Theory Comput Syst
DOI 10.1007/s00224-016-9679-z

Prediction and Welfare in Ad Auctions

Mukund Sundararajan[1] · Inbal Talgam-Cohen[2]

Abstract We study how standard auction objectives in sponsored search markets are affected by refinement in the prediction of ad relevance (click-through rates). As the prediction algorithm takes more features into account, its predictions become more refined; a natural question is whether this is desirable from the perspective of auction objectives. Our focus is on mechanisms that optimize for a convex combination of economic efficiency and revenue, and our starting point is the observation that the objective of such a mechanism can only improve with refined prediction, making refinement in the best interest of the search engine. We demonstrate that the impact of refinement on market efficiency is not always positive; nevertheless we are able to identify natural – and to some extent necessary – conditions under which refinement is guaranteed to also improve economic efficiency. Our main technical contribution is in explaining how refinement changes the ranking of advertisers by value (efficiency-optimal ranking), moving it either towards or away from their ranking by *virtual* value (revenue-optimal ranking). These results are closely related to the literature on signaling in auctions.

Keywords Mechanism design · Revenue optimization · Sponsored search auctions · Virtual values · Pareto frontier · Rearrangement inequality

✉ Inbal Talgam-Cohen
italgam@stanford.edu

[1] Google Inc., Mountain View, CA 94043, USA

[2] Computer Science Department, Stanford University, Stanford, CA 94305, USA

Published online: 02 May 2016

1 Introduction

Sponsored search is a multi-billion dollar market; it enables contextual advertising, and generates revenue that supports innovation in search algorithms. Sponsored search markets are also technically interesting and have been investigated theoretically from several perspectives [15], including auction theory [1, 7], game theory [6, 26], and bipartite matching theory [18].

Sponsored search markets exhibit an interesting interplay between auctions and machine learning. Value is realized by the combination of two processes. First, the search engine displays *relevant* ads to the user, i.e., ones that maximize the odds of the user clicking on an ad. Second, users conduct a transaction with some probability and of some value on the advertiser's website, resulting in some (expected) value per click to the advertiser. To facilitate the first process, the search engine uses a combination of machine learning and historical data to estimate the relevance of an ad to the user [12]. The second process is not directly observable by the search engine, and so it uses an auction to elicit the value per click from the advertisers [1, 7]. It then combines this with ad relevance to determine which ads to show to the user.

The recent explosion in data available to search engines makes it possible to improve relevance prediction by seemingly endless *refinements*, taking into account more and more *features* of the ad and the user. For example, consider a search query 'pizza' emanating from a user at an unspecified location within the Bay Area. By adding a feature that pinpoints the user's location within this region, it becomes clear whether advertisements showcasing San Francisco pizza merchants are more relevant than those for San Jose merchants or vice versa. This helps in deciding between these ads.

Refinement is often perceived as a positive, win-win opportunity making everyone better off – the users view more relevant ads and engage more with them, increasing overall value.[1] However, to our knowledge this has not been rigorously studied. *The focus of this paper is to explore how standard objectives of truthful auctions, specifically welfare, behave with refinement of relevance prediction*. We apply theory tools in order to understand the high-stake effects of refinement decisions carried out by sponsored search practitioners. We view this as a first step in better understanding the interaction between machine learning and market design objectives, and in particular the economic impact of more accurate learning (whether using more data as in this paper, or other improvements such as novel learning algorithms). We also discuss the connection to the signaling literature.

As our first contribution, we formalize the conventional wisdom that refinement aids optimization. While it holds generally that refinement improves the economic efficiency[2] of the efficiency-optimal mechanism, or the revenue of the revenue-optimal mechanism [10], we build upon the latter to establish this result for all truthful mechanisms that optimize some fixed convex combination of revenue and

[1]Clearly refinement should not be at the cost of using features that violate user privacy; in this work we leave aside issues of privacy to focus on welfare considerations of refinement.

[2]Throughout this paper, we use the term *efficiency* for economic rather than computational efficiency.

efficiency (*trade-off optimal* mechanisms). We discuss such mechanisms and justify why the search engine would be interested in a mechanism from this class in Section 3.[3] Thus, performing refinements always benefits the search engine.

What about the impact of refinement on social welfare? Our main contribution (Section 4) is to study conditions under which refinement is simultaneously favorable for the auctioneer and for market efficiency. Indeed, this is not always the case – the twin objectives of revenue and efficiency are not necessarily aligned in the context of refinement. We identify two assumptions under which refinement improves the efficiency of every trade-off optimal mechanism. The first assumption is fairly standard, and requires that the value-per-click distributions are i.i.d. and satisfy the monotone hazard rate condition. The second assumption is arguably more restrictive, requiring that refinements *distinguish* among advertisers, by causing the relevances of every pair of advertisers for every query to either grow further apart or switch order. We demonstrate the need for both assumptions by two examples (see Sections 2.2 and 5).

From a technical perspective, a main challenge is in understanding the mathematical effect of refinement. The revenue-optimal auction and the efficiency-optimal auction both rank advertisers by a monotone function of their bids and then use this ranking to allocate them into available ad slots. The key difference is that the two mechanisms employ different ranking functions to the bids. Refinement reduces efficiency precisely when it causes the revenue-optimal ranking to drift apart from the efficiency-optimal one. Under the assumptions mentioned above, the ranking of every trade-off optimal mechanism is guaranteed to draw closer with refinement to the efficiency-optimal ranking.

2 Model

In this section we present our model, which encompasses the standard model for position auctions [15], while capturing the effect of prediction refinement.

A search engine sells m ad slots to $n \leq m$ advertisers (also known as bidders).[4] The slots appear alongside search results for a search query q. Advertiser i has a private value $v_i \in \mathbb{R}_+$ for a *click* on his ad, and his value for an *impression* (appearance of the ad) is v_i multiplied by the corresponding *click-through rate*. This multiplicative relation is often assumed in the literature [15] and is an important feature of the model. Another standard assumption is that click-through rates are *separable*, i.e., can be multiplicatively separated into the advertiser's *relevance* to query q, and the

[3]The mechanisms used in practice, though not truthful, have equilibria that are allocation- and revenue-equivalent to the corresponding truthful mechanisms [6, 7]. Thus, we expect the gist our results to apply to practically used mechanisms in equilibrium. This raises an interesting open problem: As we show, refinement changes advertiser ranking in non-trivial ways; how do the equilibrium bids of the advertisers change in response? Will their level of granularity mirror that of the refinement? In other words, how does personalization affect the analysis of [6, 7]? The answer will depend on the informational assumptions of the model.

[4]The assumption that $m \geq n$ is without loss of generality. Advertisers/bidders are not to be confused with users, who are the ones submitting queries and not part of the auction.

effect of the slot position on the webpage. Note that if the click-through rates are 1 we get a standard m-unit auction.

