Ad Auctions with Data*

Hu Fu[1], Patrick Jordan[2], Mohammad Mahdian[3], Uri Nadav[3], Inbal Talgam-Cohen[4], and Sergei Vassilvitskii[3]

[1] Cornell University, Ithaca, NY, USA
[2] Microsoft Inc., Mountain View, CA, USA
[3] Google Inc., Mountain View, CA, USA
[4] Stanford University, Stanford, CA, USA

Abstract. The holy grail of online advertising is to target users with ads matched to their needs with such precision that the users respond to the ads, thereby increasing both advertisers' and users' value. The current approach to this challenge utilizes information about the users: their gender, their location, the websites they have visited before, and so on. Incorporating this data in ad auctions poses an economic challenge: can this be done in a way that the auctioneer's revenue does not decrease (at least on average)? This is the problem we study in this paper. Our main result is that in Myerson's optimal mechanism, for a general model of data in auctions, additional data leads to additional expected revenue. In the context of ad auctions we show that for the simple and common mechanisms, namely second price auction with reserve prices, there are instances in which additional data decreases the expected revenue, but this decrease is by at most a small constant factor under a standard regularity assumption.

1 Introduction

When an item with latent characteristics is sold, information revealed by the seller plays a significant role in the value ascribed to the item by potential buyers. For example, when booking a hotel room on a website such as Priceline.com, every extra piece of information—including the hotel's star level or its location—affects the price a buyer is willing to pay. In a similar manner, in online advertising scenarios, any information revealed about the ad opportunity—including the description of the webpage's content or the type of user—plays a crucial role in determining the ad's value, in particular because this information is extremely useful in predicting the click and conversion rate of the user.

In online display advertising settings, the publisher auctions off opportunities to show an advertisement to its users in real time, often through online ad marketplaces operated by companies such as Yahoo!, Google or Microsoft. For example, every time a user visits The New York Times website, the opportunity to show an advertisement to the user is auctioned off. Both the publisher (in this case The New York Times) and the market operator have a great deal of information about the ad opportunity, including page specific features such as layout and content, as well as user specific features such as the user's age, gender,

* Work done while the authors were at Yahoo! Reseaerch

location, etc. *How much of this information should be revealed during the auction in order to maximize revenue?* This is the question we study in this work.

While concealing information can only decrease social efficiency, it may be advantageous in terms of revenue, since releasing information may decrease competition. As an example, suppose an advertiser values males at \$2 and females at \$8. In an incentive compatible auction, the advertiser bids his value when the user's gender is known, but will bid the expected value of \$5 when the gender is not revealed (assuming each gender is equally likely). If there is a second advertiser who values males at \$8 and females at \$2, then revealing gender segments the buyers. As a result, when gender is revealed the auctioneer will face a bid of \$8 and \$2, and thus collect only \$2 in a second price auction; if the gender is kept hidden, the auctioneer will have two bids of \$5 and will collect \$5.[1]

The example above may seem to suggest that it is never in the auctioneer's interest to release information about the item. Indeed, Board [5] has shown that revealing information can only decrease the expected revenue from a second price auction with two bidders. However, the auctioneer has additional tools to increase revenue at her disposal, namely she can set a reserve price for each bidder. The right reserve price may counter the potential loss in competition, allowing the auctioneer to preserve its revenue. In the example above, a reserve price of \$8 for both advertisers would lead to a revenue of \$8 precisely in the case where gender is revealed. On the other hand, it is not obvious that using reserve prices or even applying the optimal mechanism is sufficient to recover the lost revenue from revealing data; see Example 1 in Section 4.2 for a simple case in which this does not hold.

Our Contribution In this work we study a general model of single-parameter auctions with data. We show that while revealing information can lead to a decrease in the expected revenue of second price auctions, using the revenue-optimal mechanism counteracts this trend. Our main result is that if Myerson [17]'s optimal auction mechanism is used, the expected revenue is guaranteed to (weakly) increase when more information is revealed. This result also applies to slot auctions and other settings.

We explore the assumptions of this result and show that they are necessary for revenue monotonicity to hold. In particular, if instead of Myerson's optimal mechanism, a simpler reserve price based mechanism is used, revealing information can lead to a decrease in expected revenue. However, we prove that in simple and practical second price auctions with reserve prices, fully revealing the auctioneer's information generates approximately the optimal revenue even compared with arbitrary intricate revealing schemes the auctioneer may adopt.

