Optimal Revenue Maximizing Mechanisms in Common-Value Position Auctions*

Working Paper

Maryam Farboodi†, Amin Jafarian‡

January 2014

Abstract

We study the optimal mechanism in position auctions in a common-value setting where only ordinal information about the advertisers is posted and auctioneer revenue depends on consumer belief about ad qualities. We show that when bidder valuations are correlated, existence of a simple non-ironed optimal mechanism requires a stronger condition that the usual increasing virtual valuation assumption. With fixed number of positions, the corresponding optimal decision rule (among all mixed and pure strategies) is to allocate the ad positions to advertisers in decreasing order of quality. More importantly, when the search engine also chooses the number of ads to post, ignoring the consumer belief endogeneity will lead to posting too many ads compared to what maximizes the search engine revenue. We also show that when only one ad position is available, the optimal allocation rule uses a reserve price which is *higher* than what is implied by Myerson 1981 [8] seminal work. Finally, we characterize the optimal mechanism when cardinal information about advertiser quality is posted, and we provide results on how search engine revenue compares under the two regimes.

1 Introduction

Since the huge surge in online search in the early years of 21$^{\text{st}}$ century search engines has been increasingly relying on online advertisement as one of their main sources of revenue. Online advertisements are allocated using "sponsored search" auctions, which are billions of simultaneous auctions run by search engines to allocates ad slots displayed alongside

*We are extremely grateful to Alex Frankel, Kenneth Hendricks, Laurent Mathevet, Roger Myerson, Philip Reny and Balazs Szentes, as well as seminar participants at University of Chicago and Carnegie Mellon University for numerous comments and discussions. All errors are ours.

†Princeton University. Email: farboodi@princeton.edu

‡Princeton University. Email: jafarian@princeton.edu

organic search results. The three main search engines, Google, Yahoo! and Microsoft Bing has generated billions of dollars in revenue from online auctions in 2011. Total advertising revenues constitutes more % 96 of total revenues of Google in 2011 and 2012 ($36,531 of $37,905 in 2011, and $43,686 of $46,039 in 2012). Advertising related revenue is not only generated directly from the search engine website, but also from ad slots sold on affiliated websites, shopping web pages, blogs, etc.

These sponsored search auction are generalized second price auction (GSP) run on a pay-per-click basis: advertisers bid for keywords and phrases and once that keyword is searched on the website, an online auction is run to allocated the ad slots. Bidders are ranked by complicated functions of their current bids, reputation and other indicators and some of them are displayed on the webpage along with the search results of the searcher query. Once an ad is clicked, the bidder has to pay the search engine.

Given the pay-per-click payment scheme of the sponsored search auctions, in displaying advertisements the search engine not only cares about the bids, but also about how many clicks each ad receives (*click-through-rate*). Given that consumers (searchers) are the ones who click on the displayed ads, their belief is essential to search engine revenue. Importantly, how the search engine allocate the ads to advertisers feeds back into consumer belief about the ads and affects the search engine revenue. This feed back effect has been mostly ignored in the literature on sponsored search and click-through-rates are assumed to be exogenous (for instance [3] and [4]), and has only recently been recognized ([2] and [5]). In this paper, we would like to study the optimal revenue maximizing mechanism for allocating these sponsored ads by explicitly modeling the endogenous belief of consumers (searchers) about the ads and how it affects the search engine revenue. It turns out that endogenous consumer beliefs are consequential for search engine optimal decisions rule: ignoring the endogenous beliefs and taking them as constants leads the search engine to display too many ads compared to what would maximize his expected revenue.

To study the optimal revenue maximizing mechanism when advertisers have interdependent valuations for ad slots, we build on the model by Athey-Ellison 2011 [2]. Since we am interested in the optimal decision rule designed by the search engine, we need to be very careful about how this decision affects consumer beliefs, which in turn generate clicks on different ads. In fact this endogenous determination of beliefs is what makes this problem particularly challenging, and is the main source of difference with the optimal mechanism in a similar setting where click through rates are exogenous.

Intuitively speaking, consider a setting in which the number of clicks that an ad i receive (i.e. its click through rate) depends on which ads are believed to be better than ad i, as well as consumer belief about how good ad i is. In such setting advertiser value of each slot depends on own quality and quality of better advertisers, as well as consumer beliefs own quality. The principle's choice of mechanism affects consumer beliefs, which feeds back on the advertiser value from purchasing each slot.

In the first few sections of what follows we will assume that the search engine only reveals ordinal information about ad qualities, i.e. the search engine only conveys the relative ranking of the ads to consumers but does not tell them how good each ad is. We show that with endogenous click-through-rate *hyper-regularity* of distribution of ad qualities, a strengthened version of familiar regularity (increasing virtual valuations) condition, is a sufficient condition for existence of a non-ironed optimal mechanism. More importantly, we show that the mechanism designer displays fewer ads compared to the case where click-through-rates are exogenous. Consequently, if the search engine ignores the feed back effect of the allocation rule into consumer beliefs, too many ads will be posted compared to what is revenue maximizing. Then we will consider the case where cardinal information is displayed, and finally provide results on how the revenue of the auctioneer compares across the two settings.

As mentioned earlier, the major search engines use variations of (weighted) generalized second price auctions to sell sponsored search links. [1], [11] and [3] study equilibria of such auctions in a with *exogenous* click-through rates. In contrast, [2] characterizes equilibrium strategies in a Generalized English auctions with *endogenous* click-through-rates: searchers take clicking decisions rationally and follow a sequential search procedure. We will use ingredients from this latter paper the study the optimal revenue maximizing mechanism in sponsored search auctions.

The rest of the paper is organized as follows: Section 2 lays out the environment, section 3 studies a restricted version of the optimal mechanism with ordinal information and section 4 characterizes the full fledged optimal mechanism under certain conditions and also provide results on more general cases. Section 5 characterizes the optimal mechanism when cardinal information is revealed, and section 6 compares auctioneer revenue across the two cases. Section 7 concludes.

2 Environment

There is a continuum of consumers. Each consumer has a need and visits a website to be matched with a supplier and meet the need. The consumer receives a benefit of 1 if the need is met and zero otherwise. In addition, consumer j has a privately observed search cost s_j for clicking any sponsored link. s_j's are drawn independently from distribution $G(.)$.

There are N producers who want to advertise on the website. From now on, we will simply refer to them as advertisers. Each advertiser has a quality, which is the probability of him being able to meet the consumer need. An advertiser will get a payoff of 1 each time a need is met. Let θ_i in $[0, 1]$ denote the quality of advertiser i, drawn independently from distribution $\Phi(.)$. An interesting special case which we will frequently go back to is when both θ_i and s_j are iid draws from uniform distribution on unit interval.

The search engine posts $M \leq N$ sponsored links on the a web page, and each ad appears only once on the same page. If an ad is posted on a page, the owner (advertiser) is charged according to the rule designed by the auctioneer.[1]

Consumers, once presented with a list of sponsored links, click optimally on the links until their need is met or the expected benefit from clicking an additional ad falls bellow s_j. Most importantly, consumer expected benefit from clicking each ad depends on his belief about the quality of this ad. [2]

[1]Athey-Ellison 2011[2] assumes the auctioneer runs a generalized second price auction.

[2]In order to compute the equilibrium of the model, Athey-Ellison 2011 [2] makes two important as-

A direct revelation mechanism consists of a decision (allocation) rule and a set of expected transfers for each participant. The search engine is a mechanism designer who asks the bidders for their types (here qualities). The objects ad slots are then allocated using the allocation rule and reported types, and each bidder pays or receives the designated transfer. Without loss of generality we will focus on revenue maximizing mechanisms which are truthfully implementable in Bayesian Nash Equilibria, i.e. mechanisms in which all bidders find it optimal to report their true qualities.

Before rigorously defining the optimal mechanism, we need to introduce a few pieces of notation. Let θ and $\hat{\theta}$ denote the N-vector of true and reported types, and let θ_i, θ^j and $\theta^{[k]}$ denote the type of bidder i, j^{th}-highest bidder and bidder in ad slot (position) k, respectively. Note that θ^j is the j^{th}-order statistic of the qualities.

Let CTR_k denote the number of clicks that an ad in position k receives. Note that CTR_k depends on the consumer belief about the ad quality if the consumer gets so far, as well as the percentage of needs which are not met by the higher ads.

The consumers click on the posted ads based on their belief about the (order in) quality of posted ads: they click on the ad that they believe has the highest quality first, and they continue in descending order of qualities. So an equilibrium (as described later) where the auctioneers posts a list of ads in descending order of (reported) qualities and consumers click in a top-down manner will be an equilibrium with a consistent (self-fulfilling) belief structure.

Define $r(\theta,\hat{\theta})$ as the M-vector of needs at each position; which means the k^{th} entry is the percentage of consumer needs not met by any of the ads in position $1, 2, \cdots, k-1$. With the notation introduced above, r can be written as:

$$r(\hat{\theta},\theta) = (1, (1-\theta^{[1]}), \cdots, (1-\theta^{[1]}) \cdot\cdot (1-\theta^{[k-1]}), \cdots)$$

Let $X_{N\times M}(\hat{\theta})$ denote the allocation rule. X_i, the i^{th} row of matrix X, corresponds to bidder i with $X_{ik}=1$ if bidder i is positioned in slot k. Moreover, let $t_i(X)$ be the transfer that bidder i has to make with allocation X. Note that each bidder get at most one slot and all ad positions are allocated, so we should have the following two criteria satisfied:

$$\sum_j X_{ij} \leq 1 \qquad 1 \leq i \leq N \tag{1}$$

$$\sum_i X_{ij} = 1 \qquad 1 \leq j \leq M \tag{2}$$

Also, Let $T(X(.)) = (T_1(.), \cdots, T_N(.))$ denote the set of transfers.

sumptions. First, they assume that given a list of ads, consumers believe that ads are sorted in decreasing order of quality (and they show that this belief is consistent in equilibrium). Second, for the most part of the paper they assume that *exactly* M ads are posted on each page. These two assumption guarantee a simple form for consumer expected benefit if he considers to click on an ad. Athey-Ellison 2011 [2] finds an equilibrium in which no advertiser wants to deviate from its bid. They show that advertisers bid truthfully until they get on the list and then shade their bids. They also study the welfare maximizing reserve price when there is only one ad slot available.

Using the above definitions one can compute the click through rate for each position k. Note that when a consumer j reaches an ad, he will click it only if the expected benefit from clicking this link exceeds his search cost s_j. Recall that we am looking for an equilibrium in which when consumers see an ordered list of ads, they assume that the ads are ordered in decreasing order of quality, so they believe the expected quality of the ad in position k is lower than the expected quality of all ads in higher positions. As a result they consider clicking on the k^{th} ad only if they have already clicked on the first $k-1$ ads and their need is not met yet. Furthermore, they click on this ad only if its quality conditioned on all the information so far (i.e. the fact that none of the higher ads has met their need) is still higher than their search cost.