Formally, the click-through rate for advertiser i's ad in slot j given query q is $p_{q,i}s_j$, where $1 \geq p_{q,i} > 0$ is the query-advertiser relevance,[5] and $1 \geq s_1 \geq \cdots \geq s_m \geq 0$ are the slot effects. (We omit q from the notation where clear from the context.) The relevance $p_i = p_{q,i}$ can thus be thought of as the slot-independent click-through rate. We denote the value per impression in slot j by

$$v_{i,j} = p_i v_i s_j,$$

and the value per impression without the slot effect, called the *realized value*, by

$$r_i = p_i v_i.$$

The advertisers' private values are assumed to be independently distributed according to a publicly-known distribution F with positive smooth density f. Note that the realized values are not i.i.d. and so the setting is *not* symmetric.

2.1 Prediction Schemes and Refinements

The machine learning system that predicts query-advertiser relevance has access to a set of features: keywords, geographic location, time, user demographics, search history, ad text, etc. As is standard we assume features are discretized [12]. The system partitions the set of query-advertiser pairs according to the features and produces a relevance estimate for each part. For example, a part can consist of pizzeria advertisers together with queries for "pizza" by users located in the Bay Area. We refer to the output of the machine learning system as a *prediction scheme*:

Definition 1 (Prediction scheme) A partition T of all query-advertiser pairs with a relevance prediction p_t for every part $t \in T$.[6]

Overloading notation we also denote the prediction scheme itself by T. The prediction given a search query q is *according to T* if for every advertiser i, $p_{q,i} = p_t$ where t is the part in T containing the query-advertiser pair (q, i).

Refinements A prediction scheme can be *refined* by refining its partition, i.e., dividing coarse parts into finer *subparts*. This can be achieved by taking into account additional features, such as more precise user location. For example, a subpart may consist of pizzeria advertisers together with queries for "pizza" by users located in a *specific city* within the Bay Area. We use the notational convention that $\bar{T}$ is a coarse partition and $\bar{t}$ a coarse part, whereas T is a refined partition and t is a subpart.

The relevence of a subpart can be very different from that of the original coarse part, and for this reason refinement can completely alter the outcome of the ad

[5] The assumption that $p_{q,i} \neq 0$ is without loss, to simplify the exposition.

[6] This definition matches that of Ghosh et al.'s deterministic clustering scheme [11]. In general a prediction scheme can be randomized, by including a distribution over relevance predictions for each part (cf. [8, 20]). Our results hold for randomized prediction schemes as well.

auction. However, the coarse and refined relevance predictions must maintain the following relation. Given a query-advertiser pair belonging to a coarse part $\bar{t}$, there is a certain distribution with which it falls within its different subparts. We require that in expectation over this distribution, the refined relevance prediction equals the coarse one. To summarize:

Definition 2 (Refinement) A prediction scheme T is a *refinement* of $\bar{T}$ if its partition is a refinement of $\bar{T}$'s partition, and the relevance of every coarse part $\bar{t}$ equals the expected relevance over $\bar{t}$'s subparts: $p_{\bar{t}} = \mathbb{E}_{t\subseteq\bar{t}}[p_t]$.

(If the subpart t and its coarse counterpart $\bar{t}$ are clear from context, we use p and $\bar{p}$ to denote their relevance predictions.)

Distinguishing Refinements A natural subclass of refinements is those which distinguish among advertisers, thus enabling a better matching between them and the search queries. We begin with a technical definition:

Definition 3 (Spread or flipped pairs) A pair of numbers a, b is *spread or flipped* with respect to another pair c, d if

$$\frac{a}{b} \geq \frac{c}{d} \geq 1, \text{ or } 1 \geq \frac{c}{d} \geq \frac{a}{b}, \text{ or } \frac{a}{b} \geq 1 \geq \frac{c}{d}, \text{ or } \frac{c}{d} \geq 1 \geq \frac{a}{b}.$$

Definition 4 (Distinguishing refinement) A prediction scheme T is a *distinguishing refinement* of $\bar{T}$ if T is a refinement of $\bar{T}$, and for every query q and pair of advertisers, their relevance p_1, p_2 according to T is spread or flipped with respect to their relevance $\bar{p}_1, \bar{p}_2$ according to $\bar{T}$ (Fig. 1).

Intuitively, by flipping/spreading the relevance scores, distinguishing refinements expose which advertiser is a better fit for the query, thus contributing to social welfare. On the other hand, non-distinguishing refinements drive relevance scores closer, thus revealing which advertisers are both a reasonable fit and thus competitors for the query. This competition is then exploited to increase revenue at the expense of social welfare, by sometimes allocating to the less relevant advertiser in order to extract higher payment from the leading advertiser.

Remark 1 Our model is compatible with the standard assumption in theoretical study of sponsored search auctions, by which click-through rates are estimated accurately. It does not take into account that very fine prediction schemes may be inaccurate due to the emergence of over-thin submarkets with insufficient data. This simplifying assumption helps our goal of studying how finer prediction schemes affect auction

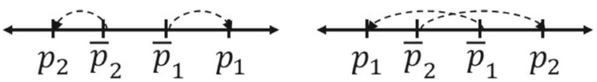

Fig. 1 An example of relevance pair p_1, p_2 spread (*left*) or flipped (*right*) with respect to $\bar{p}_1, \bar{p}_2$

objectives, by distilling this aspect of prediction refinement from various machine learning and other considerations. We expect our conclusions to apply in reality at least to *head* and *mid* search queries. Whether they apply to *tail* queries where data sparsity issues might arise is an open question to be settled by empirical research.

2.2 Examples

Example 1 (Every refinement is distinguishing) If all advertisers competing for any query q belong to the same part in $\bar{T}$ and so appear equally relevant, then every refinement of $\bar{T}$ is distinguishing.

As a concrete example, focus for simplicity on a single query. Consider the auction described in the introduction in which two pizzerias – the first located in San Francisco (SF) and the second in San Jose (SJ) – compete for a single advertisement slot next to search results for a query ‘pizza’ by a Bay Area user. Let $\bar{T}$ be a coarse prediction scheme in which both pizzerias appear equally relevant, i.e., both query-advertiser pairs belong to the same part $\bar{t} \in \bar{T}$. Let the corresponding relevance be $\bar{p}_{\bar{t}} = 3/4$. Now assume the search engine has access to a more precise location feature of the query q, indicating whether the user is in SF or in SJ, and each occurs with equal probability $1/2$. When the prediction scheme is refined by including this feature, the relevances according to the refined scheme T behave antisymmetrically, and the realized values are:

	User from SF	User from SJ	City unknown
Advertiser 1 (from SF)	$p_{\mathrm{SF},1}v_1 = v_1$	$p_{\mathrm{SJ},1}v_1 = v_1/2$	$\bar{p}_1 v_1 = 3v_1/4$
Advertiser 2 (from SJ)	$p_{\mathrm{SF},2}v_2 = v_2/2$	$p_{\mathrm{SJ},2}v_2 = v_2$	$\bar{p}_2 v_2 = 3v_2/4$

In both cases it can be observed that the refined relevances are either spread or flipped with respect to $\bar{p}_1, \bar{p}_2$.

