On the Effects of Competing Advertisements in Keyword Auctions

Aparna Das
Brown University

Ioannis Giotis
University of Washington

Anna R. Karlin
University of Washington

Claire Mathieu
Brown University

ABSTRACT

The strength of the competition plays a significant role in the efficiency of an online advertising campaign, as well as in traditional ad campaigns: the presence of a competitor's ad makes the customer less likely to buy one's own product. Models for keyword auctions usually assume that, for the advertiser, the value of a click is fixed and independent of the other sponsored ads in the auctions result. However in reality, all clicks are not created equal — the ones leading to a conversion or a purchase are definitely more important. How is the customer's decision affected by the set of competing advertisements presented to him? We propose a new valuation model for keyword auctions that is competition-aware and takes into account the entire set of sponsored results present in the auction. We study properties of our model under the two most interesting mechanisms (GSP, VCG), both analytically and in simulations.

Categories and Subject Descriptors

H.4.0 [**Information Systems Applications**]: General; G.0 [**Mathematics of Computing**]: General

General Terms

Algorithms, economics, theory

1. INTRODUCTION

Online search engine advertising is an appealing approach to highly targeted advertising, and is *the* major source of revenue for modern web search engines such as Google, Yahoo! and MSN. The most common setup is as follows: when a user performs a query at a search engine, he is shown a collection of organic search results that contains the links the search engine has deemed relevant to the search, together with a list of *sponsored links*, i.e., paid advertisements. If the user actually clicks on a sponsored link, he will be transferred to the advertiser's web site. For each such click, in which the advertiser receives a potential customer, the advertiser pays the search engine. Of course, a click is of little value to the advertiser unless the user subsequently proceeds to take an action that results in profit to the advertiser, such as buying a product, signing up for a mailing list, etc. Such a user action is called a *conversion*.

Keyword auctions determine which ads get assigned to which keywords (search terms) and how much each advertiser pays. Because of the explosive growth of online advertising and the rising economic importance of ad auctions, a great deal of recent research has focused on developing mathematical models of these systems, with an eye towards understanding their equilibria, dynamics and other properties from the perspective of users, advertisers and search engines [18, 3, 10, 15, 8].

Most keyword auction models assume that each advertisement shown has an inherent click-through rate and conversion rate that depend only on the slot allocated to that advertisement and on the advertisement itself, regardless of the other advertisements that are shown. Unfortunately, this is unlikely to reflect reality because of well-known *externality* effects: the success of an advertisement depends to a significant extent on which other advertisements are shown alongside it. There are three key reasons for this. First, the consumer has a limited attention span and will likely only focus attention on a small number of ads and even then, only in the case that the ads are highly relevant to the consumer's goals. Second, high quality, directly competitive ads placed side by side in response to a query (e.g. ads by both Honda and Toyota in response to a search on "Japanese cars") will reduce the effectiveness of each ad, as the consumer will be very unlikely to convert on both ads and each diminishes the appeal of the other. Finally, presenting the consumer with multiple comparable choices can sometimes lead to indecision or confusion, which in turn may reduce the chance of a conversion [17]. This phenomenon is well-known across all forms of advertising. For example, in television advertising, television networks go to great efforts to satisfy their advertisers by ensuring an allocation of ads to commercial breaks so that (a) competing advertisements do not appear in the same commercial break, and (b) there aren't too many commercials in a single break [20].

In this paper, we modify the traditional model of sponsored search auctions by introducing a more detailed probabilistic model of user behavior. Our goal is to get a handle on search engine revenue and equilibrium bidding in the presence of externalities. To our knowledge, this is the first sponsored search model in which the conversion rate for a particular ad depends on the set of other ads shown. After

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Fourth Workshop on Ad Auctions July 8–9, 2008, Chicago, IL.
Copyright 2008 ACM X-XXXXX-XX-X/XX/XX ...$5.00.

describing our model in Section 2, we study it under the GSP and VCG mechanisms in Section 3. In Section 3.1 we prove that it has a pure strategy Nash equilibrium under the GSP mechanism. In Section 4, we present an experimental evaluation of properties of the model from the perspective of the consumer, the search engine and the advertisers. Finally, in Section 5, we discuss the strengths and weaknesses of our models and propose some alternative directions for research.

Within the field of economics, there is an extensive literature on externalities, but to our knowledge, none of it is relevant to the problem we study here. The most closely related paper that we are aware of is a study by Ghosh and Mahdian [12] of externalities in online advertising in a setting known as online lead generation, where leads (in the form of personal information of a potential customer) are sent to companies or advertisers interested in such leads. The advertisers then contact the potential customer directly to offer quotes and information about their service. In their paper, they present models for externalities in online lead generation and study the complexity of the winner determination problem and the design of incentive compatible mechanisms.

2. MODEL AND DEFINITIONS

2.1 Motivation

We attempt to understand the effects of competitive advertisements by focusing on the most interesting searchers from the perspective of the advertisers, those who are considering making a purchase. As usual, we assume that the mechanism in use is the Generalized Second Price mechanism (GSP), with advertisements ranked by revenue (the product of relevance and bid). The "relevance" r_j of ad j is usually defined as the slot-independent probability that that particular ad will be clicked on; more broadly, relevance measures the quality of ad i. Thus, a searcher will be more likely to click on an ad with higher relevance and, more importantly, *more likely to convert on an ad with higher relevance*. However, the probability of conversion depends on the *set* of ads shown.

In our model, we postulate the following user behavior:

1. Let S be the set of displayed advertisements. We imagine the user's eyes starting at the top of S and scanning down the list, where the probability of scanning the ith ad, given that the $i-1$st ad has been scanned, is θ_i/θ_{i-1}, (with $\theta_0 = 1$).

2. Given that the user looks at (scans) the ith ad, the probability that he actually clicks on it is r_i (the relevance of the ad in slot i).

3. Finally, once the user stops looking and clicking, let C be the set of slots clicked on and let $R_C = \sum_{i \in C} r_i$. We postulate that the probability that the user then converts is R_C/R_S, where R_S is $\sum_{i \in S} r_i$, the sum of the relevances of all the displayed ads. Given that a conversion occurs, the probability that a purchase is made from advertiser $i \in C$ is r_i/R_C. We assume that there is only one purchase made.

The model is illustrated in Figure 1.

The motivation for these choices is as follows:

Our description of the user's eye movement and clicking behavior is consistent with two key facts. First, the probability that slot i is clicked on in our model is

$$\theta_1(\theta_2/\theta_1)\cdots(\theta_i/\theta_{i-1})r_i = \theta_i r_i,$$

matching the standard model for the probability that the ad in slot i gets a click (where θ_i is the ad-independent clickthrough rate of slot i and r_i is the slot-independent clickthrough rate of the advertisement allocated to slot i). Secondly, our model is consistent with eye-tracking studies that show the probability that a user looks at an ad decays with the slot number [13, 16]. Our description of the user's eye movement and clicking behavior has been studied under the term "directional market" in several economics papers [5, 4]. As stated in [4] "the directionality arises due to cognitive burden as it is cognitively "costlier" for a typical consumer to visit sellers at the bottom of the list before visiting the sellers at the top of the listing".