Following the notation introduced in Athey-Ellison 2011 [2], let $z^1, \cdots, z^M$ be Bernoulli random variables equal to one with $\theta^{[1]}, \cdots, \theta^{[M]}$ probabilities, i.e. $z^k = 1$ if the ad in position k meets the consumer need and zero otherwise. Let $\bar{\theta}^{[k]}$ denote the expected quality of the advertiser in the k^{th} position of a sorted list, given the mechanism in place and the fact that the first $k-1$ advertisers has failed to meet consumer need, i.e. $\bar{\theta}^{[k]} = \mathbb{E}[\theta^{[k]} | z^1 = \cdots = z^{k-1} = 0, X(.)]$. As consumers believe that ads are sorted, conditioned on reaching position k in the sorted list, a consumer will click on the ad if $s_j \leq \bar{\theta}^{[k]}$, which happens with probability $G(\bar{\theta}^{[k]})$. Note that $\bar{\theta}^{[k]}$ represents a *conditional* mean. [2] shows that if qualities are uniformly distributed on $[0,1]$ and precisely M ads are posted, this probability can be computed in closed form:

$$P\big(s_j \leq \bar{\theta}^{[k]}\big) = G\big(\frac{N+1-k}{N+k}\big)$$

Define $C_{1\times M}$ as the vector of the above probabilities for the M positions:

$$C = (G(\bar{\theta}^{[1]}), \cdots, G(\bar{\theta}^{[k]}), \cdots, G(\bar{\theta}^{[M]})) \tag{3}$$

Under uniform assumption for both qualities and search costs C can be written as the following:

$$C = (\frac{N}{N+1}, \frac{N-1}{N+2}, \cdots, \frac{N+1-k}{N+k}, \cdots) \tag{4}$$

The click through rate at position k is $CTR_k = r_k C_k$, so we can write the vector of click through rates as:

$$CTR = r(\hat{\theta}, \theta) \circ C$$

which is the Hadamard (element-by-element) product of the two vectors r and C. Note that unlike [3], click through rates are endogenously determined in this model. Finally, utility of advertiser i can be written as:

$$u_i\big(\theta, X(\hat{\theta})\big) = \Big(X_i(\hat{\theta}) \centerdot \big(r(\hat{\theta}, \theta) \circ C\big)\Big)\theta_i - T_i(\hat{\theta}; \theta) \tag{5}$$

$$u_i\big(\theta, X(\hat{\theta})\big) = \Big(X_i(\hat{\theta}) \centerdot \big(r(\hat{\theta}, \theta) \circ C\big)\Big)(\theta_i - t_i(X)) \tag{6}$$

One can observe that the utility of each player depends on all the reported types (through the decision rule), his own true type, and the true type of all the players positioned in higher slots. This makes it clear why we are not in a private value setting anymore.

In the next section we will first define the optimal mechanism properly and solve for a restricted version of it in order to provide some intuition for the regularity condition required for existence of optimal mechanism, and how it is different from a private value setting. With this primary intuition, we proceed to characterizing the full-fledged optimal mechanism without the restriction. We fully characterize the optimal mechanism for some parameter values and provide partial generalizations for other case.

3 Optimal Mechanism with Fixed Number of Ads

In this section, we will solve for the optimal revenue maximizing mechanism in which the mechanism designer posts and ordered list of precisely M ads, so there will not be any empty ad slots. Intuitively, posting M ads might be suboptimal if ad qualities are all fairly low, but we start with this restriction in order to show a necessary condition for existence of optimal mechanism in the common value setting with endogenous click through rates, which is stronger than what is required when click through rates are exogenous.

Note that by revelation principal we can restrict our attention to the set of direct mechanisms $\big(X(\hat{\theta}), T_1(\hat{\theta};\theta), \cdots, T_N(\hat{\theta};\theta)\big)$, in which the auctioneer (search engine) allocates M non-homogeneous goods (ad slots) to M players (advertisers). The game is the following: The search engine asks advertiser to report their types (i.e. quality of meeting the consumer's need). The reports are submitted and based in the submitted reports, M advertisers are chosen and their ads are posted on the search page ($X(.)$). In addition, each advertiser should make a transfer to the auctioneer($\{T_i(.)\}_{i=1}^N$).

Bayesian incentive compatibility requires that

$$\mathbb{E}_{\theta_{-i}}\Big[u_i\big(X(\theta_i,\theta_{-i}), T_i(\theta_i,\theta_{-i}),\theta\big)\Big] \geq \mathbb{E}_{\theta_{-i}}\Big[u_i\big(X(\hat{\theta}_i,\theta_{-i}), T_i(\hat{\theta}_i,\theta_{-i}),\theta\big)\Big] \tag{7}$$

i.e., conditioned on the fact that all other bidders are playing truthfully (reporting their true types), each player i can not benefit from misreporting his type. Note that advertiser i's utility depends on all the reports, his own true quality as well as true quality of some other advertisers, specifically advertisers who end up being listed before i. As mentioned earlier, this is why we have a common-value setting.

Finally, for any distribution $\Phi(x)$, let $V(x) = x - \frac{1-\Phi(x)}{\phi(x)}$ be the virtual valuation function, and define $I(x)$ as the ratio of the virtual valuation to quality, i.e. $I(x) = \frac{V(x)}{x}$. Distributions with non-decreasing virtual valuations are called regular distributions. [6] has defined the following strengthened version of regularity, which turns out to be a key property for the optimal mechanism here:

Definition 1. *(Hyper-regularity) [6] A hyper-regular distribution is a regular distribution with non-decreasing $I(x) = \frac{V(x)}{x}$.*

Given the above definition, the following theorem summarizes the main result of this section:

Theorem 1. *Let advertiser qualities be drawn independently at random from a distribution $\Phi(.)$ on unit interval. The auctioneer is restricted to posting an ordered list of exactly M sponsored links. Consider the allocation rule X^* in which the auctioneer posts the ordered list of M highest-quality ads in descending order of quality. Define the set of expected transfers $\{\bar{T}_i(\hat{\theta};\theta)\}_{i=1}^N$ according to Myerson 1981 [8] to satisfy Bayesian incentive compatibility.*

For any arbitrary (non-negative) distribution of search costs $G(.)$, a sufficient condition for decision rule X^* to be revenue maximizing and Bayesian incentive compatible is that $\Phi(.)$ is hyper-regular.

In the rest of this section, we will explain the main steps of the proof using the special case where both $\Phi(\theta)$ and $G(s)$ are uniform distribution. Details of the derivation can be found in the appendix.

Expected utility of advertiser i from reporting truthfully when all other players are also reporting truthfully can be written as:

$$U_i(\theta_i) = \theta_i \bar{v}_i(\theta_i) + \bar{T}_i(\theta_i)$$

The main important point is to realize that although the total transfer $T_i(.)$ depends directly on *some of* the true types (beyond its dependence on the reports), it does not depend on own true type directly since all the direct dependence comes from the click through rate at position that advertiser i gets, which in turn only depends on the true quality of advertisers listed before i and not on the true quality of i himself.

As a result, one can apply the classic result of Myerson 1981 [8] to calculate the set of Bayesian incentive compatible transfers as the following:

$$\bar{T}_i(\theta_i) = U_i(\underline{\theta}_i) - \theta_i \bar{v}_i(\theta_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s)ds$$

with $U_i(0) = 0$. Note that once the optimal allocation rule is driven, one needs to make sure $\bar{v}_i(\theta_i)$is increasing.

The auctioneer seeks to maximize his total revenue, i.e. the total transfer from all the advertisers, which can be written as $\sum_i \mathbb{E}_\theta\big[-T_i(\theta_i)\big]$, or equivalently, $\sum_i \mathbb{E}_{\theta_i}\big[-\bar{T}_i(\theta_i)\big]$. The auctioneers maximization problem will be:

$$\max_{X(.),\{U_i(.)\}_{i=1}^N} \sum_i \int_0^1 \big[\theta_i \bar{v}_i(\theta_i) - U_i(\theta_i)\big]\phi_i(\theta_i)d\theta_i \tag{8}$$

$$\text{where } \bar{v}_i(\hat{\theta}_i) = \int_{\theta_{-i}} \Big(X_i(\hat{\theta}_i, \theta_{-i}) \,\mathbf{.}\, \big(r(\hat{\theta}_i, \theta_{-i}) \circ C\big)\Big)$$

$$\begin{aligned}
\text{s.t.} \quad & \text{(i) } X(.) \quad \text{feasible} \\
& \text{(ii) } \bar{v}_i(\theta_i) \quad \text{non-decreasing} \quad \forall i \\
& \text{(iii) } U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_0^{\theta_i} \bar{v}_i(s)ds, \quad \forall i \quad \forall \theta_i \\
& \text{(iv) } U_i(\theta_i) \geq 0 \qquad \forall i \quad \forall \theta_i
\end{aligned}$$

With some further algebra, and specializing $\phi(.)$ and $G(.)$ to be the uniform distribution the above constraint maximization simplifies to

$$max_{X(.)} \int_0^1 \cdots \int_0^1 \Big(\sum_{i=1}^{N} \big(X_i(\theta) \cdot \big(r(\theta) \circ C\big)\big)\big(2\theta_i - 1\big)\Big) d\theta_N \cdots d\theta_1 \tag{9}$$

subject to constraints (i) and (ii).

We will first look for a feasible allocation rule $X(.)$, and then show that the resulting $\bar{v_i}(.)$ is increasing for each advertiser i. We claim that efficient ordering, i.e. sorting the sponsored ad list in descending order of qualities, maximizes the auctioneer's revenue. In order to prove the above claim, we will first prove two intermediate lemmas which will then entail the main theorem.

Let $J(.)$ denote the expression we are integrating over. i.e.

$$J(X(.), \theta) = \sum_{i=1}^{n} \big(X_i(\theta) \cdot \big(r(\theta) \circ C\big)\big)\big(2\theta_i - 1\big) \tag{10}$$

If we can find an allocation rule $X(.)$ which maximizes $J(.)$ for every vector of realized θ, then clearly it will maximize the expected revenue (i.e the $J(.)$ function integrated over all realizations of θ) as well. Note that the value of $J(.)$ only depends on the M qualities which are chosen by the allocation rule on the list.

Consider the following function $H(.)$, which takes n numbers $a_1, a_2, \cdots, a_n$ along with an integer $m \leq n$ as input:

$$H(a_1, \cdots, a_n; m) = \sum_{i=1}^{m} C_i \Big(\prod_{j=1}^{i-1} (1 - a_j) \Big) \big(2a_i - 1\big) \tag{11}$$

where C_i is the i^{th} entry of vector C defined earlier. There are two key facts which are worth mentioning: First, only the first m input arguments affect the value of the function. Second, if we switch the position of two (unequal) inputs, e.g a_i and a_j (where $i < j < m$), then the value of $H(.)$ function will change, so the order of inputs matters for the function.

The important point to note here is that the $H(.)$ function is actually the $J(.)$ function rewritten in a specific way. To be more precise, Let θ be the N-vector of realized qualities. For any allocation rule $X(.)$, consider the M quantities which are chosen to be posted on the list, in the same order as implied by $X(.)$; i.e $(\theta^{[1]}, \theta^{[2]}, \cdots, \theta^{[M]})$. Let $\theta^{\{M+1 \cdots N\}}$ denote the remaining θ_i's in any random order (i.e. an $(N - M)$-vector), and append it to the above list. We have:

$$J((\theta_1, \theta_2, \cdots, \theta_N, X(.)) = H(\theta^{[1]}, \theta^{[2]}, \cdots, \theta^{[M]}, \theta^{\{M+1 \cdots N\}}; M).$$

Intuitively, the k^{th} term if the $H(.)$ function corresponds to the payment from advertiser in the k^{th} position of the ad list . Since advertisers who are off the list do not make any payments, there is no term associated with them in the $H(.)$ function. The following lemmas establish two useful properties of the $H(.)$ function.