1.1 Related Work

The following scenario has been extensively studied in auction theory: The auctioneer has access to a private source of data about the item; she wishes to

[1] Perturbing this example slightly shows withholding information can decrease social welfare.

maximize her expected revenue by pre-committing to a policy of revealing or concealing data. Two effects of revealing data have been identified: the *linkage principle* by Milgrom and Weber [15], and more recently the *allocation effect* by Board [5]. The linkage principle says that when bidders' valuations are positively correlated in a specific way to the auctioneer's data, the auctioneer can increase her revenue in first or second price auctions by revealing the data. However, in ad auctions, revealing information can increase the value to some advertisers and decrease it for the rest, and so the linkage principle does not apply. The allocation effect studies the effect of information revelation on revenue in second price auctions as the number of bidders changes.

A recent line of research [9, 16] considers the computational problem of finding the optimal information revelation scheme in second price auctions. In contrast, we study optimal auctions and their implications for auctions such as second price auctions with reserve prices, and never second price auctions per se. The valuation model of ours is also different. It is neither deterministic nor arbitrarily correlated, as studied by Emek et al. [9].

We briefly mention related work further afield. Levin and Milgrom [13] highlight disadvantages of information revelation from a market design point of view—too much information leads to thin markets that are hard to operate. Several proposed mechanisms address these issues [4, 7]. Dwork et al. [8] discuss fairness concerns arising from revealing user data. A separate body of work considers cases in which bidders, not the auctioneer, have private sources of information about the item, resulting in asymmetries among them; a recent example is Abraham et al. [1]. Ghosh et al. [10] study information revelation in ad auctions through the process of cookie-matching, and its impact on the revenue of the auction.

2 Preliminaries

We briefly describe Myerson [17]'s optimal truthful mechanism, under the interpretation of Bulow and Roberts [6]. Given a valuation distribution F, each probability quantile q corresponds to a value $v = F^{-1}(1-q)$. Each value, in turn, corresponds to an expected revenue $v(1-F(v))$ generated by setting a posted price of v. A revenue curve depicts such revenue $R(q) = qF^{-1}(1-q)$ as a function of the quantile q, and the ironed revenue curve $\tilde{R}(q)$ is the concave hull of this curve. The ironed virtual valuation of v under the distribution F is then $\tilde{\varphi}(v) = \frac{\mathrm{d}\tilde{R}(q)}{\mathrm{d}q}|_{q=1-F(v)}$.

Theorem 1 (Myerson 17). *In a revenue optimal truthful auction in which bidders' valuations are independently drawn from known distributions $D_1, \cdots, D_n$, the item is allocated to the bidder with the highest non-negative ironed virtual valuation, and the expected revenue is equal to* $\mathbf{E}[\max\{0, \tilde{\varphi}_1(v_1), \ldots, \tilde{\varphi}_n(v_n)\}]$.

3 Model

We describe a general model to which our main result applies, and show that ad auctions are captured in a natural way as an instantiation of this model.

General model. n bidders compete in an auction, in which the subsets of bidders who can win simultaneously are specified by $\mathcal{I} \subseteq 2^{[n]}$, the *feasible sets*. Every bidder $i \in [n]$ has a private, single-dimensional *signal* $s_i \in \mathbb{R}_+$, drawn independently from a publicly known distribution F_i with density f_i. In addition, the auctioneer also has a private signal $u \in U$, drawn from a publicly known discrete distribution F_U with density f_U independently of the bidder signals. We also call u an item type. Denote $|U|$ by m. Bidder i's value for winning the auction is a publicly known function v_i of his own signal s_i and the auctioneer's signal u: $v_{i,u} = v_i(s_i, u)$. For every u we assume that $v_i(\cdot, u)$ is non-negative and strictly increasing in its argument s_i. Note that, under these constraints, u can affect $v_{i,u}$ in a fairly general manner.