Our assumption that the probability that the user decides to make a conversion is R_C/R_S is based on the theory that if a user bothers to click on most of the high quality ads (as measured by the total fraction of relevance clicked on), he is very seriously interested in what he is doing and more likely to decide to make a purchase. If a user clicks on only a small fraction of the high quality ads, he is not a very serious buyer, and may have just been browsing. The findings of empirical studies [1, 2] show that typically the user goes through a circular process interacting with *multiple* search engine results before making an actual conversion. And finally, given that the user has decided to make a purchase, we assume that he will choose to make the purchase from one of the ads he has clicked on and he will purchase from a particular advertiser a with probability proportional to a's quality or relevance.

It is generally accepted in the marketing world that an advertising campaign is more successful in the absence of other directly competing campaigns. In the sponsored search realm, this translates to advertisers valuing their entries more if they are presented together with a set of irrelevant entries. On the other hand, if their entry is presented alongside similarly or even more relevant advertisements, the effective advertising value is reduced. This is precisely what happens in our model. The user's final decision on whether, and if so, what to buy is dependent on the relevances of the entire list of sponsored results. The value the advertiser perceives is proportional to its relative "strength" amongst the displayed ads. Indeed, a simple calculation shows that in our model, the probability that a customer converts on the ad in slot i is

$$\theta_i r_i (R_C/R_S)(r_i/R_C) = \frac{\theta_i r_i^2}{\sum_{1 \le j \le k} r_j}.$$

2.2 Model

We now formalize the previous discussion.

Definition 1. *The* externality model *for keyword auctions is defined as follows.*

- ***The players and slots:***
  - *A set of n players (advertisers) participate in the auction where each player i has a private valuation v_i for a conversion (an end-user completing*

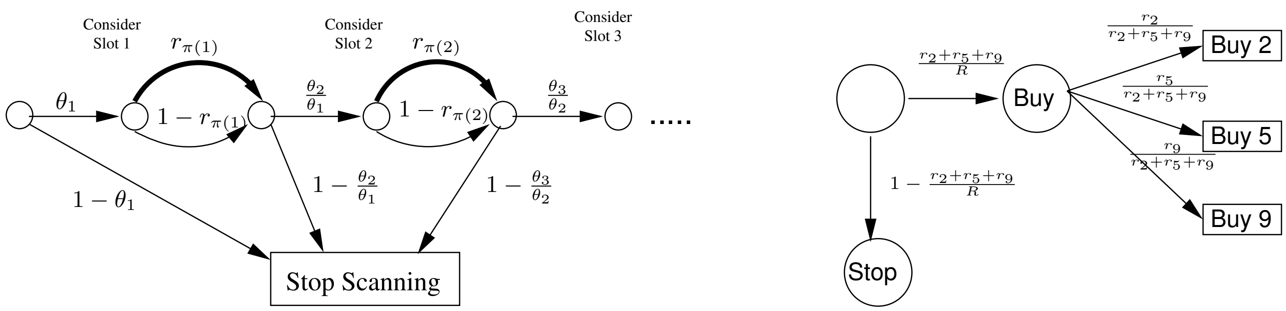

Figure 1: The Markov chain model for customer clicking behavior is shown on the left. The bold edges correspond to clicks. The right side shows the customer buying behavior after he has clicked on slots 2, 5 and 9. Here R is the sum of relevances of all the ads displayed.

a transaction through them). Each player i's ad has an associated relevance *or* quality $0 < r_i \leq 1$.

- *A set of $k \leq n$ slots with* click-through rates (CTR) $\theta_1 \geq \ldots \geq \theta_k$ *where $\theta_s \cdot r$ represents the probability that an advertisement with relevance r would receive a click when placed in slot s.*
- *With knowledge of the auction mechanism used and their own private valuations, each player submits a bid. Player i's bid is denoted by b_i.*

- ***The GSP mechanism:***
  - *computes an allocation π of the slots to k different players by ranking the players in decreasing order of $r_i \cdot b_i$. Here $\pi(s)$ is the identity of the player that is allocated slot s, and has the s highest $r_i \cdot b_i$.*
  - *charges price $p_s = b_{\pi(s+1)} r_{\pi(s+1)} / r_{\pi(s)}$ to player $\pi(s)$ when the user clicks on his ad.*

- ***The end-user:***
  - *Assuming the mechanism's output places player i's advertisement in slot s, the end-user clicks on that entry with probability $\theta_s \cdot r_i$. Let R_S be the sum of the relevances of the displayed advertisers, i.e, $R_S = \sum_{j=1}^{k} r_{\pi(j)}$. The probability that a customer converts on player i's ad, when player i is in slot s is $\theta_s r_i \frac{r_i}{R_S}$.*

- ***Players' Utility function:***
  - *For an allocation π, the expected utility of player i in slot s is given by*

$$U_{i,s} = \theta_s r_i \left(\frac{r_i v_i}{R_S} - \frac{b_{\pi(s+1)} r_{\pi(s+1)}}{r_i} \right) \quad (1)$$

Intuitively, if every customer clicked on i's ad, his revenue per click would be $v_i r_i / R_S$, as he makes v_i per conversion and a conversion occurs for him with probability r_i / R_S. To obtain the formula above, player i's payment per click is subtracted from his revenue and the result is multiplied by his expected number of clicks.

2.3 Some Properties of the Model

It is worth observing a number of properties of our model.

Externality property. Let c_i be the probability of a conversion on the ad in slot i given a click and let r_i be the relevance of the ad in slot i. Then a simple calculation shows that $\frac{dc_i}{dr_i} > 0$ and $\frac{dc_j}{dr_i} < 0$ for $i \neq j$. In other words, as the quality of ad i increases, the probability of that ad converting given that it is clicked on increases and the probability of other ads converting given that they are clicked on decreases. This is what we would expect to happen in the presence of externalities. Notice also that as r_i increases, the probability of a click on the ad in slot i increases, but the probability of a click on ads in other slots is unaffected[1].

Too-many-choices property. Let P be the overall probability that the customer ends up buying anything (i.e. converting on any of the displayed ads). Then $\frac{dP}{dr_i}$ can be either positive or negative. In particular if r_i is one of the higher quality ads, then P tends to increase as r_i increases, whereas if r_i is one of the lower quality ads, it tends to decrease as r_i increases[2].

An explanation for why this might arise in practice is that it is harder to make a decision when there are many roughly equivalent choices [17]. Thus, when a higher quality ad's relevance increases, it makes the buyer's task even easier, whereas when a lower quality ad's relevance increases, it makes the buyer's task more difficult, as he is compelled to choose among multiple equivalent options.

In Section 5, we will present further discussion of the pros and cons of this model, as well as present some alternative models of externalities that could be explored.

3. MODEL ANALYSIS

We now proceed to analyze GSP and VCG in the new model. We begin by showing that GSP still has pure Nash equilibria. It is interesting to note that in the equilibria we find the winning players may not be the same as the winning players in the equilibria of the standard model, nor, will the ordering of the players necessarily be the same. This is because in our model, relevance plays a more important role, resulting in players with lower values but higher relevances

[1]This latter property is also true in the standard keyword auction model that does not consider externalities.

[2]It's not quite this simple. The actual expression is $\frac{dP}{dr_i} = (\theta_i r_i^2 + \sum_{j \neq i} r_j (2\theta_i r_i - \theta_j r_j^2))/R^2$.

sometimes winning higher slots than in the standard model.