Lemma 1. *Consider the n-sequence $(a_1, a_2, \cdots, a_n)$. If $a_k < a_{k+1}$ $(k+1 \leq m)$, then switching entries k and $k+1$ can only increase $H(.)$; i.e.*

$$H(a_1, \cdots, a_k, a_{k+1}, \cdots, a_n, ; m) < H(a_1, \cdots, a_{k+1}, a_k, \cdots, a_n; m)$$

Proof. In the appendix. □

Lemma 2. *If $a_m < a_x$ $(x > m)$, then switching entries a_m and a_x can only increase $H(.)$; i.e.*

$$H(a_1, \cdots, a_m, \cdots a_x, \cdots, a_n; m) < H(a_1, \cdots, a_x, \cdots, a_m, \cdots, a_n; m)$$

Proof. In the appendix. □

With the above two lemmas we can prove our result in the special case of uniform qualities and search costs. We present this special case and its proof in the following corollary, and later use a very similar argument to prove the general theorem.

Corollary 1. *When advertiser qualities and consumer search costs are drawn independently at random from uniform distribution with support $[0, 1]$, efficient ordering, i.e. sorting the sponsored ad list in descending order of qualities, maximizes the auctioneer's revenue and is Bayesian incentive compatible.*

Proof. Note that we can model the auctioneer choice as a two phase process: First, M advertiser's are chosen and then they are sorted in the desired order. The proof proceeds in two steps: Step one shows that any list of chosen ads should be sorted in decreasing order of quality (phase 2), and step two shows that the M highest-quality ads must be chosen (phase 1). Both steps use the $H(.)$ function defined above.

Let us assume the revenue maximizing list of chosen ads is not the M highest qualities sorted in descending order. Let $\tilde{\theta}$ denote the list of the N qualities in the following order: The first M entries are the sorted list of ads chosen by the auctioneer, and the last $N - M$ entries are the qualities of the remaining advertisers (not posted online) in any random order.The assumption that efficient ordering is not revenue maximizing implies:

$$(\tilde{\theta}_1, \tilde{\theta}_2, \cdots, \tilde{\theta}_M) \neq (\theta^1, \theta^2, \cdots, \theta^M).$$

Recall that only $(\tilde{\theta}_1, \tilde{\theta}_2, \cdots, \tilde{\theta}_M)$ affects the revenue, and the quality of the rest of the ads are irrelevant. First assume that $\tilde{\theta}$ violates the first condition so that we don't have $\tilde{\theta}_1 > \tilde{\theta}_2 > \cdots > \tilde{\theta}_M$. Let $\tilde{\theta}_i$ and $\tilde{\theta}_j$ be a pair of entries violating the above condition, i.e. we have $i < j$ and $\tilde{\theta}_i < \tilde{\theta}_j$. By Lemma 1, exchanging $\tilde{\theta}_i$ and $\tilde{\theta}_j$ only increases the revenue, so the initial ordering could not be revenue maximizing, which contradicts the assumption. As a result, any $\tilde{\theta}$ chosen by the auctioneer should have the property that its first M elements are sorted in decreasing order.

Now assume that $\tilde{\theta}$ indeed satisfies the first condition, but it violated the second condition, so that there exists an advertiser whose quality is between the M highest qualities, but he is not included among the sponsored link ads. As a result, there exist i, j such that

$i \leq M < j$ and $\tilde{\theta}_i < \tilde{\theta}_j$. Since the first M elements are sorted in decreasing orders, we will have $\tilde{\theta}_M < \tilde{\theta}_j$. Now by Lemma 2 switching $\tilde{\theta}_M$ and $\tilde{\theta}_j$ can only increase the revenue and consequently the initial ordering could not be revenue maximizing.

The last thing that we need to show is that $\bar{v}_i(.)$ is increasing in θ_i. This is trivial since with this allocation rule, a higher θ_i results in a higher position, and so a larger click-through rate (i.e larger entries in both C and r vector). Since $\bar{v}_i(.)$ exactly corresponds to the click-through rate that the advertiser will get, it will be increasing in θ_i which completes the proof. □

In order to get the results for general distributions, we need two more lemmas. Note that these generalizations are interesting because its an inherently hard problem to approximate these distributions, specially the search cost distribution. The next two lemmas establish the result of the paper for any distribution of the search costs.

Lemma 3. *For any arbitrary distribution of qualities, $\phi(.)$, if consumers believe that ads are sorted in descending order of quality, their belief about the quality of an ad conditioned on not being fulfilled by higher ads is decreasing as they go down the list; i.e. $\bar{\theta}^{[k]} \geq \bar{\theta}^{[k+1]}$.*

Proof. In the appendix. □

Lemma 4. *When advertiser qualities are drawn independently at random from distribution $\phi(.)$, for any arbitrary (non-negative) distribution of search costs, efficient ordering, i.e. sorting the sponsored ad list in descending order of qualities, maximizes the auctioneer's revenue.*

Proof. In the appendix. □

We now have all the necessary means to prove the main result of the paper, which can be viewed as a natural extension of the proof for Theorem 1. This proof is provided in the appendix.

Theorem 1 says that there is a sufficient condition under which the search engine has no incentive to use a mix strategy in allocating ad slots, neither does he have an incentive to leave any high quality ads out of the posted list, and his decision rule is a BNE. This is the analogous result to the classic optimal mechanism design problem (as studies for instance in Myerson 1981 [8]. However, the important point is that the sufficient condition for Bayesian incentive compatibility is *not* the same in the two cases: in particular, it is stronger here. In other words, $I(\theta_i)$ increasing implies that $V(\theta_i))$ is increasing, but the reverse is not true.

The intuition behind this difference is very interesting. In most applications, when an object is allocated to a bidder, he gets the benefit with probability one, so one only needs to adjust for the rent required to induce truth telling and the adjusted value, i.e. virtual valuation, should be increasing. This is not the case in the current problem. Here we have that if advertiser i is allocated a spot, he would get a match only with probability θ_i, so the virtual valuation adjusted for this probability is what should be increasing, i.e. $I(\theta_i)$ rather than $V(\theta_i)$.

4 Optimal Number of Ads

Assume the auctioneer commits to posting *at most* M, instead of exactly M links. In this more general specification, the feasibility of an allocation rule is slightly different. Specifically, condition (1) remains the same as before, but (2) is relaxed (into two conditions). i.e, we will have:

$$\sum_j X_{ij} \leq 1 \qquad 1 \leq i \leq N \tag{12}$$

$$\sum_i X_{ij} \leq 1 \qquad 1 \leq j \leq M \tag{13}$$

$$X_{ik} > 0 \text{ only if } \sum_i X_{ij} = 1 \qquad j \leq k \leq M \tag{14}$$

13 says there can be empty slots (i.e. number of posted ads can be less than number of available positions), while 14 says slot k can be filled only if all the slots above k are already filled up.

The revenue function keeps its original form, but the auctioneer have an extra degree of freedom on the number of ads to post online, besides choosing which ads to post. As a result, the generalized maximization problem can be written as

$$max_{X(.),K\leq M} \int_0^1 \cdots \int_0^1 \Big(\sum_{i=1}^{N} \big(X_i(\theta) \centerdot \big(r(\theta) \circ C\big)\big)\big(2\theta_i - 1\big)\Big) d\theta_N \cdots d\theta_1$$

subject to the same conditions as specified in (8). Note that K is the number of posted ads implied by $X(.)$, and we have explicitly added it into choice variable only to be more clear. The change is that in this generalized version, condition (ii) in (8), i.e. the feasibility of $X(.)$ is defined by (12,13,14) instead of (1,2).

This new choice variable for auctioneer has two implications. Recall that the auctioneer's optimization problem specified by (19) is the sum of M terms, where the i^{th} term is associated with $(2\theta_i - 1)$. Note that if $\theta_i < \frac{1}{2}$, the i^{th} term along with all the following terms of the revenue function will be negative, which means they are actually reducing the auctioneers revenue, so the auctioneer is better off shortening the list and dropping these negative terms. The second implication is through vector C. The difficulty is that the number of ads posted, k, affects conditioning information for consumer in computing $\bar{\theta}^{[k]}$.

$$\bar{\theta}^{[k]} = \mathbb{E}[\theta^{[i]} | z^1 = \cdots = z^{k-1} = 0, K \text{ ads are posted}]$$

One can imagine that a larger K will potentially increase the above expectation for each $k \leq K$ and so enhance the revenue as consumers would think that there are more "high quality" ads, which gives the auctioneer an incentive to post as many ads as possible. However, the fact is that consumers will endogenize this effect and won't be "fooled". Nevertheless, the feed-back effect is in fact present, and in the next section we will show that in a special case it goes in the reverse direction. We believe the intuition carries over to the general case.

Note that since the number of bidders is N (potentially $N >> M$), the choice of K can in principal depend on all the reports. Before we proceed to the special case, we will present a useful lemma that would which restrict the number of relevant reports in determining K to M highest reports.

Lemma 5. *With N advertisers and M ads ($N > M$), at most the highest M reported qualities are relevant to determine K. Moreover, the K ads posted in equilibrium are the K highest quality ads ordered in descending order of quality.*

Proof. In the appendix. □

Next we characterize auctioneers optimal decision when there is only one possible position, i.e. $M = 1$. Note that with a single available position, the search engine can post either one ad or nothing (i.e. $K = 1$ or $K = 0$). In other words, a single available position does *not* imply a single posted ad.

By Lemma 5 we already know that only the highest reported quality is relevant for auctioneer's optimal strategy. Also, Proposition 1 ensures that bidders report truthfully if the allocation rule is monotone in each bidder's report. So we assume truthful reports and then show that the optimal auctioneer strategy is in fact monotonically increasing in each bidder report. Throughout the rest of this section, we maintain the assumption that consumer search costs are uniformly distributed. The assumption is made only for clarity and all the results go through if the assumption is relaxed.

The following theorem lays out the optimal decision rule for the auctioneer, which is basically the decision whether to post an ad or not.

Theorem 2. *When there is only one possible ad position available, the optimal decision rule for the auctioneer whether to post one ad or not is a cut-off rule. Specifically, there exists a θ^* such that he posts the highest ad if it is above θ^*, and otherwise posts no ads.*

Moreover, let $\tilde{\theta}$ be such that $V(\tilde{\theta}) = 0$. The optimal decision is such that $\theta^* > \tilde{\theta}$, i.e. in order to generate higher expected revenue from higher quality bidders the auctioneer must raise the bar; i.e. commit to post only very high quality ads (through a high reserved price).

Proof. In the appendix. □

Why is the above comparison interesting? Note that $\tilde{\theta}$ is what the cut-off would be if consumer beliefs were exogenous. With the endogenous beliefs, there is a *new trade-off* which leads the auctioneer to be more restrictive than before. The old reason to restrict the allocation of the object, i.e. setting a positive reserve price to maximize revenue, is still present here and restrict the low value bidders; but there is also a new reason to restrict even further which is improvement in beliefs conditional on seeing an ad.

The idea of the proof is the following: posting better quality ads has two positive effects for the auctioneer: First, bidders with higher qualities are willing to pay more since their expected payoff conditional on receiving a click is higher. Second, it improves consumer's expectation of the posted ad (if one is posted). As a result, the auctioneer commits not to post some *moderate* qualities which generate positive expected revenue themselves ($\tilde{\theta} < \theta < \theta^*$ as defined in the theorem) in order to improve consumer's belief about the quality of the ads who end of being posted and generate more revenue from them. In other words, the gain from improved consumer belief about better quality ads outweighs the loss from not posting the moderate ads. A rigorous proof for general distributions is provided in the appendix, but the following example can be very useful to understand the essence of the proof.