Signaling schemes. We adopt the framework for signaling schemes developed by Emek et al. [9] and Miltersen and Sheffet [16] which has its origins in Milgrom and Weber [15, see Theorem 9]. A signaling scheme is a set of m distributions over a signal set Σ of size k. On seeing type u, the auctioneer sends a signal $\sigma \in \Sigma$ with probability $\psi_{u,\sigma}$, and then bidders bid their expected value inferred from the posterior distribution on u given σ. It is important that the auctioneer commits to a signaling scheme before the auction starts. In the case of fully withdrawing the information, Σ has one element, which we call $\bar{u}$, and bidder i's posterior valuation in this case is denoted $v_{i,\bar{u}} = \mathbf{E}_{u \sim F_U}[v_{i,u}]$. In discussing this scenario, we often equivalently talk about a fictitious item type $\bar{u}$, for which each bidder i's valuation is $v_{i,\bar{u}}$.

Ad Auctions: An Instantiation. Ad auctions can be viewed as a special case of the above general model. In this case one opportunity of displaying an ad is auctioned to n bidders, and therefore the feasible sets $\mathcal{I}$ consist of single winning bidders. In this scenario, the item type u may refelect the auctioneer's information on the user to whom the ad is to be shown. (For this reason we also call u the *user type*.) A widely used model for ad auction is that a bidder i has a private value s_i for a user to click his ad, and for each user type u there is a particular probability $p_{i,u}$ with which the user does click. $p_{i,u}$ is the so-called *click-through rate*. Now a bidder's valuation is simply $v_{i,u} = v_i(s_i, u) = p_{i,u} s_i$.

4 Full Revelation in Myerson's Optimal Mechanism

In this section we present and prove our main result, which states that in Myerson's optimal mechanism, the expected revenue is monotone non-decreasing in revealed information, and so full revelation of the auctioneer's information maximizes the expected revenue.

Other results on information revelation in the private-value setting focus on the second price or English auctions [5, 14]. We first show, in the concrete context of ad auctions, why applying Myerson's mechanism gives different results. We then briefly discuss how to extend our results for ad auctions to slot auctions. In Section 4.2 we give a simple and general proof of the main result for the single-parameter model introduced in Section 3.

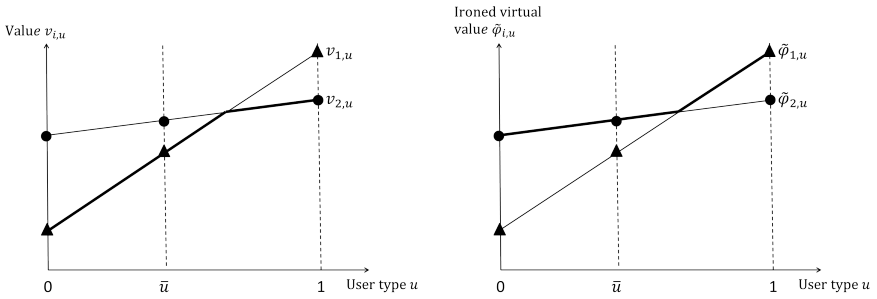
4.1 Second Price Auction vs. Myerson's Mechanism in Ad Auctions

Consider the two extreme signaling schemes of full revelation and no revelation. We show why in contrast to the result of Board [5] for the English auction, in Myerson's mechanism the former scheme is always preferable to the latter in terms of expected revenue (where expectations are taken over the random private signals and, where appropriate, over the random user type). The optimality of full revelation and monotonicity of expected revenue in information follow as corollaries.

Proposition 1. *In the ad auctions model, the expected revenue from Myerson's mechanism when the user's type u is revealed is at least as high as the expected revenue when u is not revealed.*

For completeness we include Board's result for 2 bidders (note that since $n = 2$, the second price and English auctions are the same).