3.1 The GSP Mechanism

Theorem 2. *In the externality model for keyword auctions, under the GSP mechanism, there is a pure Nash equilibrium.*

PROOF. Our proof is by reduction to the proofs found in [18, 10]. First sort all players by $v_i r_i^2$ and label the highest $k-1$ of them as $1, \dots k-1$. Let $\bar{R} = r_1 + \cdots + r_{k-1}$. Sort the remaining $n-(k-1)$ players by decreasing value of $v_i r_i^2/(\bar{R}+r_i)$ and label them as $k, \dots, n$.

Define $w_i = v_i r_i^2/(\bar{R}+r_k)$ if $i \le k$ and $w_i = v_i r_i^2/(\bar{R}+r_i)$ otherwise. Note that $w_1 \ge w_2 \ge \dots \ge w_k$ and that for $i \ge k+1$, $w_i = v_i r_i^2/(\bar{R}+r_i) \le w_k$. Let $R_S = r_1 + \cdots + r_k$. We have

$$w_i = \begin{cases} v_i r_i^2/R_S & \text{if } 1 \le i \le k, \\ v_i r_i^2/(R_S - r_k + r_i) & \text{if } k+1 \le i \le n. \end{cases}$$

By definition of our ordering, the w_i's are monotone decreasing.

For our reduction, consider the standard valuation model for k slots with player values $w_1 \ge w_2 \ge \dots \ge w_n$ under GSP. By the results of [18, 10] there exist bids c_i such that the allocation which gives slot i to player i is a GSP equilibrium and matches the VCG allocation (defined later). Since the VCG allocation orders players by their values, the c_i's must be monotone decreasing. The expected utility of player i is $\theta_i(w_i - c_{i+1})$.

By the equilibrium conditions, we have:

$$\begin{array}{rll} \theta_i(w_i - c_{i+1}) & \ge \theta_j(w_i - c_{j+1}) & \forall i < j \le k \\ \theta_i(w_i - c_{i+1}) & \ge \theta_j(w_i - c_j) & \forall j < i \le k \\ \theta_i(w_i - c_{i+1}) & \ge 0 & \forall i \le k \\ 0 & \ge \theta_j(w_{k+1} - c_j) & \forall j \le k \end{array}$$

Going back to our valuation model, let player i bid $b_i = c_i/r_i$. The GSP mechanism sorts the players by decreasing value of $b_i r_i$. In other words, by decreasing value of c_i, so we obtain the allocation where player i is in slot i. The price per click assigned to this player is going to be $b_{i+1}r_{i+1}/r_i$, and the expected utility of player i is

$$\theta_i r_i \left(\frac{r_i v_i}{R} - \frac{b_{i+1} r_{i+1}}{r_i} \right) = \theta_i(w_i - b_{i+1}r_{i+1}) = \theta_i(w_i - c_{i+1}).$$

By definition of b_i, the equilibrium constraints of our model reduce to the equilibrium constraints of the Varian model [18], therefore this is an equilibrium in our extended model. □

Our direct reduction to the regular GSP mechanism allows us not only to prove the existence of Nash equilibria in our setting but also provides a direct way to find a range of equilibria, as any equilibrium of GSP under the standard model can be converted to an equilibrium of our model. Among the most interesting of these equilibria is the *envy-free* equilibrium.

The envy free equilibrium of a particular auction mechanism must satisfy all the constraints of being an equilibrium for that mechanism. Additionally, it must be the case that no player who is assigned a slot prefers to trade his current slot for a new slot given the current prices. Thus the envy-free constraint provides a natural fairness criterion. However, the precise definition of what an envy free equilibrium entails depends on the players' valuation model. Under the standard valuation model, where the conversion rate for advertisers is assumed to be independent of the other sponsored ads which are displayed, [18, 10] showed that the GSP mechanism has a unique envy-free equilibrium which corresponds to the VCG equilibrium and can be obtained by setting the players' bids according to a specific recursive rule.

Under our valuation model the players not only care about the price of the slot they are allocated, but also about the winning set of players who are allocated slots. Thus, one natural way to define a GSP envy-free equilibrium in our model is the following:

Definition 3. *Given an allocation such that slot i is assigned to player i at price p_i, the allocation is* envy-free *if it satisfies the following conditions:*

- *Players who are assigned a slot have non-negative utility: For $1 \le i \le k$ and $1 \le j \le k$,*
$$\theta_i r_i \left(\frac{v_i r_i}{R_S} - p_i \right) \ge 0$$
- *No losing player i prefers to enter the set of winning players: For $k+1 \le i \le n$ and $1 \le j \le k$,*
$$\theta_j r_i \left(\frac{v_i r_i}{R_S - r_k + r_i} - p_j \right) \le 0$$
- *For $1 \le i \le k$ and $1 \le j \le k$, the player in slot i does not prefer slot j at current prices:*
$$\theta_i r_i \left(\frac{v_i r_i}{R_S} - p_i \right) \ge \theta_j r_i \left(\frac{v_i r_i}{R_S} - p_j \right).$$

The reduction we provided in the proof of Theorem 2 above allows us to use the bidding rule of [18, 10] to directly compute an envy-free equilibrium in our setting.

Theorem 4. *In the externality model for keyword auctions, under the GSP mechanism, there is a pure Nash equilibrium which is envy-free.*

PROOF. Define the ordering of the players as in the proof of Theorem 2, and set the losing players' $k+1, \dots, n$ bids as $b_i = v_i r_i^2/(R_S - r_k + r_i)$. Observe that this corresponds to a bid of $c_i = b_i r_i = w_i$ per the definitions in Theorem 2. We can now set the bids of the winning players by recursively setting them from player k up to player 2. The bidding rule that derives the envy-free equilibrium can be expressed in the original GSP mechanism by $\theta_i(w_i - c_{i+1}) = \theta_{i-1}(w_i - c_i)$ or equivalently

$$c_i = \left(1 - \frac{\theta_i}{\theta_{i-1}}\right) w_i + \frac{\theta_i}{\theta_{i-1}} c_{i+1}.$$

Using this rule we can recursively obtain the following bids for the winning players $2, \dots, k$.

$$b_i = \left(1 - \frac{\theta_i}{\theta_{i-1}}\right) \frac{v_i r_i^2}{R_S} + \frac{\theta_i}{\theta_{i-1}} b_{i+1} r_{i+1}.$$

Finally, player 1 can bid arbitrarily high to ensure the necessary conditions. It has been shown in [18, 10] that the resulting bids and prices of this bidding rule define the envy-free equilibrium. But the equilibrium constraints are the same in our model, therefore we get an envy-free equilibrium under our model. □