Example 1. *One available position, and N bidders.*

Take the very simple case where there are N advertisers and only one available ad position. Assume the quality of each bidder is drawn from a hypothetical distribution $\hat{\phi}(.)$ such that the first order statistic is uniformly distributed on the unit interval. The search engine can choose to post one bidder or not.

Given Lemma 5 when there is only one available position, only the first order statistic of the qualities matter for auctioneer decision whether or not to post an ad. With a little abuse of notation, let θ denote the first order statistics. The mechanism designer maximizes:

$$\frac{\int_0^1 \theta y(\theta) d\theta}{\int_0^1 y(\theta) d\theta} \int_0^1 V(\theta) y(\theta) d\theta$$

where $y(.)$ is the allocation probability. Note that the first term is the expectation of consumers about the quality of the posted ad given auctioneer allocation rule, and the second term is the "classic" expected revenue from allocating a single object to a single bidder, i.e. if there was no feedback effect from allocation rule to consumer beliefs. Let $L = \int_0^1 y(\theta) d\theta$. Note that value of L is invariant to any reordering of $y(.)$. As a result, for any *fixed* L, "swapping" high $y(.)$'s from low to high θ increases both terms in numerator (while keeping the denominator constant by assumption). Consequently, $y(.)$ is a step function and the jump (from 0 to 1) occurs at $1 - L$, i.e.

$$\hat{y}(\theta) = \begin{cases} 0 & \text{if } \theta < 1 - L \\ 1 & \text{if } \theta \geq L \end{cases}$$

So the above optimization problem boils down to choice of optimal L (equivalently the cut-off for $y(.)$).

For numerical purposes, assume $N = 1$.[3] For this particular example with uniform distribution, the jump of the step function happens at $\theta^* = \frac{\sqrt{3}}{3} > \tilde{\theta} = \frac{1}{2}$. In other words, the auctioneer chooses *not to* post some positive-revenue generating ads in order to boost consumer expectation of the posted ads.

The next theorem provides a generalization of the previous theorem to a general case of multiple ad positions. The full characterization of the optimal mechanism remains for future work.

[3] Why am we making this "extra" assumption? First of all note that conceptually, this is without loss of generality since with one position only the first order statistic matter. Second, note that with one position (and any number of bidders), the function we are integrating over is $E[\theta^{[k]}|y(.)]V(\theta)y(\theta)$ where the density in $V(\theta)$ is $\phi(\theta)$, the distribution of *each* quality. On the other hand, the distribution function we are integrating over is the distribution of *first order statistic*, $\phi_1(\theta)$. $N = 1$ means $\phi(.) = \phi_1(.) = U[0, 1]$, which gives a clean cut numerical example.

Theorem 3. *Assume there are N bidders and M ad positions, and let $\tilde{\theta}$ be such that $V(\tilde{\theta}) = 0$. No ad with quality $\theta < \tilde{\theta}$ will be posted.*

Proof. In the appendix. □

In order to build some intuition, consider the following different belief structure: Assume consumers did not know what the true M is ex-ante and they believed it is equal to the number of posted ads, i.e. they believed $M = K$. Under such belief structure the number of posted ads does not enter the consumer information set that they use to compute the expected ad qualities, so the optimal mechanism for the auctioneer is to use the mechanism devised in the previous section (i.e. order the advertisers in descending order of reported qualities), but on top of that cut out all the ads with quality θ such that $V(\theta) < 0$ since these ads generate negative revenue. This same reserve price would be optimal if click through rates were exogenous.

Comparing the original and these latter auction conveys the main intuition behind the theorem: The *feedback* effect from the optimal allocation rule to consumer belief leads the auctioneer to be more picky about the choice of ads presented to consumers. This makes the consumers more confident that they are more likely to get what they want if they click on the sponsored ad, so they would click more on these higher quality ads. This is better for the auctioneer: The auctioneer extracts more revenue from the better quality ads who are more willing to pay more at the expense of not generating any revenue from some moderate ads.[4] In other words, being able to adjust the number of ads gives the auctioneer and extra degree of freedom to signal the ad qualities to consumers and this ability incentivize the mechanism designer to be more strict in his criteria for posting ads. Without this signal, he has no incentive to leave out any ad that can generate positive expected revenue since consumer belief about the quality of the posted ads cannot be affected anyways.

The following two examples provides some more intuition about Theorem 3.

Example 2. *Two available positions with the same reserve price, and N bidders.*

Assume there are N advertisers and two possible ad positions, $M = 2$. Also, assume both qualities and consumer search costs are drawn from uniform distribution over the unit interval, $F(\theta) = \theta$. Finally, assume that the search engine uses a reserve pricing rule to post ads, and he can not set different reserve prices for the two positions.[5] Let c denote the

[4]In general expected consumer welfare, i.e. expected needs met minus expected search cost is not maximized at the same reserve price. In section 5 we show that if ad qualities are presented to consumers, cut-off $V(\tilde{\theta}) = 0$ is revenue maximizing, but welfare is not maximized at the same cut-off. The reason is that with full information, consumers are best off being presented with as many ads as possible (only constrained by maximum number of ad positions in the page); which is not optimal for the advertiser. When only ordinal information is presented the problem is more subtle: consumers get an extra signal when the auctioneer is able to choose the number of ads compared to when he is not, and in equilibrium this signal is such that it gets more consumers to click on better ads which is welfare enhancing. On the other hand, even more ads are left out which is welfare destroying.

[5]Obviously, removing this restriction can only increase the expected revenue since without this restriction, the auctioneer can still choose the same reserved price if he wants to. So this example is not characterizing the optimal reserve prices. It is designed to show that the optimal reserve price assuming independent valuations is *not* optimal now that we have interdependent values.

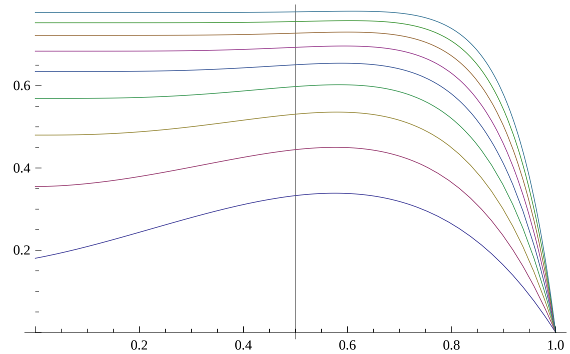

Figure 1: Expected revenue as a function of reserve price for different number of advertisers when there are two positions, and the same reserve price is used for both positions. Revenue curve shifts up and flattens as N increases.

optimal reserve price. By lemma (5), only the first order statistic (FOS) and second order statistic (SOC) of the reported qualities matter for the search engine.

The search engine solves the following maximization problem:

$$\max_c \int_c^1 \int_0^c \mathbb{E}[\theta^1|\theta^1 \geq c, \theta^2 < c] V(\theta^1) f_{1,2}(\theta^1, \theta^2) d\theta^2 d\theta^1 +$$
$$\int_c^1 \int_c^{\theta^1} \Big[\mathbb{E}[\theta^1|\theta^1, \theta^2 \geq c] V(\theta^1) + \mathbb{E}[\theta^2|\theta^1, \theta^2 \geq c, z^1 = 0](1-\theta^1) V(\theta^2)\Big] f_{1,2}(\theta^1, \theta^2) d\theta^2 d\theta^1$$

where

$\mathbb{E}[\theta^1|\theta^1 \geq c, \theta^2 < c]$: expectation of FOS when FOS/SOS is larger/smaller than c.

$\mathbb{E}[\theta^1|\theta^1, \theta^2 \geq c]$: expectation of FOS when FOS & SOS are larger than c.

$\mathbb{E}[\theta^2|\theta^1, \theta^2 \geq c, z^1 = 0]$: expectation of SOS when FOS & SOS are larger than c and first ad did not meet consumer need.

$$V(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)}$$

Note that there is a slight abuse of notation in using $V(.)$, i.e. $V(x) = x - \frac{1-F(x)}{f(x)}$ (i.e. $2x - 1$ in this particular example), regardless of the actual distribution of x. In particular, in this example θ^1 and θ^2 are first and second order statistics of N uniform random variables so they are not distributed uniformly anymore. Also, recall that z^1 is the Bernoulli random variable which takes value 0 if a match does not happen when the consumer clicks on the first ad. Finally, $f_{1,2}(\theta^1, \theta^2) = NF(\theta^2)^{N-2}$ is the joint distribution of first and second order statistic.

The first term in the sum is the expected revenue when only one ad is posted, i.e. only the highest order statistic of qualities is above the reserve price c. The second term has both highest and second highest quality ads posted since they are both above the reserve price. Note that the second term has two parts: the expected revenue from the first ad and the expected revenue from the second ad which is clicked if the first click is unsuccessful.

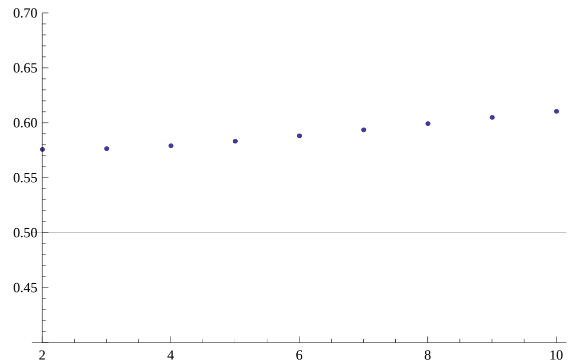

Figure 2: Optimal reserve price as a function of number of advertisers. There are two ad positions and the same reserve price is used for both.

Figure 1 shows how the expected revenue changes as a function of reserve price for different values of N. The lowest curve corresponds to $N = 2$. The revenue is maximized at a single reserve price $c > \frac{1}{2}$, which is in line with Theorem 3. Theorem 3 implies that with endogenous click through rates (due to endogenous beliefs), the mechanism designer restrict the set of chosen ads further compared to if beliefs where exogenous. The above example shows that although setting a single reserve price for two positions may not be the optimal decision rule for the mechanism designer, even in the set of such rules the best reserve price is greater than $\frac{1}{2}$, which would be the optimal reserve price if beliefs where exogenous and we where in a Myerson 1981 [8] world. As N increases, revenue curve shifts up pointwise since the expectation of first and second order statistics are increasing in number of advertiser. Moreover, the revenue curve becomes almost flat and then it drops. In other words, revenue becomes less sensitive to the choice of reserve price which is very intuitive: as N gets large it becomes more likely that the first two order statistics are quite large, so there is not much gain to setting a reserve price. However, the optimal reserve price is increasing in N, although the sensitivity of revenue to reserve price sharply decreases as N increases. Figure 2 depicts the optimal (same) reserve price (for two positions) for different number of advertisers.