Example 2 [A non-distinguishing refinement] Consider again a single-slot position auction for a query ‘pizza’. Assume now that advertiser 1 is a nationwide chain of pizzerias whose relevance does not depend on user location, while advertiser 2 is a local artisan pizzeria in SF. Consider a coarse prediction scheme $\bar{T}$ as above, and a refinement T where this time the refining feature indicates whether $q = \mathrm{SF}$ (happens with probability $1/4 - \delta$ for $\delta = 15\epsilon/(8 - 20\epsilon)$ and some small ϵ) or $q = \neg\mathrm{SF}$ (happens with probability $3/4 + \delta$). The realized values are:

	User from SF	User not from SF	City unknown
Advertiser 1 (chain)	$p_{\mathrm{SF},1}v_1 = 4v_1/5$	$p_{\neg\mathrm{SF},1}v_1 = 4v_1/5$	$\bar{p}_1 v_1 = 4v_1/5$
Advertiser 2 (from SF)	$p_{\mathrm{SF},2}v_2 = 2v_2/5$	$p_{\neg\mathrm{SF},2}v_2 = \epsilon v_2$	$\bar{p}_2 v_2 = v_2/10$

Refinement T is not distinguishing, since the relevances for $q = \text{SF}$ are neither spread nor flipped with respect to $\bar{p}_1, \bar{p}_2$.

3 Trade-off Optimal Mechanisms

3.1 Bayesian Mechanism Design

Section 3.1 contains mechanism design preliminaries in the ad auction context; the expert reader may wish to skip to Section 3.2.

Virtual Values Given a private value $v_i \sim F$, the *inverse hazard rate* is $\lambda^F(v_i) = (1 - F(v_i))/f(v_i)$ and the *virtual value* is $\varphi^F(v_i) = v_i - \lambda^F(v_i)$. Similarly, given a realized value r_i, let G with density g be the distribution from which r_i is drawn; the *realized virtual value* is then $\varphi^G(r_i) = r_i - \lambda^G(r_i)$. Since $r_i = p_i v_i$,

$$G(r_i) = F(v_i),\ g(r_i) = \frac{1}{p_i} f(v_i),\ \varphi^G(r_i) = p_i \varphi^F(v_i), \tag{1}$$

where the expression for $g(r_i)$ follows by deriving $F(\frac{r_i}{p_i})$. From now on, we omit the distribution and value from the notation where clear from context, and follow the convention that $\varphi(v_i)$ or φ_i is the virtual value, and $\varphi(r_i)$ is the realized virtual value.

A distribution F is *MHR* (*montone hazard rate*) if its inverse hazard rate function $\lambda(\cdot)$ is weakly decreasing,[7] and *regular* if its virtual value function $\varphi(\cdot)$ is weakly increasing. In other words, F is MHR if for every pair of values $v_1, v_2 \sim F$ such that $v_1 > v_2$, their inverse hazard rates λ_1, λ_2 are flipped: $v_1/v_2 > 1 \geq \lambda_1/\lambda_2$. It immediately follows that their virtual values φ_1, φ_2 are spread with respect to v_1, v_2:

$$1 < v_1/v_2 < \varphi_1/\varphi_2. \tag{2}$$

We say that values are MHR (resp., regular) if they're drawn from an MHR distribution, and that a position auction is MHR if its advertisers' values are MHR. By (1), MHR (resp., regular) values imply MHR realized values.

Efficiency-Optimal and Revenue-Optimal Mechanisms A mechanism is composed of an allocation rule, which matches bidders to slots (possibly randomly), and a pricing rule, which charges them for their allocation. Let $x_{i,j}(\mathbf{v})$ denote the probability with which bidder i wins slot j given a reported value (bid) profile $\mathbf{v} = (v_1, \ldots, v_n)$. Note that we focus on deterministic mechanisms in which $x_{i,j}(\mathbf{v})$ is an indicator $\in \{0, 1\}$. The *efficiency* (also known as *welfare*) of a mechanism with this allocation rule is

$$\sum_{i,j} x_{i,j}(\mathbf{v}) v_{i,j} = \sum_{i,j} x_{i,j}(\mathbf{v}) r_i s_j. \tag{3}$$

[7] The assumption of MHR values is common in the mechanism design literature (see, e.g., [17]). Many often-studied distributions are MHR, including the uniform, exponential and normal distributions, and those with log-concave densities [9].

A mechanism is (dominant-strategy) *truthful* if for every bidder i with true value v_i, alternative value report v_i' and reported value profile $\mathbf{v}_{-i}$ of other bidders, the utility (value minus payment) of bidder i for reporting his true value v_i is at least his utility for reporting v_i'. A mechanism is (ex post) individually rational (*IR*) if for every truthful bidder i and reported value profile $\mathbf{v}$, the utility of bidder i is non-negative. In what follows we do not distinguish between values and reported values (bids) since we focus on truthful IR mechanisms.

We say that an allocation rule $x_{i,j}(\mathbf{v})$ is *monotone* if as v_i increases, bidder i is only allocated higher slots. The following lemma is an adaptation of results by Myerson to the sponsored search context.

Lemma 1 (Myerson [21])

1. *Every monotone allocation rule can be coupled with a unique threshold pricing rule such that the resulting mechanism is truthful and IR.*[8]
2. *The expected revenue of every truthful and IR mechanism is equal to its expected realized virtual surplus, i.e.,*

$$\mathbb{E}_{\mathbf{v}}\left[\sum_{i,j} x_{i,j}(\mathbf{v})\varphi_i(r_i)s_j\right]. \tag{4}$$

The well-known Vickrey-Clarke-Groves (*VCG*) auction maximizes efficiency for every value profile [3, 13, 27]. In the context of position auctions, VCG allocates the slots to the n advertisers with highest realized values r_i, in high to low order [1]. We assume throughout that ties are broken lexicographically. Observe that this allocation rule is monotone, and so by the first part of Lemma 1, with appropriate payments we get a truthful and IR mechanism. The VCG allocation rule maximizes the efficiency as appears in (3) by the standard rearrangement inequality (see Lemma 2 below), applied to the two weakly decreasing vectors of sorted realized values $\mathbf{r}$ and slot effects $\mathbf{s}$.

The Myerson mechanism maximizes *expected* revenue among all truthful and IR mechanisms ([21], cf. [7, 14]).[9] When values are regular, the Myerson mechanism allocates slots to the $\leq n$ advertisers with highest *non-negative* realized *virtual* values $\varphi_i(r_i)$, in high to low order. By regularity this allocation rule is monotone, and so by the first part of Lemma 1, with appropriate payments we get a truthful and IR mechanism. The Myerson allocation rule maximizes the realized virtual surplus – and thus by the second part of Lemma 1 the expected revenue – by the standard rearrangement inequality applied to the two weakly decreasing vectors of sorted realized virtual values $(\varphi_i(r_i))_i$ and slot effects $\mathbf{s}$.