Proposition 2 (Board [5]). *In a generalization of the ad auctions model with $n = 2$ bidders, the expected revenue from the second price or English auction when the user's type u is not revealed is at least as high as the expected revenue when u is revealed.*



(a) Second price auction [5] (b) Myerson's mechanism

Fig. 1: The Effect of Information Revelation

Figure 1, adapted from Board [5], provides intuition for the difference between the above propositions (see also Palfrey [18], McAfee [14]). In the second price auction, for every signal profile of the bidders, the revenue is the minimum of their values and so a concave function. Therefore, while revealing information produces the average of pointwise minimums, no revelation does at least as well by producing the minimum of averages. By contrast, recall that applying Myerson's mechanism in our setting means that given user type $u \in U \cup \{\bar{u}\}$, the auctioneer transforms the advertisers' values $\{v_{i,u}\}$ into the corresponding ironed virtual values $\{\tilde{\varphi}_{i,u}\}$, and then allocates the impression to the advertiser

with highest non-negative ironed virtual value. The expected revenue is equal to the expected ironed virtual surplus, and for every signal profile the ironed virtual surplus is convex, so the effect of data revelation is reversed.

To formalize this intuition we need to show that the same relation that holds for values before the transformation to ironed virtual values, when a bidder's value under no revelation is equal to his expected value under full revelation, continues to hold after the transformation as well. This is established in Observations 1 and 2. The proof of Proposition 1 then applies convexity and Jensen's inequality to get the result.

Observation 1 *Let advertiser i's value be $v_{i,u} = p_{i,u}s_i$, where $u \in U \cup \{\bar{u}\}$ is the user type and $s_i \sim F_i$. Then $v_{i,u}$ is distributed according to $F_{i,u}(x) = F_i(x/p_{i,u})$, and the corresponding ironed virtual value function is $\tilde{\varphi}_{i,u}(x) = p_{i,u}\tilde{\varphi}_i(x/p_{i,u})$.*

Proof. The derivation of $F_{i,u}$ is straightforward. The expression for the ironed virtual value follows by looking at the revenue curves R_i and $R_{i,u}$ corresponding to distributions F_i and $F_{i,u}$ respectively:

$$\begin{aligned} R_{i,u}(1 - F_{i,u}(x)) &= x(1 - F_{i,u}(x)) \\ &= p_{i,u} \cdot \frac{x}{p_{i,u}}(1 - F_i(x/p_{i,u})) \\ &= p_{i,u}R_i(1 - F_i(x/p_{i,u})). \end{aligned}$$

The ironed revenue curves are concave hulls of the revenue curves, and therefore preserve the same relationship $\tilde{R}_{i,u}(1 - F_{i,u}(x)) = p_{i,u}\tilde{R}_i(1 - F_i(x/p_{i,u}))$. The ironed virtual valuations, which are their derivatives, also satisfy the same linear relationship. □

We can now compare the ironed virtual values with and without information revelation $\tilde{\varphi}_{i,u}$ and $\tilde{\varphi}_{i,\bar{u}}$. We show the latter equals the former in expectation.

Observation 2 $\tilde{\varphi}_{i,\bar{u}}(v_{i,\bar{u}}) = \mathbf{E}_{u\sim F_U}[\tilde{\varphi}_{i,u}(v_{i,u})]$.

Proof. We have

$$\begin{aligned} \tilde{\varphi}_{i,\bar{u}}(p_{i,\bar{u}}s_i) = p_{i,\bar{u}}\tilde{\varphi}_i(s_i) &= \mathbf{E}_{u\sim F_U}\left[p_{i,u}\right]\tilde{\varphi}_i(s_i) \\ &= \mathbf{E}_{u\sim F_U}\left[p_{i,u}\tilde{\varphi}_i(s_i)\right] = \mathbf{E}_{u\sim F_U}\left[\tilde{\varphi}_{i,u}(p_{i,u}s_i)\right], \end{aligned}$$

where the first and last equalities are by Observation 1, the second is by definition of $p_{i,\bar{u}}$, and the third is by linearity of expectation. □

Proof of Proposition 1. The expected revenue of Myerson's mechanism is equal to its expected ironed virtual surplus [17] [see also 11, Theorem 13.10]. We use this result by Myerson to prove the proposition as follows. We show that pointwise for every fixed profile of values per click $(s_1, \ldots, s_n)$, the ironed virtual surplus of Myerson's mechanism when u is revealed is at least as high as when u is not revealed, in expectation over u. Taking expectation over profiles $(s_1, \ldots, s_n)$ and applying Myerson's result completes the proof.