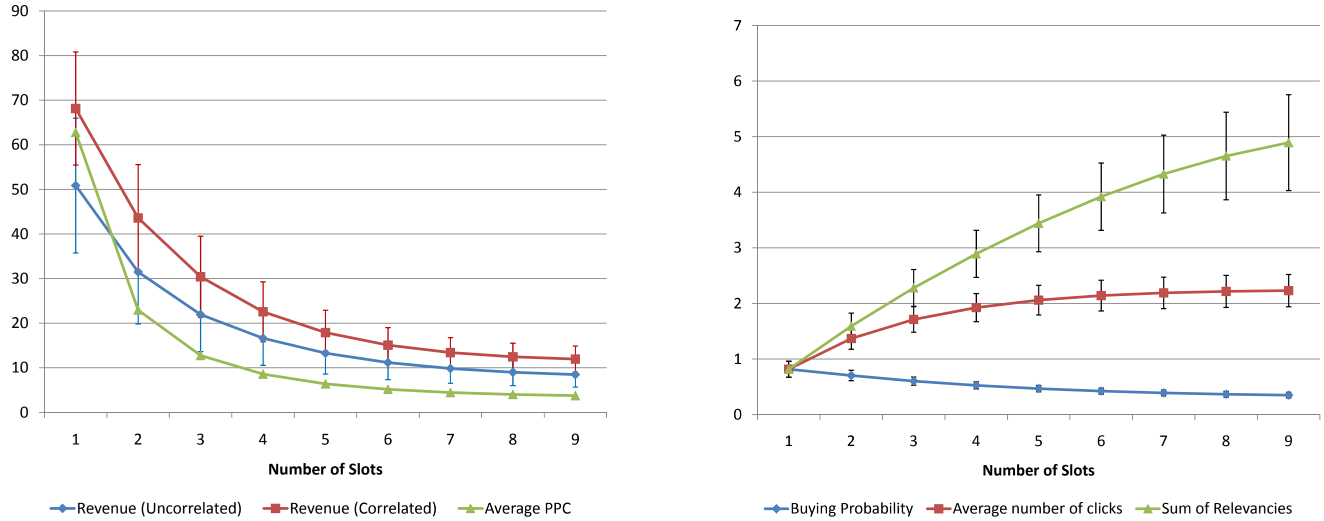

Figure 2: Varying the number of available slots. 10 players, 1-9 slots, Values: Uniform distribution [1-100], Relevances: Uniform distribution (0,1).

3.2 Efficiency and the VCG mechanism

A fundamental property of a mechanism is the *efficiency* it achieves in equilibrium. As usual, the efficiency of an allocation π is defined as the total value it gives to the players who are allocated a slot which, in the externalities model, is:

$$\text{eff}(\pi) = \frac{\sum_{1\leq i\leq k} \theta_i v_{\pi(i)} r^2_{\pi(i)}}{\sum_{1\leq j\leq k} r_{\pi(i)}} \tag{2}$$

The Vickrey-Clarke-Groves or VCG mechanism [19, 9, 14] is the mechanism which chooses the allocation so as to maximize efficiency. With VCG, it is in the best interest of participating advertisers to bid their true valuation so the input consists of the true values of the n players. (We assume that the relevances are public information.) VCG then charges each player an amount equal to how much his presence affects the maximum efficiency achievable in the system. Specifically, let π_{-i} be the maximum efficiency allocation when player i does not participate in the auction. The VCG price for slot s, allocated to player $i = \pi(s)$, is

$$p(s) = \text{eff}(\pi_{-i}) - \text{eff}(\pi) + \frac{\theta_s v_i r_i^2}{\sum_{1\leq j\leq k} r_{\pi(j)}}. \tag{3}$$

Since GSP does not maximize efficiency, a basic question one can ask is how much worse can it be. Unfortunately, it turns out that GSP has equilibrium allocations with arbitrarily worse efficiency than VCG. This result is perhaps not surprising given the fact that there exist VCG allocations which are not achievable as an equilibrium in GSP. (See Claim 8 in the appendix.)

Claim 5. *GSP equilibria can have efficiency arbitrarily lower than that of the efficiency-maximizing allocation.*

Proof. We show this for a simple example with 2 slots and 3 players. Consider the players below,

Players	v_i	r_i	$v_i r_i^2$
A	$1000/x^2$	x	1000
B	999	1	999
C	998	1	998

The envy free GSP equilibrium (defined above) always picks allocation AB: player A is picked for slot 1 as A maximizes $v_i r_i^2$, player B is picked for slot 2 as $v_b r_b^2/(r_a + r_b) > v_c r_c^2/(r_a + r_c)$. When $x > 1$ (sufficiently), then the efficiency-maximizing allocation will be BC. Thus we have

$$\frac{\text{eff(OPT)}}{\text{eff(GSP)}} = \frac{\text{eff}(BC)}{\text{eff}(AB)} = \frac{(999\theta_1 + 998\theta_2)/2}{(1000\theta_1 + 999\theta_2)/(x+1)} \approx \frac{x}{2}$$

Setting x arbitrarily proves the claim. □

In the experiments described in Section 4 we see that for random instances, the efficiency of the GSP equilibrium is usually more than 75% of the VCG efficiency.

We now delve into VCG under the externalities model in a bit more detail. We begin by studying the implementability of VCG. Note that our model is not a standard unit-demand combinatorial auction, since the players' values depend on the set of winners. Nonetheless, the following theorem indicates that it is possible to implement VCG efficiently even in our model.

Theorem 6. *Assume that all input parameters are integers. Then the VCG mechanism can be run in polynomial time.*

Proof. See Appendix. □

Finally, we observe that VCG in this setting has some undesirable properties.

Claim 7. *VCG has the following three unappealing properties:*

1. *The VCG prices may be negative.*

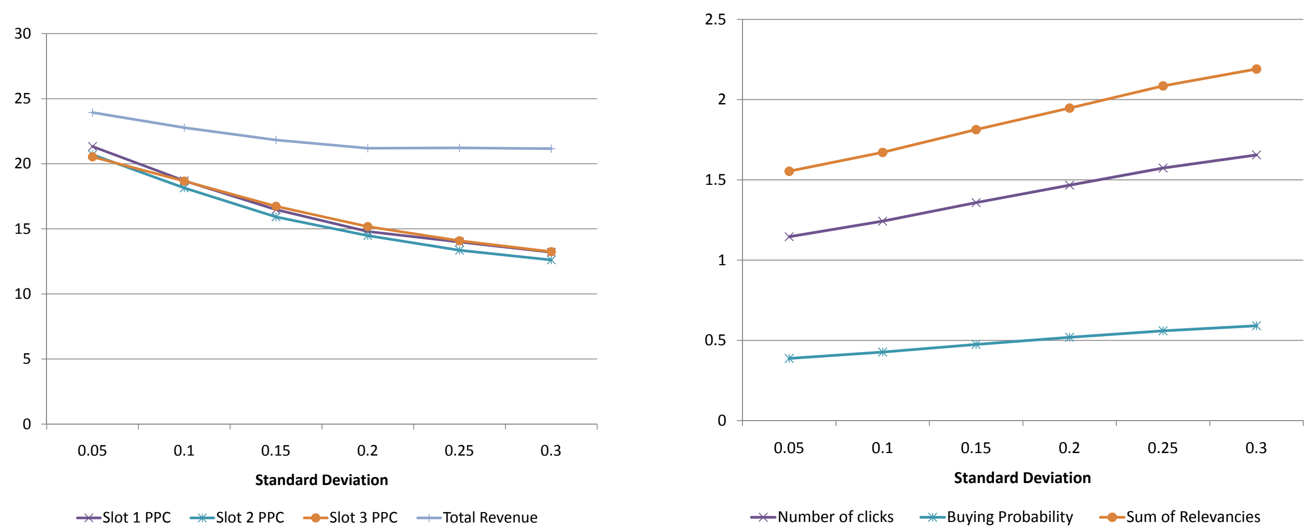

Figure 3: Varying the standard deviation of the relevances distribution. 10 players, 3 slots, Values: Uniform distribution [1-100], Relevances: Normal distribution with mean 0.5 and standard deviation 0.05-0.30.

2. *The output of VCG may not be envy free.*
3. *When the number of slots is unconstrained, VCG allocates only one slot.*

PROOF. See Appendix. □

The first part of the claim shows that under the VCG mechanism, it is possible to charge negative prices for slots. Thus under some circumstances the search engine would be required to pay advertisers to participate in the auction. Note that this relies on the fact that we only consider allocations that allocate all k slots, and do not consider allocations of less than k slots to be feasible. In contrast, the GSP mechanism never charges negative prices to any advertiser.