For completeness we now state (a version of) the rearrangement inequality. We say that π is a *partial ranking* of an n-element vector $\mathbf{x}$ if it ranks a subset of $n' \leq n$

[8] For our purpose we need not specify the pricing rule, because the second part of this lemma gives us a handle on revenue even without knowing the precise price form.

[9] In fact, it maximizes expected revenue among a larger class of mechanisms – *Bayesian* truthful and IR mechanisms.

elements, which we refer to as *acceptable*. We denote by $\mathbf{x}(\pi)$ a vector of length n in which the first n' entries are the acceptable elements ranked by π, and the rest are zero entries. We denote by $\mathbf{x}^+$ a vector in which the negative entries of $\mathbf{x}$ have been replaced by zeros.

Lemma 2 (Rearrangement inequality) *Let* $\mathbf{s} \geq 0, \mathbf{x}$ *be two weakly decreasing vectors. For every partial ranking* π *of* $\mathbf{x}$ *such that* $\mathbf{x}(\pi) \geq 0$ *it holds that* $\mathbf{s} \cdot \mathbf{x}(\pi) \leq \mathbf{s} \cdot \mathbf{x}^+$.

3.2 Trade-off Optimality

We now define a class of virtual value based mechanisms, of which the VCG auction and the Myerson mechanism are extremal members. We apply a result of Myerson and Satterthwaite to the sponsored search context, showing that mechanisms in this class optimize any efficiency-revenue trade-off [16, 22]. Such mechanisms are termed *trade-off optimal*, and their outcomes lie on the Pareto frontier of efficiency and revenue.[10]

Our interest in trade-off optimal mechanisms stems from our belief that search engines aim to optimize some convex combination of efficiency and revenue. While commercial search engines are "revenue maximizers", the "revenue" they aim to maximize is not just the short-term revenue referred to in this paper; rather, they care about a combination of revenue and efficiency, due to their interest in the long-term health and efficiency of sponsored search markets. As discussed in [16] in the context of multi-unit auctions, trade-off optimal mechanisms can also be used to to maximize expected welfare subject to a minimum constraint on the expected revenue.

Definition 5 [α-virtual value] For $\alpha \geq 0$, the α*-virtual value* of $v \sim F$ is $\varphi^\alpha(v) = v - \alpha\lambda^F(v)$ (where λ^F is the inverse hazard rate).

The α-virtual value of v can be rewritten as a combination of v and its corresponding virtual value: $\varphi^\alpha(v) = (1-\alpha)v + \alpha\varphi(v)$. The following definition encompasses the VCG auction ($\alpha = 0$) as well as the Myerson mechanism ($\alpha = 1$).

Definition 6 (α-virtual value based mechanism) For $\alpha \geq 0$, the α*-virtual value based mechanism* is a deterministic mechanism which asks the advertisers to report their values v_i, ranks them according to their realized α-virtual values $p_i\varphi_i^\alpha$, and allocates the slots to those ranked highest with non-negative $p_i\varphi_i^\alpha$. The allocation rule of the mechanism is coupled with the threshold pricing rule.

[10]Mechanisms on the efficiency-revenue Pareto frontier are not to be confused with mechanisms that generate Pareto optimal outcomes, in which no bidder's utility can be increased without decreasing another's. Diakonikolas et al. study computational complexity aspects of the Pareto frontier; the difference between their work and ours is that we focus on trade-off optimal mechanisms, which are not required to realize *every* point on the Pareto optimal curve.

Lemma 3 (Truthfulness) *For* $0 \le \alpha \le 1$ *and regular values, the* α*-virtual value based mechanism is truthful and IR.*

Proof By regularity, $\varphi^\alpha = (1-\alpha)v + \alpha\varphi(v)$ is weakly increasing in v when $0 \le \alpha \le 1$, thus the allocation rule is monotone; truthfulness and IR follow from Lemma 1. □

Lemma 4 (Trade-off optimal mechanisms) *Consider a regular position auction.*[11] *For* $0 \le \alpha \le 1$*, the optimal mechanism for the objective*

$$(1-\alpha)\mathbb{E}[\mathit{efficiency}] + \alpha\mathbb{E}[\mathit{revenue}] \tag{5}$$

among all truthful and IR mechanisms is the α*-virtual value based mechanism.*

Proof From Myerson's results applied to sponsored search (Lemma 1), it follows that the optimal mechanism for the objective maximizes the realized α-virtual surplus. The standard rearrangement inequality (Lemma 2) ensures that the optimal mechanism is the α-virtual value based mechanism, which is truthful and IR (Lemma 3).

In more detail, the optimal mechanism for the objective in (5) must maximize the following expression, obtained by applying the second part of Lemma 1 to get an expression for the expected revenue:

$$\begin{aligned}(1-\alpha)\mathbb{E}_{\mathbf{v}}[\sum_{i,j} v_{i,j}x_{i,j}(\mathbf{v})] + \alpha\mathbb{E}_{\mathbf{v}}[\sum_{i,j} s_j\varphi_i(r_i)x_{i,j}(\mathbf{v})] \\ = \mathbb{E}_{\mathbf{v}}[\sum_{i,j} x_{i,j}(\mathbf{v}) \cdot \big((1-\alpha)s_jr_i + \alpha s_j\varphi_i(r_i)\big)]\end{aligned}$$

Therefore, by the rearrangement inequality shown in Lemma 2, the optimal allocation rule picks up to m bidders with highest non-negative combinations $(1-\alpha)r_i + \alpha\varphi_i(r_i)$, and assigns them one by one to the highest slots. This is precisely the α-virtual value based mechanism, which is guaranteed to be truthful by Lemma 3. □

4 Refinement Effects on Auction Objectives

Our starting point is an observation regarding the search engine's incentive to perform refinement. Recall from Section 3.2 that we assume the search engine aims to optimize a fixed trade-off between revenue and efficiency. We observe that in expectation, refinement helps this objective, thus generalizing a previous result of

[11] A similar result holds for irregular position auctions, by replacing realized α-virtual values with their ironed counterparts.

Fu et al. beyond revenue maximization [10].[12] This indicates that up to practical limitations, the search engine would prefer as refined a prediction scheme as possible.

Before we present the formal statement of this result, we argue that it has merit beyond the intuitive fact that more information (in the form of refinement) can only help since it can always be discarded. This intuitive argument is not quite sufficient to prove the statement in Lemma 5 below, because we are considering a particular mechanism – the trade-off optimal mechanism – which was not explicitly designed to optimize the use of available information. The trade-off optimal mechanism uses given distributions to compute realized α-virtual values (see Definition 5), and it is not clear a priori whether the performance of this mechanism would benefit or not from using refined/coarse distributions. The result in [10] and its generalization in Lemma 5 show that *on average* over the extra information, using more information and thus refined distributions is better for the objective.