Fix $(s_1, \ldots, s_n)$ and let $u \in U$ be the user's type. The ironed virtual surplus of Myerson's mechanism when u is revealed is $\max\{0, \tilde{\varphi}_{1,u}(p_{1,u}s_1), \ldots, \tilde{\varphi}_{n,u}(p_{n,u}s_n)\}$. We will omit the term 0 from this point on, since we can always add a dummy bidder whose valuation (and virtual valuation) is constantly 0. Taking expectation over u gives

$$\mathbf{E}_{u\sim F_U}\left[\max\{\tilde{\varphi}_{1,u}(p_{1,u}s_1), \ldots, \tilde{\varphi}_{n,u}(p_{n,u}s_n)\}\right]. \tag{1}$$

If u is not revealed, the ironed virtual surplus of Myerson's mechanism is

$$\max\{\tilde{\varphi}_{1,\bar{u}}(p_{1,\bar{u}}s_1), \ldots, \tilde{\varphi}_{n,\bar{u}}(p_{n,\bar{u}}s_n)\}.$$

By Observation 2, this is equal to

$$\max\left\{\mathbf{E}_{u\sim F_U}\left[\tilde{\varphi}_{i,u}(p_{i,u}s_i)\right]\right\}_{i=1}^{n}. \tag{2}$$

Since max is a convex function, by Jensen's inequality (1) $\geq$ (2). We conclude that in expectation over u, revealing the user's type u does not reduce the ironed virtual surplus. □

So far we have considered only two possible signaling schemes for the auctioneer: either to fully reveal the user's type or to conceal it. A direct corollary of Proposition 1 is that the full revelation strategy yields the highest expected revenue among *all* possible signaling schemes.

Corollary 1. *In the ad auctions model, the expected revenue from Myerson's mechanism when the user type is revealed is optimal among all signaling schemes.*

Proof. Consider a signaling scheme $\{\psi_{u,\sigma}\}_{u\in U,\sigma\in\Sigma}$. Condition on the revealed signal σ. Recall that together with the scheme $\{\psi_{u,\sigma}\}$ and the distribution F_U, it induces an ex post distribution $F_{U|\sigma}$ on the user types. We can now apply Proposition 1 to the setting in which $u \sim F_{U|\sigma}$, and conclude that the expected revenue from full revelation of u is at least as high as the expected revenue from revealing σ. Taking expectation over $\sigma \in \Sigma$ completes the proof. □

Furthermore, Proposition 1 implies *monotonicity* of optimal expected revenue in information release—adding any signaling scheme to Myerson's mechanism can only improve expected revenue.

Corollary 2. *In the ad auctions model, the expected revenue from Myerson's mechanism with a signaling scheme is at least the expected revenue from Myerson's mechanism with no signaling.*

Proof. Let f_U be the density of the user types in the original setting, and let $\{\psi_{u,\sigma}\}_{u\in U,\sigma\in\Sigma}$ be the signaling scheme. Now consider the following alternative setting: A user type σ is sampled from Σ with probability $\sum_{u\in U} f_U(u)\psi_{u,\sigma}$, and the bidders' values are $\{v_{i,\sigma}\}$. Observe that the expected revenue from Myerson's mechanism with signaling scheme $\{\psi_{u,\sigma}\}$ in the original setting equals the expected revenue from Myerson's mechanism with full revelation in the new setting. Similarly, the expected revenue from Myerson with no signaling is the same in both settings. Applying Proposition 1 to the new setting we get that the expected revenue from full revelation of σ is at least as high as the expected revenue from no revelation, completing the proof. □

Generalization to Slot Auctions A particular case of practical interest is slot auctions, in which the auctioneer has k slots $\{1, \cdots, k\}$ to sell to the advertisers, and an advertiser's value for winning depends on the particular slot he gets. We extend our result for ad auctions to show that full information revelation is the optimal signaling scheme for optimal slot auctions. The main thing to show is that the ironed virtual surplus remains convex.

Formally, the slots have intrinsic click through rates $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k$. An advertiser's valuation for a user of type u at slot j is $\alpha_j p_{i,u} s_i$. By the same argument as in Observation 1, his ironed virtual valuation is $\alpha_j p_{i,u} \tilde{\varphi}_i(s_i)$. The optimal auction ranks the k bidders with highest non-negative ironed virtual valuations and maps them to the k slots accordingly (if there are fewer than k bidders with non-negative ironed virtual valuations then the remaining slots are not sold). The auctioneer's expected revenue is then the expected sum of the k highest ironed virtual valuations. Just as the proof of Proposition 1 and its corollaries relies on the fact that taking maximum is a convex function, a similar full revelation statement for slot auctions follows from the next observation, whose proof is a consequence of the rearrangement inequality and appears in the appendix.