Example 3. *Two available positions with (potentially) different reserve prices, and N bidders.*

If the mechanism designer is restricted to use reserve pricing, but can set different reserve prices for the two slots, the optimal rule would be of the following form:

$$y(\theta^1, \theta^2) = \begin{cases} \text{2 ads posted} & \theta^1 \geq c_2, \theta^2 \geq c_2 \\ \text{1 ad posted} & \theta^1 \geq c_1, \theta^2 < c_2 \\ \text{No ad posted} & \text{Otherwise} \end{cases}$$

where $c_1(c_2)$. The above rule implies that c_1 (c_2) is the reserve price on the first (second) position. It is important to note that $c_1 \leq c_2$, otherwise the decision rule is inconsistent. Also, since by definition $\theta^2 \leq \theta_1$, with some more algebra it can be shown that the above rule is the only consistent rule with two (potentially) different reserve prices.

With this rule in mind, the search engine objective function is very similar to Example 2:

$$\max_{c_1,c_2} \int_{c_1}^{1} \int_{0}^{\min\{c_2,\theta^1\}} \mathbb{E}[\theta^1|\theta^1 \geq c_1, \theta^2 < c_2] V(\theta^1) f_{1,2}(\theta^1,\theta^2) d\theta^2 d\theta^1 +$$
$$\int_{c_2}^{1} \int_{c_2}^{\theta^1} \Big[\mathbb{E}[\theta^1|\theta^1,\theta^2 \geq c_2]V(\theta^1) + \mathbb{E}[\theta^2|\theta^1,\theta^2 \geq c_2, z^1 = 0](1-\theta^1)V(\theta^2)\Big] f_{1,2}(\theta^1,\theta^2) d\theta^2 d\theta^1$$

Numerical maximization shows that for small values of N, the mechanism designer chooses $c_1 \approx c_2$. As N gets large the gap between the two reserve prices widen, but as mentioned in Example 2 the left tail of the revenue curve is almost flat for large N, so the effect of lowering c_1 compared to c_2 on expected revenue is almost negligible.

So what is the punch line? we have shown that similar to a setting with private valuations, with common values the auctioneer always chooses to post the highest quality ads, which is not surprising (although the required regularity condition is more strict here). Where the two settings differ when it comes to setting the reserve price: with common valuation the optimal reserve price is higher. So if a search engine is setting reserve prices for sponsored search ads using traditional formulas, they are doing it wrong! They should set higher reserve prices for a very clear reason: The search engine wants a reserve price not just to eliminate negative marginal revenue bidders at the current beliefs, but also to improve beliefs about ad quality.

5 Posting List of Reported Qualities

In this section, we study the scenario in which the auctioneer posts the qualities reported by the advertisers along with their ads on the search page. Note that the cardinal information imposes a natural ordering on the ads (order of being clicked). We start with uniform distribution of qualities, but arbitrary distribution $G(.)$ for search costs. We can formulate the advertiser's utility function in a similar fashion as in the previous section with one modification: As the consumers observe the exact qualities, their expectation of link qualities does not affect their search behavior, and they click only if the posted quality is larger than their search cost. As a result, we will have:

$$u_i\big(\theta, X(\hat{\theta})\big) = \big(X_i(\hat{\theta}) \cdot r(\hat{\theta},\theta)\big) G(\hat{\theta}_i)\big(\theta_i - t_i(X)\big) \tag{15}$$

$$u_i\big(\theta, X(\hat{\theta})\big) = \big(X_i(\hat{\theta}) \cdot r(\hat{\theta},\theta)\big) G(\hat{\theta}_i)\theta_i - T_i(X) \tag{16}$$

Observe that as in the previous section, advertiser utility depends on his own true type only linearly (although it depends on his announced type in a more complicated fashion than before). Consequently, all the arguments of the previous section hold here as well. We will state the main theorem of this section, which is analogous to Theorem 1 in a general setting, and then derive the special case of uniform qualities and uniform search costs as a corollary.

There are a few subtleties involved with this problem. The first one relates to the definition of the allocation rule $X(.)$. Recall that in the previous case, the order in which ads were shown determined the order in which they were being clicked. Note that here the auctioneer only decides on which M ads to post online, and after that there is no choice of ordering, since the posted qualities themselves impose a specific order of being clicked. So basically in our notation for $X(.)$, the auctioneer chooses which rows he want to put a 1 in (i.e. which ads he wants to post online), and the column of the 1 in each row is determined by the rank of the quality of the advertiser who corresponds to that row, relative to the rank of the other chosen advertisers. This means that the auctioneer is not "free" to choose any allocation that he wants, i.e. since the consumers directly observe the cardinal information on the reported qualities, the auctioneer can not attempt to increase his revenue by misreporting the ordinal information. As an example, if $\hat{\theta}_i > \hat{\theta}_j$, the following allocation is invalid: $X_{ik_1} = X_{jk_2} = 1$, $k_1 > k_2$, because observing the $\hat{\theta}_i$'s, the consumers first click on i and then on j. As a result, in this case $X_{ij} = 1$ means that the auctioneer has chosen advertiser i, and $j-1$ other advertisers with higher quality.

The next subtle point is how to interpret Lemma 1 and Lemma 2 which we proved in the previous section. There, we basically devised a multi-step algorithm which the auctioneer was to take and go from a non-efficient allocation to the efficient one and only increase the revenue along the way. Here, those steps can not be materialized anymore, i.e. they don't have real world realization (i.e., the auctioneer can not enforce the consumers to click on a lower-quality ad before a higher quality one). As a result, these steps only correspond to intermediate values which are used to show that revenue in one case is higher than the other case. Having these in mind, we state our last theorem which is analogous to Theorem 1:

Theorem 4. *Let advertiser qualities be drawn independently at random from a distribution $\Phi(.)$ with non-negative support $[\underline{\theta}, \bar{\theta}]$ and increasing virtual valuations, $V(\theta_i) = \theta_i - \frac{1-\Phi(\theta_i)}{\phi(\theta_i)}$. Also, Let $\tilde{\theta}$ be such that $V(\tilde{\theta}) = 0$. Let K be the number of ads with quality higher than $\tilde{\theta}$. Consider the decision rule $X^*(.)$ in which the auctioneer sets the reserved quality to be $\tilde{\theta}$, and posts the list of $min(M, K)$ highest-quality ads along with the their corresponding reported qualities. For any arbitrary (non-negative) distribution of search costs, $G(.)$, a sufficient condition for X^* to be revenue maximizing is that $\Phi(.)$ is hyper-regular for $\theta_i > x^*$; which is the same condition as in Theorem 1.*

Proof. In the appendix. □

The following corollary states the result for the special case where both search costs and qualities are uniform random variables:

Corollary 2. *Let advertiser qualities and consumer search costs be drawn independently at random from uniform distribution with support $[0, 1]$. Moreover, assume the auctioneer commits to posting at most M ads on the search page along with the reported qualities of the advertiser. In this scenario, the revenue maximizing allocation rule is to post $K \leq M$ highest-quality ads such that $\theta^K > 1/2$ and $\theta^{K+1} < 1/2$.*

6 Revenue Comparison

In this section, we study how the revenue generated from the mechanism with only ordered list of ads posted (section 4) compares to that of the mechanism in which ad qualities posted as well (section 5). This question is very important since it sheds light on how willing search engines are in providing information to consumers and has important welfare implications. In order to analyze this question we restrict my attention to the case where there are N advertisers and only one available ad position.

First note that there are two channels through which revenue can be enhanced: higher volume of search, and higher per-click payments made by advertisers. In addition, higher volume of search can itself come from two sourced: showing an ad more frequently, or more search conditional on an ad being shown.

Let SV_O (Search Volume Ordinal) and SV_Q (Search Volume Quality) denote the expected total search volume when ordinal and quality information about qualities are posted, respectively.[6]

$$SV_O = G(\mathbb{E}[\theta^1|\theta^1 > c)P(\theta^1 > c)$$
$$SV_Q = \int_{\frac{1}{2}}^{1} G(\theta^1) f_1(\theta^1) d\theta^1$$

where c is the optimal reserved price and $f_1(.)$ is the distribution of first order statistic.

Consider the simple case when qualities and search costs are uniformly distributed, $G(s) = s$, $\Phi(\theta) = \theta$. We have already proved that $c > \frac{1}{2}$, so the total volume of search is higher when quality is posted. Intuitively, when quality is not posted the search engine commits to restricting himself further by setting a higher reserve price in order to get more clicks when an ad is posted, but this commitment has a negative effect on the revenue as well since it means the search engine posts an ad less frequently. On the other hand, the number of clicks conditional on an ad being posted with no quality is higher than the expected number of clicks conditional on an ad being posted along with its quality. However with uniform search costs the first effect dominates and total search volume is always higher if quality is posted.

Difference in revenue depends on concavity of both quality and search cost distributions as well as number of bidders. Increasing the number of bidders acts through two different channels: first, it increases the probability of the best ad being good enough to be posted which enhances the revenue in both cases. Second, it enhances the expected quality of the first order statistic, which *only* improves the revenue when quality is not posted.

Numerical analysis using the following functions provides some intuition about the relative total search and revenue. We have used $F(\theta) = \theta^x$, and $G(s) = s^m$ or $G(s) = \frac{1}{1-e^{-m}}(1 - e^{-ms})$.[7] First note that $\tilde{(}\theta)$ is increasing in x (recall that $V(\tilde{\theta}) = 0$, so $\tilde{(}\theta)$ is the reserve price is quality is posted).

[6]In other words, SVO is the expected total search volume when only and ordered list of ads is posted, and SVC is the expected search volume when qualities are also posted.

[7]This is a variation of exponential distribution adjusted to be a distribution over the unit interval.

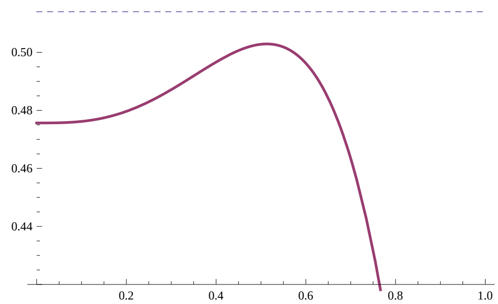

Figure 3: Rev_O (solid) and Rev_Q (dashed) as a function of reserve price c. Dashed and solid lines are the expected revenue when qualilies are and are not posted, respective. There are $N = 5$ adverstisers, quality distribution $F(\theta) = \theta^0.6$ and search cost distribution $G(s) = \frac{1}{1-e^{-2}}(1 - e^{-2s})$.

This analysis shows that for the aforementioned classes of functions, both total search volume and expected revenue is higher when quality is posted. Let Rev_O and Rev_Q denote the expected revenue when ordinal and quality information is posted, respectively. Recall that c is the reserve price used by the search engine if quality is not posted. Note that if quality is posted, search engine will use $\tilde{\theta}$ such that $V(\tilde{\theta}) = 0$ as the reserve price, and given the particular form used for $F(\theta)$, $\tilde{\theta}$ can be computed in closed form, $\tilde{\theta} = \frac{1}{1+x}^{-\frac{1}{x}}$.

The following two figures provide a schematic view of this result. The first figure plots the expected revenue when quality is posted and when it is not as a function of c for a set of randomly chosen parameters. The second figure shows the behavior of difference in revenue and total search as a function of reserve price c (used for posting ads when quality is not posted) for the following parameter values: $N = 5$, $x = 0.6$ and $m = 2$.

The dashed line is the expected revenue when quality is posted (Rev_Q) and the solid curve is the same thing when quality is not posted (Rev_O). Note that Rev_Q is independent of choice of c, so it is a constant line. Rev_O changes with c, and it is maximized at θ^* which is the optimal reserve price and satisfies the theoretical prediction: $\theta^* > \tilde{\theta}$.