4. EMPIRICAL ANALYSIS

In this section we empirically study how competition among advertisers affects the various participants in the auction.

We vary the competition within a keyword auction in one of two ways: by varying the number of available advertisement slots, and by varying the relevances of advertisers competing in the auction. In practice, the number of slots is chosen by the search engine and usually differs per keyword without the search engine revealing its decision process. The relevances of the competitive advertisements also depend on the keyword and which search engine the auction is run on, but it is not clear if and how these relevances are related to the advertisers' values and bids.

Under various settings of competition, we look at the following metrics:

- *search engine revenue*, which can be broken down into click probabilities and prices per click;
- *user satisfaction*, as measured by both buying (conversion) probability and the sum of relevances of displayed ads (a measure of the quality of the set of results shown); and
- *efficiency*, where we compare GSP and VCG.

We also consider the impact of the recently implemented change in which search engines sometimes divide sponsored search results into two sets: those shown at the top, which are required to have a minimum quality or relevance threshold and those on the side, which are chosen according to strict rank by revenue without concern for quality. The motivation for this change was to ensure that the most prominently displayed ads would be of high quality and thus be satisfying to users. We study experimentally how such a decision affects the properties of our system by varying the minimum relevance required to win a slot.

4.1 Experimental Setup

In all experiments, we define the keyword auction as follows: we use three slots and ten players. Feng et al.[11] state that the click-through rates are typically well fitted to the geometrically decreasing sequence $\theta_i = (.7)^{i-1}$, and this is the sequence we use. Each plotted point represent the average of 1000 instances where for each instance, the values of the players are chosen independently and uniformly from the range $[0, 100]$. We always take the outcome of the GSP keyword auction to be the envy-free equilibrium defined previously.

In experiments where we analyze a particular quantity (such as search engine revenue) with respect to the number of slots, the relevances of the players are chosen independently and uniformly from the range $[0, 1]$ and the number of slots varies from 1 to 9. We have two scenarios for creating competition in our experiments by varying the relevance of players. In the first scenario the relevances are chosen independently from a normal distribution with mean .5 and standard deviation σ. We analyze the metrics described above by varying σ from .05 to .3 in .05 increments. In the second scenario we create a homogeneous set of players by choosing every player's relevance to have correlation c to

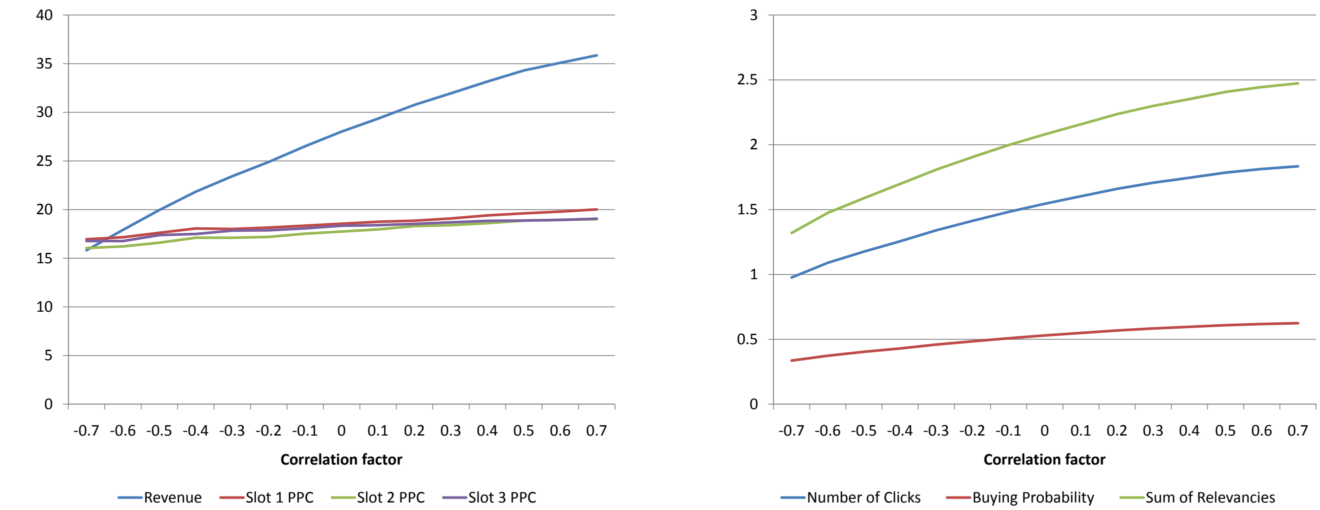

Figure 4: Varying the correlation factor. 10 players, 3 slots, Values: Uniform distribution [1-100], Relevances: Drawn randomly with correlation factor -0.7 to 0.7.

that player's value[3]. We analyze quantities of interest while varying the correlation c from -0.7 to 0.7 in .1 increments.

4.2 Slot Availability

A critical choice that significantly affects the auctions for a given keyword is the number of available advertisement slots. In Figure 2, we show the various metrics as the number of slots is varied from 1 to 9.

As expected, this figure shows that as the number of slots increases, the total quality displayed to the user (sum of relevances) increases and the average number of received clicks increases. This would seem to suggest that it is in the best interest of the search engine to display many highly relevant ads. However, increasing the number of displayed ads also increases the competitiveness of the market and could lead to lower utility for advertisers and hence cause their bids to go down. And indeed this is exactly what we see. As the number of slots increases, slot prices drop and the total revenue for the search engine *decreases* with more slots.

On the left side of Figure 2, we consider search engine revenue under two different scenarios: when the value of a conversion is positively correlated with the relevance of the advertiser and when it is uncorrelated. Not surprisingly, search engine revenue is higher when these values are positively correlated, but in both cases, the effects of competition result in decreasing revenue with increasing numbers of slots.

Focusing on the uncorrelated case[4], the figure shows that the buying (conversion) probability of the end-user also decreases with more slots. This behavior is consistent with customer behavioral studies (e.g., [17]) which claim that customers become confused when presented with too many appealing or similar choices, and thus have a harder time finalizing their decisions. However when a few sponsored results look noticeably more relevant than the rest, or only a few sponsored results are shown, the user is more likely to make a purchase.

Putting these observations together, we see that there is a tradeoff when it comes to the number of slots to allocate. User happiness is likely to be some combination of total relevance seen and buying probability. Keeping users happy keeps them coming back which ultimately increases search engine revenue. Thus, even if our model reflects reality and it is the case that displaying multiple ads lowers revenue in a single keyword auction, it is unclear how revenue is affected over time. Resolving this question is an interesting direction for future research.

4.3 Distribution of Relevances

We next examine how a varying distribution of relevances affects the participants of the system for a fixed number of slots. The results are shown in Figure 3. Not surprisingly, as the variance of the relevances increases, we find that the sum of the relevances presented to the user and the average number of received clicks increase. More interesting, but not unexpected, is the fact that the buying probability increases. Again, this corresponds to the observation that having roughly equal choices (low variance in relevances) makes it more difficult for the user to make a decision to buy.

On the other hand, we see that increasing the variance of relevances causes a decrease in the slot prices, enough to make the total revenue remain roughly at the same level. Thus, if the results in Figure 3 are representative of reality, higher variance of the relevances' distribution does not affect the search engine's revenue significantly, but increases user satisfaction. This could have implications for how the search engine should choose which advertisers are allowed to participate in each auction.