Observation 3 *The function* $M_k(v_1, \ldots, v_n) = \sum_{j=1}^{k} \alpha_j \max\text{-}j\{v_1, \ldots, v_n\}$ *is a convex function in* $(v_1, \ldots, v_n)$*, where* $\max\text{-}j\{v_1, \ldots, v_n\}$ *is the* j*-th largest element from the set* $\{v_1, \ldots, v_n\}$*.*

For space consideration we omit the proof of this, which is a simple application of the rearrangement inequality.

By the same argument as before using Jensen's inequality, we obtain

Corollary 3. *In slot auctions, the expected revenue from Myerson's mechanism when the user type is revealed is optimal among all signaling schemes.*

4.2 General Model

We extend the optimal revelation results in Section 4 to the general single-parameter model introduced in Section 3. The proof there uses the specific form of ironed virtual values found in this model (also used in Section 5.1). The specific ironed virtual values form is not necessary for the result to hold, and here we prove a general full revelation result for Myerson's mechanism based only on its optimality and monotonicity, and not on the details of its allocation rule.

Proposition 3. *In the general single-parameter model with values* $v_{i,u} = v_i(s_i, u)$*, where* v_i *is non-negative, strictly increasing in* s_i *and continuously differentiable for every* i*, the expected revenue from Myerson's mechanism when the auctioneer's information* u *is revealed is optimal among all signaling schemes.*

Proof. Similarly to ad auctions, it is sufficient to compare full revelation to no revelation. Assume full revelation, and fix the revealed signal of the auctioneer to be $u \in U$. We define the following auxiliary mechanism M. Mechanism M

receives reported values $\{v_{i,u}\}$ from the bidders. By the assumption that $v_{i,u} = v_i(s_i, u)$ is strictly increasing in bidder i's signal s_i, for every i and u there is a one-to-one relation between bidder i's signals and values. Therefore mechanism M may recover the bidders' signals from their reported values. It then finds $\{v_{i,\bar{u}}\}$, the set of values that would have been reported by the bidders if no data had been revealed. Finally, M runs Myerson's mechanism on these values, assuming they're drawn from distributions $\{F_{i,\bar{u}}\}$.

We first claim that the auxiliary mechanism M is truthful, i.e., that its allocation rule is monotone in the reported values $\{v_{i,u}\}$. Fix i and s_{-i}. We want to show that increasing $v_{i,u}$ can only cause M to allocate to bidder i more often. By truthfulness of Myerson's mechanism, we know that M is monotone in $v_{i,\bar{u}}$. Again by the assumption that $v_{i,u}$ is strictly increasing in s_i for every u, the expectation $v_{i,\bar{u}}$ is also strictly increasing in s_i. So M is monotone in s_i, and thus also in $v_{i,u}$, as required.

Consider the expected revenue of the auxiliary mechanism M. On one hand, in expectation over $u \sim F_U$ and the signal profile $\boldsymbol{s}$, its revenue equals that of Myerson's mechanism with no data revelation. On the other hand, for every fixed $u \in U$, Myerson's mechanism with full revelation does at least as well as M in terms of expected revenue over the signal profile, simply because it is optimal. We conclude that the expected revenue of Myerson with full revelation is at least as high as with no revelation, completing the proof. □

The following example shows that optimality of full revelation does not hold without the assumption that a bidder's value is strictly increasing in his signal.

Example 1. Assume u is distributed uniformly over $\{0, 1\}$, and there's a single bidder whose private signal s is distributed uniformly over a discrete support $\{1, 2, 3\}$. When $u = 0$, the bidder's value is just his signal, i.e., $v(s, 0) = s$. When $u = 1$, the values are $v(s, 1) = 4 - s$. Then with full revelation, the maximum expected revenue is $\frac{4}{3}$ by setting a reserve price of 2. When no information is revealed, the bidder's value is $v_{\bar{u}} = 2$ and so the expected revenue is 2.