The above parameter values were randomly chosen, but the relative placement of the two curves is robust to change in parameters, so the optimal mechanism if quality is not posted generates weakly lower expected revenue compared to the optimal mechanism if qualities are posted. As m gets larger and search cost distribution becomes more concave, search volume increases when quality is not posted, and the expected revenues converge. $N = 5$, $x = 0.6$ and $m = 2$ is one such example.

The dashed curve is the difference in expected revenue when quality is posted and when it is not, while the solid curve is the difference in total search volume among the two cases.

The above analysis suggests that to the extent that the above classes of distributions are good approximations to advertiser and consumer characteristics, search engines have a tendency to disclose the information that they obtain from advertisers (through their bidding mechanism) to the consumers. More precise characterization of classes of distributions for which this result holds remains for future research.

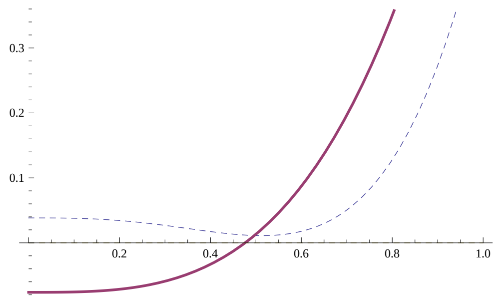

Figure 4: ΔRev=$Rev_Q - Rev_O$ (dashed) and ΔSearch=$SV_Q - SV_O$ (solid) as a function of reserve price c. Although posting ad qualities can decrease the search volume, but in enhances the revenue since search is better directed. There are $N = 5$ adverstisers, quality distribution $F(\theta) = \theta^0.6$ and search cost distribution $G(s) = \frac{1}{1-e^{-2}}(1 - e^{-2s})$.

7 Conclusion

In this paper we study the optimal revenue maximizing mechanisms in common value position auctions. For the most part of the paper, we assume that the search engine only reveals ordinal information about ad qualities, i.e. the search engine only conveys the relative ranking of the ads to consumers but does not tell them how good each ad is. We show that with endogenous click-through-rate *hyper-regularity* of distribution of ad qualities, a strengthened version of familiar regularity (increasing virtual valuations) condition, is a sufficient condition for existence of a non-ironed optimal mechanism. More importantly, we show that the mechanism designer displays fewer ads compared to the case where click-through-rates are exogenous. Consequently, if the search engine ignores the feed back effect of the allocation rule into consumer beliefs, too many ads will be posted compared to what is revenue maximizing.

We also study an extension in which cardinal information about ad qualities are posted and we show how the optimal mechanism differs with the former case. Finally, we provide some interesting revenue comparison results across the two cases, which can shed light on the amount of information about advertisers that search engines are willing to share with consumers.

References

[1] Aggarwal, G., A. Goel, and J. Motwani (2006) "Truthful Auctions for Pricing Search Keywords," *7th ACM Conference on Electronic Commerce (EC06).*

[2] Athey, S., and G. Ellison (2011) "Position Auctions with Consumer Search", *Quarterly Journal of Economics*, 126(3), 1213-1270.

[3] Edelman, B., M. Ostrovsky, and M. Schwarz (2007) "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords", *American Economic Review*, 97(1), 242-259.

[4] Edelman, B. and M. Schwarz (2010) "Optimal Auction Design and Equilibrium Selection in Sponsored Search Auctions," *American Economic Review*, 100(2), 597-602.

[5] Gomes, R.(2011) "Optimal Auction Design in Two-Sided Markets," Working Paper.

[6] Kleinberg, R., and Y. Yuan (2013) "On the Ratio of Revenue to Welfare in Single-Parameter Mechanism Design," *14th ACM Conference on Electronic Commerce (EC13).*

[7] McAfee, R. P. and P. Reny (1992) "Correlated Information and Mechanism Design", *Econometrica*, 60(2), 395-421.

[8] Myerson., R. (1981) "Optimal Auction Design", it Mathematics of Operations Research 6, 58-73

[9] Mas-colell A., Whinston, M. D. and Green J. R. (1995) *Microeconomics Theory*, New York, N.Y., Oxford University Press, Inc.

[10] Rayo, L., and I. Segal (2010) "Optimal Information Disclosure," *Journal of Political Economy*, 118(5), 949-987.

[11] Varian, H. R. (2007) "Position Auctions", *International Journal of Industrial Organization*, 25(6), 1163-1178.

[12] Weyl, G. (2010) " A Price Theory of Two-Sided Markets," *American Economic Review*, 100(4), 1642-1672.

8 Appendix

8.1 Derivations of Section 3

Substitute 5 in 7 to get the truth-telling incentive compatibility constraint:

$$\mathbb{E}_{\theta_{-i}}\left[\Big(X_i(\theta_i,\theta_{-i}) \centerdot \big(r((\theta_i,\theta_{-i}),\theta)\circ C\big)\Big)\theta_i - T_i(\theta_i,\theta_{-i})\right] \geq$$
$$\mathbb{E}_{\theta_{-i}}\left[\Big(X_i(\hat{\theta}_i,\theta_{-i}) \centerdot \big(r((\hat{\theta}_i,\theta_{-i}),\theta)\circ C\big)\Big)\theta_i - T_i(\hat{\theta}_i,\theta_{-i})\right]$$

Note that the mechanism designer can have a per-click payment rule $t_i(\hat{\theta})$ so that we have:

$$T_i\big(\theta, X_i(\hat{\theta})\big) = \Big(X_i(\hat{\theta}) \centerdot \big(r(\hat{\theta},\theta)\circ C\big)\Big)t_i(\hat{\theta})$$

Further define the "conditional benefit" as:

$$v_i\big(\theta, X(\hat{\theta})\big) = \Big(X_i(\hat{\theta}) \centerdot \big(r(\hat{\theta},\theta)\circ C\big)\Big)$$

So that each advertiser utility function can be written as:

$$u_i\big(\theta, X(\hat{\theta})\big) = \theta_i v_i\big(\theta, X(\hat{\theta})\big) + T_i\big(\theta, X_i(\hat{\theta})\big)$$

Next we will use the following result from Myerson 1981 [8] to compute the Bayesian incentive compatible set of transfers.

Proposition 1 (Myerson [8])**.** *The social choice function $f(.) = (k(.), t_1(.), \cdots, t_N(.))$ is Bayesian incentive compatible if and only if, for all $i = 1, \cdots, N$,*

- $\bar{v}_i(.)$ *is non-decreasing.*
- $U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s)ds \qquad$ *for all* θ_i.

However, in order to be able to use this result we need to ascertain that $v_i(.)$ and $T_i(.)$ are in fact (directly) independent of θ_i itself, i.e. they are independent of own true type but can depend on other players true type (as well as all the reported types).

Let us focus on $v_i(.)$, and then the exact same argument holds for $T_i(.)$ as well. If the ad of advertiser i is not chosen to be posted on the search page, then X_i is the all-zero vector, and so $v_i(.)$ is zero and independent of θ_i. To see why the above independency condition holds when slot j is allocated to advertiser i, i.e. $X_{ij} = 1$ and $\theta^{[j]} = \theta_i$,note that the $v_i(.)$ function depends on θ only through $r(\theta, \hat{\theta})$. Since each advertiser gets at most one position, $\theta^{[k]} \neq \theta_i$ for all $k \neq j$, and as a result $r_j = (1 - \theta^{[1]}) \cdots (1 - \theta^{[j-1]})$, i.e. the only entry of vector r which appears with a non-zero coefficient in $v_i(.)$, is independent of θ_i. Therefore, $v_i(.)$ is independent of θ_i as well.

The rest of the derivation is straight forward. One can further refine the utility function as the following:

$$u_i\big(\theta_{-i}, X(\hat{\theta})\big) = \theta_i v_i\big(\theta_{-i}, X(\hat{\theta})\big) + T_i\big(\theta_{-i}, X_i(\hat{\theta})\big)$$

i.e. player i utility can be written as a linear function of his own type, θ_i. We am interested in a truth-telling equilibrium, so it is appropriate to define $\bar{v}_i(\hat{\theta}_i)$ and $\bar{T}_i(\hat{\theta}_i)$ as the expected "benefit" and transfer of player i given that he announces his type to be $\hat{\theta}_i$ and that all players $j \neq i$ truthfully reveal their types.

$$\bar{v}_i(\hat{\theta}_i) = \mathbb{E}_{\theta_{-i}}\big[v_i\big(\theta_{-i}, X(\hat{\theta}_i, \theta_{-i})\big)\big] = \mathbb{E}_{\theta_{-i}}\big[v_i\big(\hat{\theta}_i, \theta_{-i}\big)\big]$$
$$\bar{T}_i(\hat{\theta}_i) = \mathbb{E}_{\theta_{-i}}\big[T_i\big(\theta_{-i}, X(\hat{\theta}_i, \theta_{-i})\big)\big] = \mathbb{E}_{\theta_{-i}}\big[T_i\big(\hat{\theta}_i, \theta_{-i}\big)\big]$$

Recall that the advertisers qualities are independent. Using our new notation, advertiser i's expected utility from social choice function $f(.)$ when his type is θ_i, he announces $\hat{\theta}_i$ and everyone else reports truthfully can be written as

$$\mathbb{E}_{\theta_{-i}}\big[u_i\big(f(\hat{\theta}_i, \theta_{-i}), \theta_i|\theta_i\big)\big] = \theta_i \bar{v}_i(\hat{\theta}_i) + \bar{T}_i(\hat{\theta}_i)$$

Moreover, let $U_i(\theta_i)$ denote advertiser i's expected utility from the mechanism conditional on his type being θ_i and everyone (including i) reporting truthfully:

$$U_i(\theta_i) = \theta_i \bar{v}_i(\theta_i) + \bar{T}_i(\theta_i)$$

Finally, using Proposition 1 the expected transfer function can be written as:

$$\bar{T}_i(\theta_i) = U_i(\underline{\theta}_i) - \theta_i \bar{v}_i(\theta_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s)ds$$

and per-click transfers $(t_1(\theta), \cdots, t_n(\theta))$ should be chosen such that

$$\mathbb{E}_{\theta_{-i}}\Big[X_i(\hat{\theta}) \centerdot \big(r(\hat{\theta}, \theta) \circ C\big) t_i(\hat{\theta})\Big] = \bar{T}_i(\theta).$$

Derivation of Search Engine Objective Function.

Consider the maximization problem 8. Feasibility of $X_i(.)$ means (1) and (2) are satisfied. Conditions (ii) and (iii) ensure that the allocation rule and transfer functions are Bayesian incentive compatible, and (iv) is the participation constraint, i.e. each advertiser's utility from participating in the auction must be higher than his utility from not doing so, which is zero.