4.4 Value-Relevance Correlation

It is unknown to us whether or not in practice there is any correlation, positive or negative, between the values of advertisers and their relevances. In order to understand the

[3] c refers to the Pearson correlation factor.

[4] The results for the correlated case are very similar.

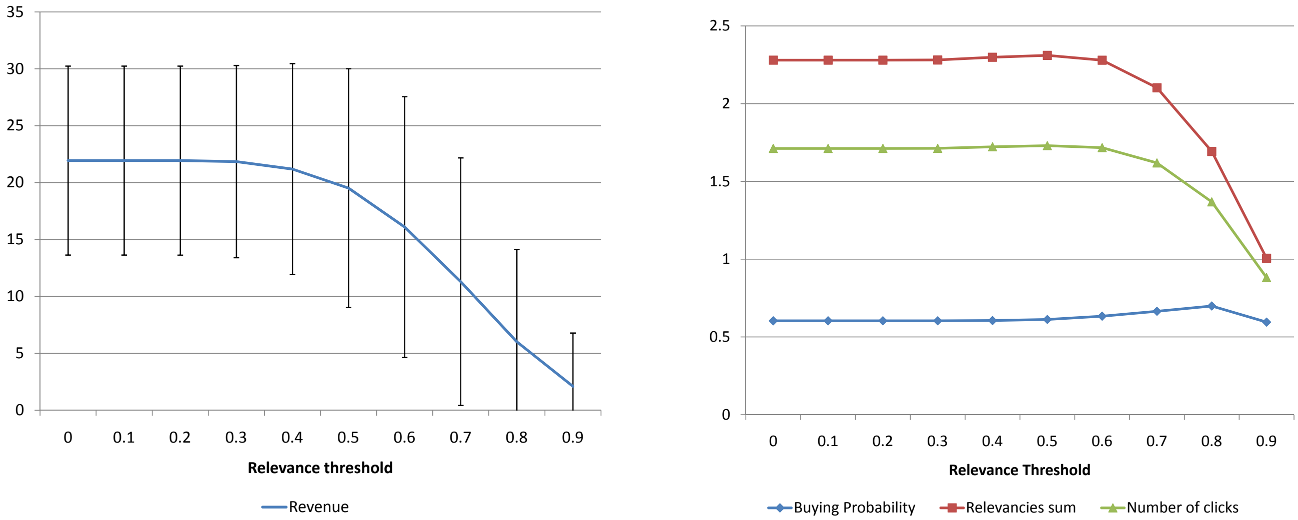

Figure 5: Using a minimum relevance factor. 10 players, 3 slots, Values: Uniform distribution [1-100], Relevances: Uniform distribution (0,1).

effects of any such correlations, should they exist, we study our metrics as the correlation varies from negative to positive. The results are shown in Figure 4.

The results in Figure 4 suggest that the slot prices per click remain fairly unaffected by correlation. (Notice that these pay per click prices might not be monotone in the order of slots; however the effective payments will be, as they are multiplied by the relevance of the players occupying the respective slots.) We expect that positive correlation results in advertisements allocated to higher slots having both higher values and higher relevances. Consequently, as correlation increases, the number of received clicks increases, resulting in higher search engine revenue.

Overall, it seems that a positive correlation between the values and relevances is quite critical in the system's performance and a study to reveal what kind of correlation is exhibited in practice could be a key step in optimizing such systems.

4.5 Relevance Threshold

We next explore the impact of the recent introduction of a minimum relevance threshold for a "prime" set of slots. In these experiments, we studied the various metrics as the minimum threshold of relevance to be considered for the auction varies from 0 (no threshold) up to around the maximum relevance value. The results are shown in Figure 5.

What we see here is that, for our choice of parameters, none of the metrics change much until the threshold is around 0.4. This is mostly attributable to the fact that the winners of the auction often don't change. Starting at a minimum relevance threshold of 0.4, however, the search engine starts trading off revenue for user satisfaction, because at this point, and until a threshold of about 0.6, both buying probability and the sum of relevances displayed is essentially unchanged. Beyond a threshold of 0.6, the sum of relevances starts to drop because there simply aren't enough advertisers of high relevance. Clearly, choosing the relevance threshold is an important optimization question for the search engine.

4.6 Market Efficiency

Finally, we compare efficiency of the GSP envy-free equilibrium to the efficiency of VCG, when both are required to allocate 3 slots.

Figure 6 plots the efficiency of VCG and GSP as a function of the correlation between players' values and relevances. The first thing we see is that the bad example given in Claim 5 showing that the efficiency of the envy free equilibrium could be very poor compared to the VCG efficiency is a rarity among random instances. In fact the efficiency of the GSP equilibrium is nearly comparable to the efficiency of VCG. The GSP efficiency is most competitive with the VCG efficiency when there is positive correlation between players' values and relevances ($c > .3$), and slightly less competitive when there is a negative correlation. (Note that in our bad example from Claim 5 the values and relevances are negatively correlated.)

The efficiency of the GSP equilibrium is less robust when the advertisers have more diverse relevances. Figure 6 plots the efficiency of VCG and GSP for this setting. Generally as the relevances of players get more varied, the efficiency of GSP gets farther from the VCG efficiency.

5. DISCUSSION

We have presented a model that incorporates the effects of competing advertisements in sponsored search auctions. We developed this model by doing a thought experiment exploring how we thought a user interested in making a purchase might behave when presented with a set of advertisements. Indeed, the model we propose has many of the qualitative properties we expect in this setting.

On the other hand, our model may not reflect reality in some ways. First, there is the fact that search engine revenue, efficiency and the probability that the user converts all decrease with increasing numbers of slots. We do not know if this is reasonable or is a property that is seen in practice. Second, there appears to be empirical evidence that conver-

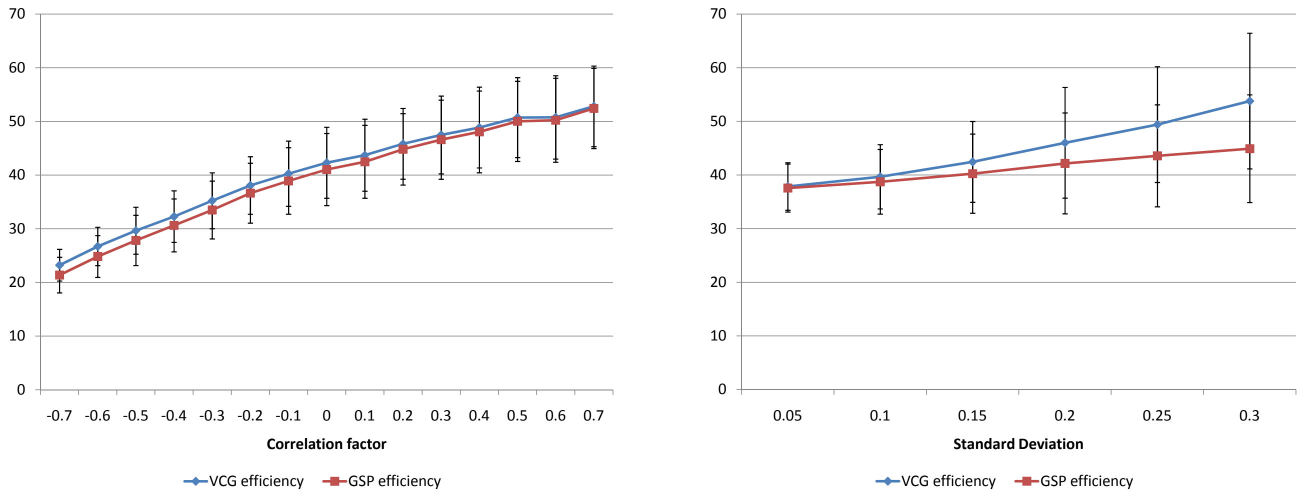

Figure 6: Comparing efficiency of GSP and VCG varying standard deviation of relevances distribution and correlation factor.

sion rates are higher for lower slots than for higher slots [6]. This will not happen in our model unless, for some reason, relevances and values are inversely correlated in such a way that relevances tended to be *higher* for lower slots. Third, there is some anecdotal evidence [7] that a user is more likely to buy something from an ad he more recently clicked on. Finally, our model does not reflect the fact that some users are purely browsers and have no intention of buying, but still generate clicks.