Lemma 5 (Refinement improves trade-off) *Let prediction scheme T be a refinement of scheme $\bar{T}$, and let q be a query belonging to a coarse part $\bar{t} \in \bar{T}$. Then with respect to its objective, a trade-off optimal mechanism M performs as well for q with scheme T as with $\bar{T}$, in expectation over value profiles and over the refined part $t \in T$ to which q belongs.*

Proof The proof follows from the fact that the trade-off optimal mechanism M is α-virtual value based (Lemma 4), combined with the proof of Proposition 3 of Fu et al. [10] in which, *mutatis mutandis*, the notion of value is replaced with that of α-virtual value. □

We now turn to the effect of refined relevance prediction on the efficiency guarantees of trade-off optimal mechanisms. In our main technical result, we identify natural conditions under which refining the prediction improves the efficiency of any trade-off optimal mechanism. The proof appears in Section 4.1.

Theorem 1 (Refinement improves efficiency) *Let prediction scheme T be a distinguishing refinement of scheme $\bar{T}$. Consider a set of bidders whose values are i.i.d. and satisfy MHR, and a position auction for a query q. Then with respect to social efficiency, a trade-off optimal mechanism M performs as well for q with scheme T as with $\bar{T}$, for every value profile of the advertisers.*

It is instructive to compare the two above results. Lemma 5 is less conditional, that is, the conditions of i.i.d., MHR values and distinguishing refinement are not required. On the other hand, Theorem 1 holds entirely *pointwise*, that is, it does not require averaging over the value profiles or query types. The fact that Theorem 1 requires more conditions, whose necessity is discussed in Section 5 by analyzing Examples 1 and 2, indicates a non-trivial trade-off between efficiency and revenue:

[12] Note however that the result of Fu et al. [10] applies to completely general signals whereas we focus on the linear form standard in the context of sponsored search.

When the search engine is optimizing for a combination of efficiency and revenue, refining "ad infinitum" will not always be the right thing to do in terms of social efficiency. This can be the case, for example, if the refinement is indistinguishing. On the flip side, when the conditions of Theorem 1 hold, the social interest is aligned with that of the search engine; prediction refinement is in both their best interests since it simultaneously increases social efficiency and its combination with revenue. This is formalized in Corollary 1, which is a direct consequence of Lemma 5 and Theorem 1.

Corollary 1 *Let prediction scheme T be a distinguishing refinement of scheme $\bar{T}$, and let q be a query belonging to part $\bar{t} \in \bar{T}$. Consider a set of bidders whose values are i.i.d. and satisfy MHR, and a position auction for q. Then with respect to both its objective and social efficiency, a trade-off optimal mechanism M performs as well for q with scheme T as with $\bar{T}$, in expectation over value profiles and the part $t \in T$ to which q belongs.*

In particular, mechanism M in Corollary 1 can be the revenue-optimal Myerson mechanism, for which a distinguishing refinement improves both efficiency and revenue. It is an interesting question whether there are additional mechanisms for which this desirable property of simultaneous improvement holds.

What is the importance of simultaneous improvement? Understanding the alignment between the objectives of welfare and revenue has attracted much attention in the theoretical literature (see, e.g., [2, 4, 5, 16, 23, 25]). This is grounded in the understanding that auctioneers aim to optimize their short-term goals without compromising their contribution to welfare (achieved, in this case, by delivering the most relevant set of ads to the users). Corollary 1 shows when refinement is in line with this two-fold aim.

Application to Signaling The above results are closely related to *signaling* of seller information in auctions, studied in the economic literature since the seminal work of Milgrom and Weber [19], and more recently in the computer science literature starting with [8, 11, 20]. The seller can adopt a *signaling scheme* by which he communicates his information to the bidders, who adjust their realized values accordingly. In the sponsored search context, the features which determine advertiser relevance can be viewed as the seller's information, making it a special case in which the effect of the information on values is multiplicative, and refinement is equivalent to revealing more of the seller's information. To our knowledge, this mathematical equivalence between prediction and signaling schemes has not been previously observed.

One difference between prediction and signaling is in who updates the values according to the information. In prediction, the seller internalizes this process, combining the information with the bidders' reports to get realized values. Thus refinement must affect values in a way that is completely known to the seller; this is not necessarily the case in the classic signaling context.

Our results apply to settings to which the fundamental *Linkage Principle* does not, due to the inherent asymmetry of advertiser relevance (indeed, it is not hard to see that refinement may decrease the expected revenue of mechanisms such as the second-price auction). Our main result stated in the context of signaling is that if releasing information distinguishes among i.i.d. MHR bidders, then it improves both the expected outcome of a Pareto optimal mechanism and its efficiency.

4.1 Proof of Theorem 1

Refinement has a delicate effect in the context of ad auctions. The transformation of values to realized values using different relevance terms causes the revenue-optimal ranking to differ from the efficieny-optimal one, even under assumptions of i.i.d. and MHR. This is in contrast to simple single-item multi-unit settings, where the revenue- and efficiency-optimal rankings both order bidders in the same way, and differ only in that the former excludes bidders with negative virtual values. In this section we show that the difference between the two rankings in sponsored search diminishes with refinement, as long as the conditions stated in Theorem 1 hold. In fact we show this for any trade-off optimal ranking according to α-virtual values, in addition to the revenue-optimal one where $\alpha = 1$.

Before proving Theorem 1 we present two lemmas. Throughout, fix a query q and let prediction scheme T be a distinguishing refinement of a scheme $\bar{T}$. The first lemma shows that if according to T, advertiser 1 has lower realized value but higher realized α-virtual value in comparison to advertiser 2, then the same holds according to $\bar{T}$. This indicates that any inefficiency due to the trade-off optimal ranking according to T occurs according to $\bar{T}$ as well, and so refinement can only increase efficiency.

Lemma 6 (Inefficient allocation) *Consider two advertisers with i.i.d. MHR values $v_1 \neq v_2$, and α-virtual values $\varphi_1^\alpha, \varphi_2^\alpha$. Let p_1, p_2 be their relevance predictions according to T, and $\bar{p}_1, \bar{p}_2$ their relevance predictions according to $\bar{T}$. Then*

$$p_1 < p_2 \text{ and } p_1\varphi_1^\alpha \geq p_2\varphi_2^\alpha > 0 \implies \bar{p}_1\varphi_1^\alpha \geq \bar{p}_2\varphi_2^\alpha.$$

Proof First observe that by the i.i.d. MHR assumption, the α-virtual values $\varphi_1^\alpha, \varphi_2^\alpha$ are spread with respect to the values v_1, v_2 (cf. (2)). We rewrite the two inequalities on the left-hand side as:

$$\varphi_1^\alpha/\varphi_2^\alpha \geq p_2/p_1 > v_1/v_2 > 1, \tag{6}$$

where the last inequality follows since otherwise the α-virtual values would not be spread with respect to the values. Since T is a distinguishing refinement, the pair p_1, p_2 is spread or flipped with respect to $\bar{p}_1, \bar{p}_2$, and so (recall Definition 3)

$$p_2/p_1 > 1 \implies p_2/p_1 \geq \bar{p}_2/\bar{p}_1. \tag{7}$$

Equations (6) and (7) combined show that $\bar{p}_1\varphi_1^\alpha \geq \bar{p}_2\varphi_2^\alpha$, completing the proof. □

The second lemma is a generalization of the standard rearrangement inequality (as stated in Lemma 2). Let π_1, π_2 be two partial rankings of an n-element vector $\mathbf{x}$. We say π_1 is *more ordered* than π_2 if: (1) the same elements are acceptable in both; and

(2) for every pair of acceptable elements $x_i > x_j$ such that x_i appears before x_j in $\mathbf{x}(\pi_2)$, this pair also appears in the correct order in $\mathbf{x}(\pi_1)$.