Note that if constraint (iii) is satisfied, then (iv) will be satisfied if and only if $U_i(\underline{\theta}_i) \geq 0$ for all i. As a result, we can replace constraint (iv) with the following constraint:

$$\text{(v) } U_i(\underline{\theta}_i) \geq 0 \qquad \forall i$$

To solve the optimization problem, substitute into the objective function for $U_i(\theta_i)$ using constraint (iii) and consider only the payment from a single advertiser for the moment:

$$\begin{aligned}
\mathbb{E}_{\theta_i}\big[-\bar{T}_i(\theta_i)\big] &= \int_{\underline{\theta}_i}^{\bar{\theta}_i} \big(\theta_i \bar{v}_i(\theta_i) - U_i(\theta_i)\big)\phi_i(\theta_i)d\theta_i \\
&= \int_{\underline{\theta}_i}^{\bar{\theta}_i} \Big(\theta_i \bar{v}_i(\theta_i) - U_i(\underline{\theta}_i) - \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s)ds\Big)\phi(\theta_i)d\theta_i \\
&= \Bigg[\int_{\underline{\theta}_i}^{\bar{\theta}_i} \Big(\theta_i \bar{v}_i(\theta_i) - \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s)ds\Big)\phi_i(\theta_i)d\theta_i\Bigg] - U_i(\underline{\theta}_i) \qquad (17)
\end{aligned}$$

Use integration by parts to simplify:

$$\begin{aligned}
&\int_{\underline{\theta}_i}^{\bar{\theta}_i} \Big(\int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s)ds\Big)\phi_i(\theta_i)d\theta_i = \Big(\int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{v}_i(\theta_i)d\theta_i\Big) - \Big(\int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{v}_i(\theta_i)\Phi_i(\theta_i)d\theta_i\Big) \\
&= \int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{v}_i(\theta_i)(1 - \Phi_i(\theta_i))d\theta_i
\end{aligned}$$

and substitute in (17) to get:

$$\begin{aligned}
&\mathbb{E}_{\theta_i}\big[-\bar{T}_i(\theta_i)\big] = \Bigg[\int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{v}_i(\theta_i)\Big(\theta_i - \frac{1 - \Phi_i(\theta_i)}{\phi_i(\theta_i)}\Big)\phi_i(\theta_i)d\theta_i\Bigg] - U_i(\underline{\theta}_i) = \\
&\Bigg[\int_{\underline{\theta}_1}^{\bar{\theta}_1} \cdots \int_{\underline{\theta}_n}^{\bar{\theta}_N} v_i(\theta)\Big(\theta_i - \frac{1 - \Phi_i(\theta_i)}{\phi_i(\theta_i)}\Big)\Big(\prod_{j=1}^{N} \phi_j(\theta_j)\Big)d\theta_N \cdots d\theta_1\Bigg] - U_i(\underline{\theta}_i)
\end{aligned}$$

Substitute for $v_i(.)$ and add up the transfers from all the advertisers to get the auctioneer's total revenue:

$$\int_{\underline{\theta}_1}^{\bar{\theta_1}} \cdots \int_{\underline{\theta}_n}^{\bar{\theta_n}} \Bigg[\Big(\sum_{i=1}^{N} \big(X_i(\theta) \centerdot \big(r(\theta) \circ C\big)\big) \big(\theta_i - \frac{1 - \Phi_i(\theta_i)}{\phi_i(\theta_i)}\big) \Big) \Big(\prod_{j=1}^{N} \phi_j(\theta_j) \Big) d\theta_N \cdots d\theta_1 \Bigg] - \sum_{i=1}^{N} U_i(\underline{\theta}_i) \tag{18}$$

So the auctioneer chooses $X(.), U_1(\underline{\theta}_1), \cdots, U_N(\underline{\theta}_N)$ to maximize the above expression subject to constraints (i), (ii) and (v). For each advertiser's transfer, if $\theta_i = \underline{\theta}_i$ the first term of the transfer would be zero, so the auctioneer must set $U_i(\underline{\theta}_i) = 0$ for all $i = 1, \cdots, N$ to maximize the revenue. Hence, the auctioneer problem reduces to choosing $X(.)$ to maximize

$$\int_{\underline{\theta}_1}^{\bar{\theta_1}} \cdots \int_{\underline{\theta}_n}^{\bar{\theta_n}} \Bigg[\Big(\sum_{i=1}^{N} \big(X_i(\theta) \centerdot \big(r(\theta) \circ C\big)\big) \big(\theta_i - \frac{1 - \Phi_i(\theta_i)}{\phi_i(\theta_i)}\big) \Big) \Big(\prod_{j=1}^{N} \phi_j(\theta_j) \Big) d\theta_N \cdots d\theta_1 \Bigg]$$

subject to constraints (i) and (ii). As mentioned earlier, one interesting case is uniform distribution for both qualities and search costs where have $\Phi(\theta_i) = \theta_i$ and $\phi(\theta_i) = 1$. With these assumptions, the auctioneer's maximization problem further simplifies to

$$max_{X(.)} \int_0^1 \cdots \int_0^1 \Big(\sum_{i=1}^{N} \big(X_i(\theta) \centerdot \big(r(\theta) \circ C\big)\big) \big(2\theta_i - 1\big) \Big) d\theta_N \cdots d\theta_1 \tag{19}$$

8.2 Proofs

Lemma 1.

The $H(.)$ function in the two cases can be written as:

$$\begin{aligned} H_{k,k+1} = & \sum_{i=1}^{k-1} C_i \Big(\prod_{j=1}^{i-1} (1 - a_j) \Big) \big(2a_i - 1\big) \\ & + C_k \Big(\prod_{j=1}^{k-1} (1 - a_j) \Big) \big(2a_k - 1\big) + C_{k+1} \Big(\prod_{j=1}^{k} (1 - a_j) \Big) \big(2a_{k+1} - 1\big) \\ & + \sum_{i=k+2}^{m} C_i \Big(\prod_{j=1}^{i-1} (1 - a_j) \Big) \big(2a_i - 1\big) \end{aligned}$$

$$
\begin{aligned}
H_{k+1,k} = &\sum_{i=1}^{k-1} C_i\Big(\prod_{j=1}^{i-1}(1-a_j)\Big)\big(2a_i - 1\big) \\
&+ C_k\Big(\prod_{j=1}^{k-1}(1-a_j)\Big)\big(2a_{k+1}-1\big) + C_{k+1}\Big(\prod_{j=1}^{k-1}(1-a_j)\times(1-a_{k+1})\Big)\big(2a_k - 1\big) \\
&+ \sum_{i=k+2}^{m} C_i\Big(\prod_{j=1}^{i-1}(1-a_j)\Big)\big(2a_i - 1\big)
\end{aligned}
$$

It is clear that switching a_k and a_{k+1} leaves the first $k-1$ terms of the $H(.)$ function , as well as the last $n-k+1$ terms $(k+2, k+3, \cdots, n)$ intact. So we only need to compare the k^{th} and $k+1^{th}$ term:

$$
\begin{aligned}
H_{k+1,k} - H_{k,k+1} = &\Big(\prod_{j=1}^{k-1}(1-a_j)\Big)\Big[\big(C_k(2a_{k+1}-1) + C_{k+1}(1-a_{k+1})(2a_k - 1)\big) \\
&- \big(C_k(2a_k - 1) + C_{k+1}(1-a_k)(2a_{k+1}-1)\big)\Big] > 0
\end{aligned}
$$

It is enough to show

$$
\begin{aligned}
&\big(C_k(2a_{k+1}-1) + C_{k+1}(1-a_{k+1})(2a_k - 1)\big) - \\
&\big(C_k(2a_k - 1) + C_{k+1}(1-a_k)(2a_{k+1}-1)\big) \\
&= 2C_k(a_{k+1}-a_k) - C_{k+1}(a_{k+1}-a_k) = (2C_k - C_{k+1})(a_{k+1}-a_k) > 0
\end{aligned}
$$

The latter holds since both $C_k > C_{k+1}$ and $a_{k+1} > a_k$, which completes the proof.

For the case of general distribution $\Phi(\theta)$, the term $2a_i - 1$ is substituted by $V(a_i) = a_i - \frac{1-\Phi(a_i)}{\phi(a_i)}$, so we get:

$$
H_{k+1,k} - H_{k,k+1} = (C_k - C_{k+1})[V(a_{k+1}) - V(a_k)] + C_{k+1}[a_k V(a_{k+1}) - a_{k+1}V(a_k)]
$$

The first teem in the sum is positive, so a sufficient condition for $H_{k+1,k} - H_{k,k+1} > 0$ is

$$
\frac{V(a_{k+1})}{a_{k+1}} > \frac{V(a_k)}{a_k}
$$

□

Lemma 2.

With the same argument as in the previous lemma, switching these two entries only affect the last term in the $H(.)$ function, so it is enough to show that

$$
\Big(\prod_{j=1}^{m-1}(1-a_j)\Big)\Big[C_m(2a_x - 1) - C_m(2a_m - 1)\Big] > 0
$$

which is true since $a_x > a_m$. □

Lemma 3.

First note that since consumers believe that ads are sorted in descending order of qualities, $\mathbb{E}[\theta^{[i]}] = \mathbb{E}[\theta^i]$, and consequently the "unconditional" expectation of ad qualities decreases further down the list. But recall that $\bar{\theta}^{[i]}$ is a conditional mean, so we need to show:

$$\mathbb{E}[\theta^i | z^1 = \cdots = z^{i-1} = 0] \geq \mathbb{E}[\theta^{i+1} | z^1 = \cdots = z^{i-1} = z^i = 0]$$

Note that $\mathbb{E}[\theta^{i+1} | z^1 = \cdots = z^{i-1} = 0] \geq \mathbb{E}[\theta^{i+1} | z^1 = \cdots = z^{i-1} = z^i = 0]$. In addition, by definition of order statistic we have $\mathbb{E}[\theta^i | z^1 = \cdots = z^{i-1} = 0] \geq \mathbb{E}[\theta^{i+1} | z^1 = \cdots = z^{i-1} = 0]$, which establishes the result. □

Lemma 4.

First note that the search costs enter the advertisers utility and auctioneer revenue functions through vector C. The only step of the proof which uses properties of C is Lemma 1, which only uses the fact that for every i, $C_i > C_{i+1}$. Recall that C_i is the probability that consumer search cost is lower than his belief about the quality of the ad in position i condition on reaching this position, i.e. conditioned on the fact that all the higher ads failed to meet his need. Let $G(.)$ denote the cumulative distribution of consumer search costs. As a generalization to vector C defined in (4), the generalized vector C^G can be written as:

$$C^G = \big(G(\bar{\theta}_1), G(\bar{\theta}_2), \cdots, G(\bar{\theta}_N)\big)$$

cince by lemma 3, $\bar{\theta}^{[i]}$ is non-increasing for any distribution of θ_i, $C_i^G > C_{i+1}^G$ regardless of what distributions of search costs and qualities are.

□

Theorem 1.

With firm qualities independently drawn from distribution $\Phi(.)$, and an arbitrary distribution $G(.)$ for the search costs, the generalized C vector, and $J(.)$ and $H(.)$ functions defined by (10,11) will take the following form:

$$C^G = \big(G(\bar{\theta}_1), G(\bar{\theta}_2), \cdots, G(\bar{\theta}_N)\big)$$

$$J^G(X(.), \theta) = \sum_{i=1}^{N} \big(X_i(\theta) \cdot \big(r(\theta) \circ C^G\big)\big)\big(\theta_i - \frac{1 - \Phi(\theta_i)}{\phi(\theta_i)}\big)$$

$$H^G(a_1, \cdots, a_n; m) = \sum_{i=1}^{m} C_i^G \Big(\prod_{j=1}^{i-1} (1 - a_j)\Big)\big(a_i - \frac{1 - \Phi(a_i)}{\phi(a_i)}\big)$$

We need analogous arguments as those of Lemma 1 and 2 here to get the desired result. Since virtual valuation is increasing in θ_i, we know that if $\theta_i > \theta_j$, $V(\theta_i) > V(\theta_j)$. It is easy to see that with non-decreasing virtual valuations, Lemma 2 still holds in the general case. In order for Lemma 1 to hold here we should have that for $\theta_i > \theta_j$ which are in two consecutive position:

$$C_k^G V(\theta_i) + C_{k+1}^G (1 - \theta_i) V(\theta_j) \geq C_k^G V(\theta_j) + C_{k+1}^G (1 - \theta_j) V(\theta_i)$$

which along with lemma 3 and 4 gives the condition specified in the theorem. □

Lemma 5.