There are of course many other possible models that would be interesting to explore to help deal with these issues. For example, here are two possibilities:

- Consider the following simple variation on the model described in this paper: The click probabilities and the overall probability of buying are the same as what we presented here. However, if the user clicks on slots $i_1, i_2, \ldots, i_s$, in that order, and then decides to buy, then the probability that the purchase is made from slot i_s is proportional to r_{i_s}, the probability that the purchase is made from slot i_{s-1} is proportional to $\frac{r_{i_{s-1}}}{2}$, the probability that the purchase is made from slot i_{s-2} is proportional to $r_{i_{s-2}}/3$, and so on. In other words, while the probability of making a purchase from a particular advertiser is related to their quality, it also diminishes the farther in the past the click on the associated advertisement occurred.

- Another possibility is to consider two types of users: browsers and buyers. For each user, there is some probability α that that user is a browser, and corresponding probability $1 - \alpha$ that the user is a buyer. Associated with browsers is a sequence of clickthrough rates $\theta'_1 > \theta'_2 > \ldots > \theta'_k$, such that the probability that a browser clicks on the ad in slot i with relevance r_i is $\theta'_i r_i$. However, a browser will never makes a purchase. Associated with buyers is a different sequence of clickthrough rates $\theta''_1 > \theta''_2 > \ldots > \theta''_k$, such that the probability that the buyer clicks on the ad in slot i with relevance r_i is $\theta''_i r_i$. Once the potential buyer finishes clicking, they choose whether to convert, and if so, from which advertiser to buy, according to the model in this paper.

  With this model, the probability that a user converts on player i's ad in slot i given that he clicks on it is
  $$(1-\alpha)\theta''_i \frac{r_i}{R_S} / (\alpha\theta'_i + (1-\alpha)\theta''_i).$$
  In addition to capturing both types of users, this model has the property that if the θ''_i sequence decays faster with i than the θ'_i sequence (and it seems likely that buyers would be more prone to clicking than browsers), then the factor multiplying $\frac{r_i}{R_S}$ in the above expression would *increase* with i. This corresponds to the observed property that conversion rates may increase with slot number.

6. CONCLUSIONS

In this paper, we have taken first steps towards understanding the effects of competing advertisements in sponsored search auctions. We have studied this model using a combination of theoretical analysis and empirical evaluation and have shown that it has many of the qualitative properties we expect.

Clearly, this is not the only reasonable model – we have outlined some alternatives in the previous section. We plan to explore these other models in the near future.

7. REFERENCES

[1] B2b survey summary, November 2004. Prepared by Hotchkiss, G.

[2] Inside the mind of the searcher, November 2004. Prepared by Hotchkiss, G.

[3] Aggarwal, G., Goel, A., and Motwani, R. Truthful auctions for pricing search keywords. In *EC '06: Proceedings of the 7th ACM conference on Electronic commerce* (New York, NY, USA, 2006), ACM, pp. 1–7.

[4] Animesh, A., V. S., and Agarwal, R. t.
[5] Arbatskaya, M. Ordered search. 2004.
[6] Blumrosen, L., Hartline, J., and Nong, S. Conversion rates in sponsored search auctions. Preliminary version in *Dagstuhl workshop on Computational Social Systems and the Internet*, 2007.
[7] Broder, A. Private Communication, 2008.
[8] Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A., Mathieu, C., and Schwarz, M. Greedy bidding strategies for keyword auctions. In *EC '07: Proceedings of the 8th ACM conference on Electronic commerce* (New York, NY, USA, 2007), ACM, pp. 262–271.
[9] Clarke, E. H. Multipart pricing of public goods. *Public Choice 11* (1971), 17–33.
[10] Edelman, B., Ostrovsky, M., and Schwarz, M. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. To appear in *American Economic Review*, 2006.
[11] Feng, J., Bhargava, H. K., and D.M., P. Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *INFORMS Journal on Computing* (2006). Forthcoming.
[12] Ghosh, A., and Mahdian, M. Externalities in online advertising. In *Proceedings of the 17th InternationalWorld Wide Web Conference* (2008).
[13] Granka, L. Eye-r: Eye tracking analysis of user behavior in online search. Masters Thesis, Cornell University Library Press, 2004.
[14] Groves, T. Incentives in teams. *Econometrica 41* (1973), 617–631.
[15] Lahaie, S. An analysis of alternative slot auction designs for sponsored search. In *EC '06: Proceedings of the 7th ACM conference on Electronic commerce* (New York, NY, USA, 2006), ACM Press, pp. 218–227.
[16] Prepared by Hotchkiss, G. Eye tracking report: Google, msn and yahoo! compared, November 2006.
[17] Schwartz, B. The paradox of choice: Why more is less. HarperCollins Publishers.
[18] Varian, H. Position auctions. To appear in *International Journal of Industrial Organization*, 2006.
[19] Vickery, W. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance* (1961), 8–37.
[20] Zhang, X. Finding edgeworth cycles in online advertising auctions. MIT Sloan School of Management, Working Paper, 2005.

APPENDIX

Claim 8. *There exist a VCG allocation which cannot be realized in any GSP equilibrium.*

Proof. We will give an example for 2 slots and 3 bidders. Suppose the VCG allocation is to allocate the bidders in order ABC. As $\theta_1 > \theta_2$ this implies that $v_a r_a^2 > v_b r_b^2$ as otherwise VCG would reverse AB in the allocation.