Lemma 7 (Generalized rearrangement inequality) *Let* $\mathbf{s} \geq 0, \mathbf{x}$ *be two weakly decreasing vectors, and let* π_1, π_2 *be two partial rankings of* $\mathbf{x}$ *such that* π_1 *is more ordered than* π_2*, and* $\mathbf{x}(\pi_1), \mathbf{x}(\pi_2) \geq 0$*. Then* $\mathbf{s} \cdot \mathbf{x}(\pi_1) \geq \mathbf{s} \cdot \mathbf{x}(\pi_2)$*.*

Proof We first prove the lemma assuming that π_1, π_2 are identical except for two acceptable elements $x_i > x_j$, which appear consecutively in both $\mathbf{x}(\pi_1)$ and $\mathbf{x}(\pi_2)$, but in flipped order. Note that since π_1 is more ordered, x_i must appear before x_j in $\mathbf{x}(\pi_1)$.

Using $\pi(x_i)$ to denote the rank of element x_i according to π, if $k = \pi_1(x_i)$, then $\pi_1(x_j) = \pi_2(x_i) = k + 1$ and $\pi_2(x_j) = k$. For rankings as above, to prove that $\mathbf{s} \cdot \mathbf{x}(\pi_1) \geq \mathbf{s} \cdot \mathbf{x}(\pi_2)$ it's sufficient to show that $s_k x_i + s_{k+1} x_j \geq s_k x_j + s_{k+1} x_i$. Since $x_i > x_j \geq 0$ and $\mathbf{s}$ is decreasing, this holds by the standard rearrangement inequality (Lemma 2).

We now turn to general partial rankings π_1, π_2 where π_1 is more ordered, and conceptually run a "bubble sort" on $\mathbf{x}(\pi_2)$ to turn π_2 step by step into π_1 (this is possible since the same elements are acceptable in both rankings). In every step, we compare a pair of adjacent acceptable elements in $\mathbf{x}(\pi_2)$ and swap them if their order does not match their order according to π_1. This results in a new partial ranking π_2'. Notice that two advertisers are swapped only if they're in the wrong order, and so π_2' is more ordered than π_2. We can thus apply the above proof for identical rankings up to consecutive flipped elements to π_2 and π_2', and get that $\mathbf{s} \cdot \mathbf{x}(\pi_2') \geq \mathbf{s} \cdot \mathbf{x}(\pi_2)$. Thus $\mathbf{s} \cdot \mathbf{x}(\pi_2)$ is weakly increasing with each step of the bubble sort, completing the proof. □

We are now ready to prove Theorem 1.

Proof of Theorem 1 Consider n advertisers with i.i.d. MHR values $\mathbf{v}$, competing for $m \geq n$ ad slots to appear along search results for a query q. For every advertiser i, let $\bar{p}_i$ be the relevance prediction according to $\bar{T}$, and let p_i be the prediction according to the distinguishing refinement T. We want to show that with respect to social efficiency, the trade-off optimal mechanism M performs better with T than with $\bar{T}$.

For simplicity, rename the advertisers such that their true realized values, i.e., those according to the refined scheme T, are in decreasing order $p_1 v_1 \geq \cdots \geq p_n v_n$. (These are the true realized values since they are based on an accurate prediction of the click-through rates, and so reflect the true added efficiency from allocating a slot to each advertiser). The advertisers are now ordered according to the efficiency-optimal ranking.

We know that M is α-virtual value based for some α (Lemma 4). It thus ranks the advertisers by their realized α-virtual values – either $\bar{p}_i \varphi_i^\alpha$ if using scheme $\bar{T}$, or $p_i \varphi_i^\alpha$ if using scheme T. Let π_1 be the partial ranking of advertisers with non-negative α-virtual values according to T, and let π_2 be the same according to $\bar{T}$.

We first claim it is sufficient to show that, as partial rankings of the advertisers and their true realized values $\mathbf{r} = (p_1 v_1, \ldots, p_n v_n)$, π_1 is more ordered than π_2. The sufficiency follows from the generalized rearrangement inequality (Lemma 7): Both $\mathbf{r}$ and the vector of slot effects $\mathbf{s}$ are decreasing, so if π_1 is more ordered than π_2 it holds that $\mathbf{s} \cdot \mathbf{r}(\pi_1) \geq \mathbf{s} \cdot \mathbf{r}(\pi_2)$, i.e., the efficiency of M with T is at least its efficiency with $\bar{T}$.

It's left to show that π_1 is more ordered than π_2. First observe that advertiser i is acceptable according to either partial ranking if and only if his (non-realized) α-virtual value φ_i^α is non-negative, and so π_1 and π_2 rank the same subset of advertisers as acceptable. Now consider a pair of acceptable advertisers i, j ($\varphi_i^\alpha, \varphi_j^\alpha \geq 0$), whose realized values are $r_i = p_i v_i < p_j v_j = r_j$. We claim that if their ranking according to π_1 is reversed (i appears before j even though his realized value is lower), then this will also be the case according to π_2, and so π_1 is indeed more ordered.

If the ranking according to π_1 is reversed then we know that $p_i \varphi_i^\alpha \geq p_j \varphi_j^\alpha$. We can now invoke Lemma 6 to get $\bar{p}_i \varphi_i^\alpha \geq \bar{p}_j \varphi_j^\alpha$ (note that while Lemma 6 requires that $\varphi_i^\alpha, \varphi_j^\alpha$ are both positive, if at least one of these is zero then the inequality holds trivially). We have shown that advertiser i is ranked before j in π_2, completing the proof. □

5 Discussion and Future Directions

In this section we discuss the necessity of Theorem 1's assumptions, namely, i.i.d. MHR values and a distinguishing prediction refinement. We show the assumptions are necessary in a "worst-case" sense: if an assumption is violated, there's a setting in which prediction refinement harms expected efficiency, where expectation is taken over value profiles as well as over the refined subpart to which the query belongs (this rules out even a non-pointwise but less conditional version of Theorem 1). However, we expect that in many other non-worst-case settings, refinement will still contribute to efficiency despite the violated assumption. We demonstrate this below but leave the question of identifying when the assumptions can be weakened as an open problem. The settings below are based on Examples 1 and 2 in Section 2.2, in which there are two advertisers and a single ad slot, and use the revenue-optimal Myerson mechanism. We conclude with additional directions for future research.

