuation, it is always the winner. Then the price payed must be in the given range, yielding the stated maximum income. □

Proof of Corollary 2:

The values of $G_l(1) > G_l(2)$ are independent of q. When $q = 0$, $G_l(i) = G_w(i)$ and as q increases $G_w(1)$ decreases. Let q^* be the unique solution of

$$R((1+q)I(1)\gamma_1\alpha_1 + (1+q)\gamma_1\beta\alpha_1) - R((1+q)I(1)\gamma_1\alpha_1 + q\gamma_1\beta\alpha_1) =$$
$$= R((1+q)I(2)\gamma_2\alpha_2 + \gamma_2\beta\alpha_1) - R((1+q)I(2)\gamma_2\alpha_2).$$

Then, for $0 < q < q^*$, we get the first case in Proposition 4 and for $q^* < q$, we get the second case. □

Proof of Corollary 3:

Fixing $G_l(2)$ in Proposition 4, we can establish

$$\lim_{G_w(1)-G_l(2)\to 0^+} M_1 = \frac{G_l(2)}{1-\delta},$$

$$\lim_{G_w(1)-G_l(2)\to 0^-} M_2 = \frac{G_l(1)+\delta G_l(2)}{1-\delta^2} = \frac{G_l(2)}{1-\delta} + \frac{G_l(1)-G_l(2)}{1-\delta^2}.$$

Hence, the difference is

$$0 < \frac{G_l(1)-G_l(2)}{1-\delta^2} = \frac{G_l(1)-G_w(1)}{1-\delta^2},$$

which clearly increases in q and δ. $\square$

References

Athey, Susan, Glenn Ellison. 2008. Position auctions with consumer search. *Working Paper, MIT.*

Bhargava, Bharat, Mamata Jenamani, Yuhui Zhong. 2005. Counteracting shill bidding in online english auction. *International Journal of Cooperative Information Systems* **14**(2&3) 245–263.

Chen, Yongmin, Chuan He. 2006. Paid placement: Advertising and search on the Internet. *Working Paper, U. of Colorado Boulder.*

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.

Edwards, Steven M., Hairong Li, Joo-Hyun Lee. 2002. Forced exposure and psychological reactance: Antecedents and consequences of the perceived intrusiveness of pop-up ads. *Journal of Advertising* **31**(3) 83–95.

Feng, Juan. 2007. Optimal mechanism for selling a set of commonly-ranked objects. *Working Paper, U. of Florida.*

Feng, Juan, Zuo-Jun Max Shen, Roger L. Zhan. 2006. Ranked items auctions and online advertisements. *Working Paper, UC. Berkeley.*

Goldfarb, Avi, Catherine Tucker. 2007. Search engine advertising: Pricing ads to context NET Institute Working Paper No. 07–23.

Katona, Zsolt, Miklos Sarvary. 2008. Network formation and the structure of the commercial world wide web. *Marketing Science* forthcoming.

Lutz, Richard J. 1985. Affective and cognitive antecedents of attitude towards the ad: A conceptual framework. *in Psychological Processes and Advertising Effects* Edited by Linda F. Alwitt and Andrew A. Mitchell.

Maskin, Eric, John Riley. 2000. Asymmetric auctions. *Review of Econmic Studies* **67** 413–438.

Myerson, Roger B. 1981. Optimal auction design. *Mathematics of Operations Research* **6**(1) 58–73.

Ofek, Elie, Miklos Sarvary. 2003. R&D, marketing and the success of next generation products. *Marketing Science* **22**(3) 355–370.

Rutz, Oliver J., Randolph E. Bucklin. 2007a. From generic to branded: A model of spillover dynamics in paid search advertising. *Working paper, UCLA.*

Rutz, Oliver J., Randolph E. Bucklin. 2007b. A model of individual keyword performance in paid search advertising. *Working paper, UCLA.*

Schlosser, Ann E., Sharon Shavitt, Alaina Kanfer. 1999. Survey of internet users' attitude towards internet advertising. *Journal of Interactive Marketing* **13**(3) 34–54.

Varian, Hal. 2007. Position auctions. *Internation Journal of Industrial Organization* forthcoming.

Villas-Boas, Miguel. 1993. Predicting advertising pulsing policies in an oligopoly: A model and empirical test. *Marketing Science* **12**(1) 88–102.

Wilbur, Kenneth C., Yi Zhu. 2008. Click fraud. *Marketing Science* forthcoming.

Zeithammer, Robert. 2006. Forward-looking bidding in online auctions. *Journal of Marketing Research* **43**(3) 462–76.

Zeithammer, Robert, Pengxuan Liu. 2008. When is auctioning preferred to posting a fixed selling price? *Working paper UCLA.*