What conditions must be satisfied for this allocation to be a GSP equilibrium? WLOG we can assume that player A will get the top slot by bidding infinity so that players B and C will not target slot 1. Let $R = r_a + r_b + r_c$, and let the prices of slot 1 and 2 be $p_1 = b_b r_b$, $p_2 = b_c r_c$. In order for this allocation to be an equilibrium in GSP, it is necessary and sufficient to satisfy the following constraints:

$$\begin{aligned} p_2 &\leq p_1 \\ p_1 &\leq v_a r_a^2/(R-r_c) \\ p_2 &\leq v_b r_b^2/(R-r_c) \\ p_1 - (\theta_2/\theta_1)p_2 &\leq (1-(\theta_2/\theta_1))v_a r_a^2/(R-r_c) \\ v_c r_c^2/(R-r_b) &\leq p_1 \end{aligned}$$

The second inequality follows from the third and fourth inequalities (as $v_b r_b^2 \leq v_a r_a^2$) and can be eliminated. Observe that p_2 is upper bounded in two equations and lower bounded on only one equation, so for any feasible solution (p_1, p_2) there exists a feasible solution (p_1, p_2') such that the lower bound is tight. In other words, without loss of generality we can assume that

$$p_2 = (\theta_1/\theta_2)p_1 - (\theta_1/\theta_2 - 1)v_a r_a^2/(R-r_c)$$

Substituting in for p_2, the constraints above are equivalent to the following equation which must be satisfied to get allocation ABC as a GSP equilibrium:

$$\frac{v_c r_c^2}{R-r_b} \leq \frac{\theta_2}{\theta_1}\frac{v_b r_b^2}{R-r_c} + \left(1-\frac{\theta_2}{\theta_1}\right)\frac{v_a r_a^2}{R-r_c}. \tag{4}$$

Any VCG allocation (for 2 slots, 3 bidders) which violates equation 4 is not achievable as a GSP equilibrium. The example given in Claim 7, part (2), the VCG allocation is ABC. For click-through-rates $\theta_1 = 1$ and $\theta_2 = .92$) the right hand side of equation 4 is $(.95(39) + .05(48))/3 = 13.15$ and the left hand side is $80/6 = 13.333$. Thus Equation 4 is violated implying that the ordering A, B, C cannot be an equilibrium in GSP. $\square$

A. PROOF OF RESULTS OF SECTION 3.2

Proof of Theorem 6. For any set of values $v_1, \ldots v_n$ and relevances $r_1, \ldots, r_n$, we will prove that the allocation of maximum efficiency assigning k slots can be found in polynomial time. Thus, for any given set of players, the allocation of maximum efficiency can be found in polynomial time. This implies the Theorem, as the VCG prices can be calculated from the maximum efficiency allocations for two sets of players per slot.

First consider the decision version of the problem: Does there exist an allocation σ with efficiency greater than equal to W? Going back to Equation (2), this is equivalent to asking whether there exists an allocation σ such that

$$\sum_{1\leq i\leq k} \theta_i r_{\sigma(i)}^2 v_{\sigma(i)} - W r_{\sigma(i)} \geq 0 \tag{5}$$

To answer this question, define a complete bipartite graph G with n vertices on the right side (representing the players), k vertices on the left side (the slots), such that the edge between player j and slot s has weight $\theta_s r_j^2 v_j - W r_j + B$. Here $B = W \max_i r_i$ is a value to ensure that all weights are positive. Compute a maximum weight matching on G. Since all edge weights are positive, the maximum weight matching of G has k edges, hence corresponds to an allocation of the k slots; and it has weight at least kB iff there exists an allocation σ which assigns k slots and has efficiency at least W. As maximum weight matching in G can be found in

$O(n^3 + k^2)$ steps, we can answer the decision version of the problem in poly-time.

The optimization problem can be solved by doing binary search along a bounded interval of possible efficiency values, using the decision version algorithm to check for the existence of allocations with particular efficiencies. The inequality $(a_1+\ldots+a_n)/(b_1+\ldots+b_n) \le \max a_i/b_i$ can be used on Equation (2) to derive an upper bound for the efficiency of any allocation of $\max_{1\le j\le n} v_j r_j$. Thus the initial binary search interval is $[0, \max(v_j r_j)]$. Now, in order to bound the number of steps of the binary search, we compute a lower bound on the difference in efficiencies of two allocations π and σ. Each input parameter is an integer. If π and σ have distinct efficiencies, then

$$|eff(\pi) - eff(\sigma)| \ge \frac{1}{(\sum_{1\le j\le k} r_{\pi(j)})(\sum_{1\le j\le k} r_{\sigma(j)})} \ge \frac{1}{R^2},$$

where R denotes the sum of k highest relevances in the input. Thus we can stop the binary search when the interval is of size $1/R^2$, and return the last allocation where the decision version algorithm returned a positive answer. Going from an interval of size $\max v_j r_j$ to one of size $1/R^2$ using binary search can be done by solving $O(\log(R^2 \max v_i r_i))$ decision problems, giving a polynomial time algorithm for finding the allocation of maximum efficiency.

Proof of Claim 7, part (1). We demonstrate this using a simple example with 2 slots having click through rates $\theta_1 = 1$ and $\theta_2 = 1/2$. Consider the following players:

	Value	Relevance
Player A	1	1
Player B	1-ε	1
Player C	1/8	2

It is easy to check that allocation AB has the highest efficiency of $(\theta_1 v_a r_a^2 + \theta_2 v_b r_b^2)/(r_a + r_b) \sim 3/4$. Now let us examine the price of the second slot p_2. If player B was not participating in the auction, the most efficient allocation of the remaining players would be AC, with efficiency approximately $\frac{5}{12}$. Finally the value B obtains from allocation ABC is $\theta_2 v_2 r_2^2/(r_1 + r_2) \sim 1/4$. Therefore using the vcg pricing equation 3, the total price for slot 2 is: $p_2 \simeq \frac{5}{12} - \frac{3}{4} + \frac{1}{4} = -\frac{1}{12}$.

It is not hard to generalize this example to larger instances where the VCG prices are negative. Since VCG is forced to fill k slots, this situation may occur when a low relevance high valued bidder (such as B) is part of the VCG allocation and his next best replacement is a bidder of with much higher relevance.

Proof of Claim 7, part (2). Once again we demonstrate this with a simple 2 slots example, with click through rates $\theta_1 = 1$ $\theta_2 = .95$. Consider the following set of bidders [5]

Player	Value	Relevance
A	12	2
B	39	1
C	5	4

Examining all possible allocations it is easy to check that the allocation AB has efficiency of $(\theta_1 v_a r_a^2 + \theta_2 v_b r_b^2)/(r_a + r_b) \sim 3/4$ which is the highest. Now let us calculate the VCG prices for slot 1 and 2:

$$\begin{aligned} p_1 &= eff(CB) - eff(AB) + v_a r_a^2/(r_a + r_b) = 11.06 \\ p_2 &= eff(CA) - eff(AB) + \theta_2 v_b r_b^2/(r_a + r_b) = 4.90 \end{aligned}$$

This is not envy free as A would rather have slot 2 at price \$4.90 then his current slot at price \$11.06. Indeed, A's utility at his current slot is $\theta_1(v_a r_a^2/(R-r_c)-p_1) = 4.94$ whereas his utility in slot 2 at price p_2 would be $\theta_2(v_a r_a^2/(R-r_c)-p_b) = 10.54$.

Proof of Claim 7, part (3).

Suppose not and let σ a more efficient allocation containing $k \ge 2$ slots and let $S = \sum_{i=1}^{k} r_{\sigma(i)}$. The efficiency of σ is $eff(\sigma) = (\theta_1 v_{\sigma(1)} r_{\sigma(1)}^2 + \ldots + \theta_k v_{\sigma(k)} r_{\sigma(k)}^2)/S$.

Let m be the player who maximizes $v_i r_i$ so that $v_m r_m \ge v_{\sigma(i)} r_{\sigma(i)}$. Using this and the fact that $\theta_1 > \theta_s$ for all slots $s > 1$, we get

$$eff(\sigma) < \theta_1 v_m r_m (r_{\sigma(1)} + \ldots + r_{\sigma(k)})/S = \theta_1 v_m r_m$$

The right hand side of the inequality is the efficiency of allocation the top slot is allocated to player m.

[5]We can always scale all relevance in this example to be ≤ 1 and the same results would hold.