5.1 Revisiting Example 2

We return to Example 2 in which the refinement was non-distinguishing. Recall that taking into account the user location could possibly make the two advertisers seem more alike, and thus in more direct competition. To take advantage of this, the revenue-optimal mechanism sometimes allocates to the advertiser with lower realized value, increasing the expected revenue but decreasing the efficiency.

As a result, the assumption of a distinguishing refinement is necessary in a strong sense for a *pointwise* statement such as Theorem 1. That is, for every (non-degenerate) MHR distribution, there exist valuations for which efficiency of the revenue-optimal mechanism falls with non-distinguishing refinement.

Moreover, for a specific selection of relevance parameters, refinement and value distribution, efficiency loss can happen *in expectation* when the distinguishing refinement assumption is violated. We use the relevance parameters from Example 2, and assume that the values v_1, v_2 are drawn independently from the MHR uniform distribution over range $[3, 5]$. The ranges of the realized values and virtual values are as follows: For advertiser 1, since his relevance prediction is 0.8 whether or not the prediction scheme is refined, his realized value range is $[2.4, 4]$ and his virtual value range is $[0.8, 4]$. As for advertiser 2, there are three cases to consider:

1. User from SF and refined scheme T applied – realized value range is $[1.2, 2]$ and virtual value range is $[0.4, 2]$;
2. User not from SF and refined scheme T applied – realized value range is $[3\epsilon, 5\epsilon]$ and virtual value range is $[\epsilon, 5\epsilon]$;
3. Coarse scheme $\bar{T}$ applied – realized value range is $[0.3, 0.5]$ and virtual value range is $[0.1, 0.5]$.

We conclude that applying the refined prediction scheme T which uses the location data lowers the expected welfare: Observe that the realized value of the advertiser 1 is always higher, and when $\bar{T}$ is applied his virtual value is always higher as well, guaranteeing an efficient allocation. But when the user is from SF, the relevance predictions of the advertisers become closer, and the range of advertiser 2's refined virtual value overlaps that of advertiser 1, and so advertiser 2 sometimes wins despite this being inefficient.

How often would we expect such inefficienies due to non-distinguishing refinements in general? Assuming the setting of parameters as above, suppose we vary the second advertiser's relevance from 0 towards 0.8, and plot efficiency loss against the optimally efficient outcome; see Fig. 2. As the figure shows, several refinements that are not distinguishing would still result in an efficiency increase. For instance, any refinement where $\bar{p}_2 > 0.4$ and p_2 is in the range $[\bar{p}_2, 0.8]$ would cause an efficiency increase. However, when $\bar{p}_2 < 0.4$ and p_2 is in the range $[\bar{p}_2, 0.4]$, the corresponding refinement causes an efficiency drop. Thus, we would expect that if the relevances of advertisers were initially roughly comparable (recall that for the first advertiser, $p_1 = \bar{p}_1 = 0.8$), any refinement ought to improve expected-efficiency.

Remark 2 (Non-I.i.d.) Example 2 can also be adjusted such that the resulting setting is completely equivalent, but now T is a distinguishing refinement of $\bar{T}$. This is by noticing that if advertisers are allowed to have non-i.i.d. values, a distinguishing refinement can actually make them more similar. For example, let advertiser 2's value be uniform over $[1.2, 2]$ instead of $[3, 5]$, and assume that finding out the user is from SF makes the advertisers' relevances flip from, say, 0.8,0.25 to 0.8,1. The ranges of their realized values however get closer, and the virtual value ranges overlap as above, leading to inefficiency. Thus the assumption of i.i.d. values is also in some sense necessary.

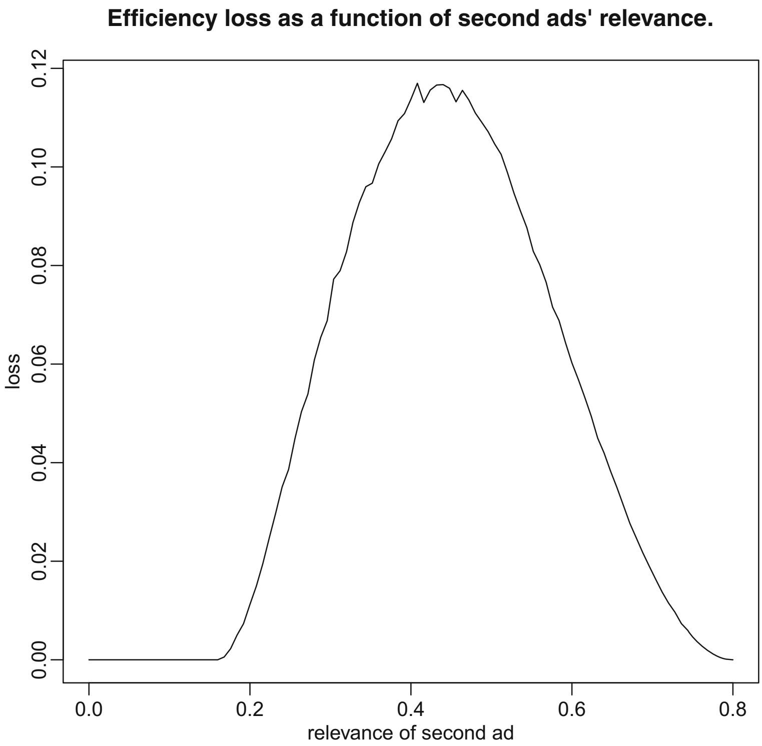

Fig. 2 Efficiency loss as a function of relevance

5.2 Revisiting Example 1

We demonstrate that refinement may decrease efficiency when values are non-MHR. We emphasize that this is not equivalent to demonstrating the *necessity* of the MHR assumption; indeed, there may be an alternative assumption that more tightly encircles the scenarios under which our main result holds.

To show what goes wrong, we analyze the expected loss in efficiency due to refinement and due to coarseness – see (8) and (9) below. This enables us to calculate and compare these two expected losses for a specific value distribution F, a truncated and shifted variant of the equal revenue distribution $\tilde{F}(v) = 1 - \frac{1}{v}$ which is non-MHR but regular.

Specifically, let $H = 10^3$ be a truncating parameter and $b = -1$ a shifting parameter. Let F be the variant of the equal revenue distribution achieved by truncating its support from $[1, \infty)$ to $[1, H]$ (the truncation ensures that F has finite expectation), and shifting it to the right by $|b|$. (Note that if $H = \infty$ and $b = 0$ we get the standard equal revenue distribution.)