The second claim follows directly from the last two lemmas. Specifically, we can model the optimal mechanism as a two-step procedure of first deciding on the optimal number of ads (k) and then choosing which ads to post. Given k, using Lemma 3 and 4 it is always optimal for the auctioneer to post the k highest (reported) quality ads in decreasing order of quality. In other words for any choice of k the auctioneer always prefer to swap a candidate lower quality ad for a higher quality one to be presented to consumers. So for every bidder i, allocation rule is monotone in his reported type (quality).

Now, assume that the auctioneer uses a strategy which does not satisfy the criterion stated in the lemma, i.e. k depends on some l^{th}-order statistics of reported qualities where $M < l \leq N$. We will show that this rule can be replaced with another one which does not have such dependence and does equally well in expectation.

The new strategy is implemented through a black-box which receives all the reports and then pass them on to the auctioneer. The black-box works as following: It does not change the highest M reports and pass them to the auctioneer as they are. Let the M^{th} highest report be c, i.e. $\theta^{[M]} = c$. The black-box replaces each of $\theta^{[M+1]}, \cdots, \theta^{[N]}$ with a draw from its corresponding conditional distribution conditional on $\theta^{[M]} = c$.[8] With this transformation, the distribution of reports received by the auctioneer is the same as the distribution of reports reported by bidders. As a result, although the auctioneer decision rule generate different outcomes for each realization of reports with and without the black-box, but the distribution of outcomes remain unchanged (since input distribution remains unchanged) which means expected revenue remains unchanged. Since the $M + 1^{st}$-N^{th} reported qualities has in fact been replaced by the black-box, the new strategy does not depend on them and it satisfies the property explained in the lemma. □

Theorem2.

As mentioned in the text, given Lemma 5 when there is only one available position, only the first order statistic of the qualities matter for auctioneer decision whether or not to post an ad. With a little abuse of notation, let θ and $\phi_1(.)$ denote the first order statistics and

[8] Equivalently, it replaces the sub-vector $[\theta^{[M+1]}, \cdots, \theta^{[N]}]$ with a draw from its joint conditional distribution conditional on $\theta^{[M]} = c$.

its distribution respectively. The auctioneer maximization problem can be written as the following:

$$\max_{y(\theta)} \int_0^1 \mathbb{E}[\theta|1 \text{ ad is posted using } y(\theta)] V(\theta) y(\theta) \phi_1(\theta) d\theta$$
$$= \frac{\int_0^1 \theta y(\theta)\phi_1(\theta)d\theta}{\int_0^1 y(\theta)\phi_1(\theta)d\theta} \int_0^1 V(\theta) y(\theta)\phi_1(\theta)d\theta$$

where $V(\theta) = \left(\theta - \frac{1-\Phi(\theta)}{\phi(\theta)}\right)$. Note that the first fraction is the conditional expectation of the quality of the only posted ad if it is posted.

The same intuition as the one for the example in the text applies for general $\phi_1(\theta)$. The only twist is that in the general case, $K = \int_0^1 y(\theta)\phi_1(\theta)d\theta$ which is not invariant to reordering of $y(\theta)$. Note that since $0 \leq y(\theta) \leq 1$, $y(\theta)\phi(\theta) \leq \phi(\theta)$. Consider the following construction of $\hat{y}(.)$ where $\hat{y}(.)$ as large as possible for higher θ's keeping K fixed. Intuitively, let $\hat{y}(\theta) = 1$ starting from $\theta = 1$ and moving down to $\theta = 0$ until $K = \int_0^1 \hat{y}(\theta)\phi_1(\theta)d\theta$.

Define $g(\theta) = \frac{y(\theta)\phi_1(\theta)}{\int_0^1 y(\theta)\phi_1(\theta)d\theta}$ (and similarly $\hat{g}(\theta)$ for $\hat{y}(.)$). Note that $g(\theta)$and $\hat{g}(\theta)$ are PDFs themselves, and let $G(\theta)$ and $\hat{G}(\theta)$ denote their CDFs respectively. By construction, $\hat{G}$ FOSD G, so transforming $y(\theta)$ to $\hat{y}(\theta)$ increases both terms in the objective function. As a result, the optimal decision rule should take the same form as $haty(\theta)$, i.e. a step function.

Finally, assume $\theta^* < \tilde{\theta}$.[9] Construct the following $\bar{y}(.)$ from $y(.)$:

$$\bar{y}(\theta) = \begin{cases} 0 = y(\theta) & \text{if } \theta < \theta^* \\ 0 \neq y(\theta) & \text{if } \theta^* \leq \theta < \tilde{\theta} \\ 1 = y(\theta) & \text{if } \tilde{\theta} \leq \theta \end{cases}$$

Define $g(\theta)$and $\bar{g}(\theta)$ as above and note that $\frac{\bar{g}(\theta)}{g(\theta)}$ is increasing in θ, so it satisfies the monotone likelihood ratio property. Equivalently, $\bar{G}$ FOSD G, so substituting $\bar{y}(.)$ for $y(.)$ increases both the first and second term in the objective function, so $y(.)$ can not be the optimal solution, i.e. $\theta^* \geq \tilde{\theta}$ □

Theorem 3.

Assume in the original setting the auctioneer optimal allocation rule is such that some ad with quality $\theta < \tilde{\theta}$ will be posted in position j. If the auctioneer changes the rule such that ad i can possibly by posted only if $\theta_i > \tilde{\theta}$, there are 3 effects we have to consider: First is the effect of the allocation rule on expectation of consumers on quality of each posted ad, second the revenue generated by each particular ad and third the effect of each ad on the click-through-rate of all ads which are posted bellow that one.

[9] θ^* is where the jump in the optimal solution $y(\theta)$ happens, and $\tilde{\theta}$ is such tat $V(\tilde{\theta}) = 0$ as defined in the theorem.

First consider the last two effects: an ad with negative virtual valuation generates negative expected revenue, so the proposed modification to the existing rule will enhance the revenue through this channel. Moreover since by Lemma 5 ads are posted in descending order of quality, all ads in position $j+1, \cdots, k$ have negative virtual valuations as well. AS a result changing the rule as described above would cut these ads as well and increase the auctioneer revenue, and the third effect is irrelevant.

The most complicated effect is the first one, which is the feedback effect on consumer expectations. Recall that the consumer expectation of the ad posted in position l can be written as:

$$\bar{\theta}^{[l]} = \mathbb{E}[\theta^{[l]} | z^1 = \cdots = z^{l-1} = 0, k \text{ ads are posted}]$$

Where k *ads are posted* represents the auctioneer rule which maps the reported qualities to optimal number of ads. In other words, the expectation of l^{th} order-statistic is taken over its distribution conditional on the other order statistics being posted. Since the distribution of the l^{th} order-statistic conditional on higher values of j^{th} order-statistic FOSD the one conditional on lower values of j^{th} order-statistic, the proposed changed in the allocation rule can only improve the consumer expectations about the quality of each posted ad and enhance the revenue. As a result the modifies rule generates higher revenue and the original ad could not be optimal. □

Theorem 4.

The proof method is the same as in Theorem 1, i.e. we will show that any choice of ads that replaces a higher quality ad with a lower quality one is sub-optimal in terms of revenue. Throughout this proof, when we refer to the ad in position i (since physical position is irrelevant here), it means the ad which is i^{th} option of the consumer to click, i.e. the i^{th} highest-quality posted ad.

Note that if there are at most $k < M$ reported qualities higher than the threshold x^*, the auctioneer is obviously better off posting only those k highest-quality ad. The reason is that all the additional ads are going to be clicked after these highest k by the consumers, so they generate negative revenue themselves, and only affect the click-through rate of some other negative-revenue-generating ads (i.e. they have no effect on any positive revenue generated from this search page), so it is better for the auctioneer to just cut the list at position $k+1$ (i.e post k ads). The adverse effect of decreasing click-through-rate of positive revenue generating ads by cutting negative-revenue-generating ads is not present when the qualities are posted, or when consumers stop at ad j with some constant probability π.

So the only remaining case is when exactly k ads are posted. Assume that the revenue maximizing list is not composed of the M highest quality ads. In this case the ad in position k (with quality θ_j) is certainly not one of the k highest quality one, and at least one of the best k ads (with quality θ_i) is out of the list. We show that swapping θ_i and θ_j and keeping all the rest of the ads the same can only increase the revenue.

The main conceptual difference between this proof and that of 1 is that here, the intermediate steps can not be materialized. Recall that the intermediate steps were picking a high quality ad who is outside and swapping it with the ad currently in the last position, and then moving it up to its optimal position. Here, for instance, swapping θ_i and θ_j "in the last position" does not have an outside world realization unless θ_i is lower than the quality of all the other ads currently on the list. Otherwise, consumers will automatically click on θ_i before some of the other ads, so it will not be in position k any more. So this last position swap only corresponds to an intermediate number, which is higher than the revenue of the auctioneer with θ_j in the list, and lower than his revenue with θ_i in. If the auctioneer has the power to enforce the consumers to click on θ_i after all other ads, although it has a higher quality than some, the result of this last position swap would correspond to that state.

In order to get Lemma 1 and 2 to work here we need to show that for $\theta_i > \theta_j$:

(i) $\big[V(\theta_i)G(\theta_i) + (1-\theta_i)V(\theta_j)G(\theta_j)\big] - \big[V(\theta_j)G(\theta_j) + (1-\theta_j)V(\theta_i)G(\theta_i)\big] > 0$
(ii) $V(\theta_i)G(\theta_i) - V(\theta_j)G(\theta_j) > 0$

The second condition holds trivially when $\theta_i > \theta_j$, and in order for the first condition to hold independent of $G(.)$, we should have $\theta_j V(\theta_i) > \theta_i V(\theta_j)$, which means $I(.)$ is increasing.

For the $\bar{v}_i(.)$ function to be increasing, we need $G(\theta_i)V(\theta_i)$ to be increasing if $\theta_i > x^*$ (i.e if the ad has a chance of being on the list). Note that we have:

$$\big(G(\theta_i)V(\theta_i)\big)' = V'(\theta_i)G(\theta_i) + V(\theta_i)G'(\theta_i)$$

and we know that $V'(\theta_i) > 0$ and $G'(\theta_i), G(\theta_i) \geq 0$ (G' is the pdf of the search costs). In addition, only θ_i's with $G(\theta_i) > 0$ are interesting, because otherwise the ad will not be posted. So at any θ_i where $V(\theta_i) > 0$, i.e. for $\theta_i > x^*$, the above derivative is positive, and as a result the function $G(\theta_i)V(\theta_i)$ will be increasing, which is the desired result.

Note that here, we don't have two forces in opposite directions in determining the reservation wage. Here, vector C is determined by the posted qualities that consumers observe, so it does not depend on the number of posted ads, so the only force driving the reserve price is that the auctioneer does not want to post ads with negative virtual valuation, which gives the desired result. $\square$