We revisit the setting of Example 1 in which advertiser 1 is from SF and advertiser 2 from SJ. We assume the user is from SJ (the other case is symmetric). For the analysis, we fix values $v \geq v' > 0$ but do not specify which value belongs to which advertiser; by symmetry both possibilities are equally probable. Let φ, φ' be the virtual values, where by regularity $\varphi \geq \varphi' \geq 0$ (we assume φ' is non-negative and φ is positive, since otherwise refinement changes neither allocation nor efficiency). Then the expected loss in efficiency from inefficiencies due to refinement is

$$\frac{1}{2}\int_{v_{\min}}^{v_{\max}}\int_{2v'}^{\bar{v}}(v/2 - v')f(v)f(v')dvdv', \tag{8}$$

where $[v_{\min}, v_{\max}]$ is the range of distribution F, and $\bar{v}$ is defined to be the value such that the corresponding virtual value $\varphi(\bar{v})$ is equal to $2\varphi'$. The expected loss in efficiency from inefficiencies due to coarseness is

$$\frac{1}{2}\int_{v_{\min}}^{v_{\max}}\int_{v_{\min}}^{2v'}(v' - v/2)f(v)f(v')dvdv'. \tag{9}$$

Calculation of the integrals show that the expected efficiency loss due to refinement surpasses in this case the loss due to coarseness.

5.3 Implications for Real-World Search Auctions

We conclude with two future research directions aimed at closing any gaps between the theoretical contribution in this paper and actual implications for real-life search auctions.

First, in practice, not all improvements to prediction accuracy are achieved by adding features. Certainly, the addition of new features can be a dominant factor, but there are also algorithmic improvements, more accurate click models, etc. This paper provides a starting point for studying the high-level question of whether predicting more accurately leads to better auction outcomes, and an open question is to provide a generalized model and analysis in order to understand whether our results extend to other improvements in prediction accuracy besides refinement.

Second, this paper provides theoretical guidelines to search auction designers when contemplating the interaction between click prediction and auctions. Namely, it introduces a desired "flip-spread" effect of refinement that should serve as a guiding factor in deciding whether to refine. An open direction is to test these guidelines empirically, by comparing between bidder relevance before and after adding features, analyzing how often the effect is a distinguishing one, and testing the correlation between this and between how well the auction achieves its objectives.

Acknowledgments The authors wish to thank Amir Najmi for suggesting the problem, Mohammad Mahdian for suggesting that Pareto optimal mechanisms are virtual value-based, and Qiqi Yan for many helpful comments. I. Talgam-Cohen gratefully acknowledges the support of the Hsieh Family Stanford Interdisciplinary Graduate Fellowship. For an early version see [24].

References

1. Aggarwal, G., Goel, A., Motwani, R.: Truthful auctions for pricing search keywords. In: Proceedings of 7th ACM Conference on Economics and Computation (EC), pp. 1–7 (2006)

2. Aggarwal, G., Goel, G., Mehta, A.: Efficiency of (revenue-)optimal mechanisms. In: The 10th ACM Conference on Electronic Commerce, pp. 235–242 (2009)
3. Clarke, E.H.: Multipart pricing of public goods. Public Choice **11**, 17–33 (1971)
4. Daskalakis, C., Pierrakos, G.: Simple, optimal and efficient auctions. In: Proceedings of 7th Conference on Web and Internet Economics (WINE) (2011)
5. Diakonikolas, I., Papadimitriou, C., Pierrakos, G., Singer, Y.: Efficiency-revenue trade-offs in auctions. In: Proceedings of ICALP (2012)
6. Edelman, B., Ostrovsky, M., Schwarz, M.: Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. Am. Econ. Rev. **97**(1), 242–259 (2007)
7. Edelman, B., Schwarz, M.: Optimal auction design and equilibrium selection in sponsored search auctions. Am. Econ. Rev. **100**(2), 597–602 (2010)
8. Emek, Y., Feldman, M., Gamzu, I., Paes Leme, R., Tennenholtz, M.: Signaling schemes for revenue maximization. ACM Trans. Econ. Comput. **2**(2), 5 (2014)
9. Ewerhart, C.: Optimal design and ρ-concavity, working paper (2009)
10. Fu, H., Jordan, P., Mahdian, M., Nadav, U., Talgam-Cohen, I., Vassilvitskii, S.: Ad auctions with data. In: Proceedings of 5th International Symposium on Algorithmic Game Theory (SAGT), pp. 168–179 (2012)
11. Ghosh, A., Nazerzadeh, H., Sundararajan, M.: Computing optimal bundles for sponsored search. In: Proceedings of 3rd Conference on Web and Internet Economics (WINE), pp. 576–583 (2007)
12. Graepel, T., Candela, J.Q., Borchert, T., Herbrich, R.: Web-scale Bayesian click-through rate prediction for sponsored search advertising in Microsoft's Bing search engine. In: 27th International Conference on Machine Learning (ICML), pp. 13–20 (2010)
13. Groves, T.: Incentives in teams. Econometrica **41**, 617–631 (1973)
14. Lahaie, S., Pennock, D.M.: Revenue analysis of a family of ranking rules for keyword auctions. In: Proceedings of 8th ACM Conference on Economics and Computation (EC), pp. 50–56 (2007)
15. Lahaie, S., Pennock, D.M., Saberi, A., Vohra, R.V.: Sponsored search auctions. In: Nisan, N., Roughgarden, T., Tardos, É., Vazirani, V. (eds.) Algorithmic Game Theory, chap. 28, pp. 699–716. Cambridge University Press (2007)
16. Likhodedov, A., Sandholm, T.: Auction mechanism for optimally trading off revenue and efficiency. In: Proceedings of 4th ACM Conference on Economics and Computation (EC), pp. 212–213 (2003)
17. McAfee, R., McMillan, J.: Auctions and bidding. J. Econ. Lit. **25**(2), 699–738 (1987)
18. Mehta, A., Saberi, A., Vazirani, U., Vazirani, V.: Adwords and generalized on-line matching. J. ACM **54**(5), 22 (2007)
19. Milgrom, P., Weber, R.J.: A theory of auctions and competitive bidding. Econometrica **50**, 1089–1122 (1982)
20. Miltersen, P.B., Sheffet, O.: Send mixed signals: Earn more, work less. In: Proceedings of 13th ACM Conference on Economics and Computation (EC), pp. 234–247 (2012)
21. Myerson, R.: Optimal auction design. Math. Oper. Res. **6**(1), 58–73 (1981)
22. Myerson, R., Satterthwaite, M.: Efficient mechanisms for bilaterial trade. J. Econ. Theory **29**(1), 265–281 (1983)
23. Neeman, Z.: The effectiveness of English auctions. Games Econ. Behav. **43**(2), 214–238 (2003)
24. Sundararajan, M., Talgam-Cohen, I.: Prediction and welfare in ad auctions. In: Proceedings of 7th International Symposium on Algorithmic Game Theory (SAGT), pp. 267–278 (2014)
25. Taleghan, M.A., Hamdaoui, B.: Efficiency-revenue optimality tradeoffs in dynamic spectrum allocation. In: IEEE Global Telecommunications Conference (GLOBECOM) (2010)
26. Varian, H.R.: Position auctions. Int. J. Ind. Organ. **25**(6), 1163–1178 (2007)
27. Vickrey, W.: Counterspeculation, auctions, and competitive sealed tenders. J. Financ. **16**, 8–37 (1961)