5 Full Revelation in Simple Auctions with Reserve Prices

In this section we show several results relating to simple, commonly-used ad auctions, namely second price auctions with *anonymous reserves*, and second price auctions with *monopoly reserves*. In the former, a single reserve price is applied to all advertisers, and only those who bid above the reserve compete in a second price auction. In the latter, a distinct monopoly reserve price is applied to each advertiser, and advertisers who bid above their respective reserves enter the second price auction. The monopoly reserve price for a bidder with regular distribution is the optimal price for the auctioneer to set in an auction where only this bidder participates. Equivalently, it is equal to the value v whose corresponding virtual value $\varphi(v)$ is 0.

First, in Section 5.1, we complement our results for optimal ad auctions by showing that in second price auctions with reserves, fully revealing information

is approximately optimal among all signaling schemes, provided that advertisers' distributions are regular. This is encouraging in light of previous results on signaling in second price auctions without reserves: Emek et al. [9] showed that finding the optimal signaling scheme is NP-hard, and no approximation algorithm is known yet. We note that in practice, second price auctions with reserve are more common than those without reserve.

In Sections 5.2 and 5.3, we demonstrate that full revelation in simple auctions can be sub-optimal. In fact, revealing no information at all can sometimes leave the auctioneer better off, even for distributions such as the uniform distribution, although by no more than a small constant factor, as we show in Section 5.1.

5.1 Approximation Guarantee in Simple Ad Auctions

We recall the following result of Hartline and Roughgarden [12] on the performance of second price auctions with reserves.

Theorem 2. *For every single-item setting with values drawn independently from regular distributions,*

1. *the expected revenue of the second price auction with the optimal anonymous reserve price is a* 4*-approximation to the optimal expected revenue; and*
2. *the expected revenue of the second price auction with monopoly reserves is a* 2*-approximation to the optimal expected revenue.*

Corollary 4. *In ad auctions, when bidders' valuations per click s_i are independently drawn from regular distributions, fully revealing the type in a second price auction with anonymous reserve (monopoly reserves, resp.) is a* 4*-approximation (*2*-approximation, resp.) to the expected revenue of the optimal signaling scheme.*

Proof. Consider an optimal signaling scheme in a second price auction with reserves. Under the same signaling scheme, running Myerson's optimal auction would extract at least the same expected revenue. By Corollary 1, fully revealing the user type is optimal among all signaling schemes in Myerson's auction. Then for every user type u, we apply Observation 1 and the regularity of the s_i's to establish regularity of the $v_{i,u}$'s, and so a second price auction with reserves extracts a 4 (or 2)-approximation by Theorem 2. We conclude that fully revealing the information u in a second price auction with reserves extracts a 4 (or 2)-approximation to the revenue obtained by Myerson's optimal auction with full information revelation. The corollary follows from this chain of bounds. □

5.2 Revenue Loss with Anonymous Reserve

This section gives an example in which announcing the item type decreases the revenue of the second price ad auction with the optimal anonymous reserve price.

The example has $n = 2$ bidders and $m = 2$ item types, with F_U being uniform between the two types. Bidder 1's valuation for a "high" type is uniformly drawn from $[0, 2]$, and for a "low" type is 0. Bidder 2 is not sensitive to the types and her valuation is drawn uniformly from $[0, 1]$ regardless of the type. When the type is not announced, the optimal auction is a second price auction with reserve

price 1/2, and the optimal revenue is 5/12. When the type is revealed to be low, the optimal auction is a second price auction with a reserve price 1/2, and the revenue is 1/4. When the type is revealed to be high, if we set a reserve price of $x \in [0, 1]$, the revenue is

$$x\left[x(1-\frac{x}{2})+\frac{x}{2}(1-x)\right]+\int_x^1 y(1-\frac{y}{2})+\frac{y}{2}(1-y)\,dy=\frac{3}{4}x^2-\frac{2}{3}x^3+\frac{5}{12}.$$

This is maximized at $x = 3/4$, yielding a revenue of $\frac{9}{64}+\frac{5}{12}$. Setting a reserve price in $[1, 2]$ does no give a revenue better than 0.5. Therefore, for a high type, the revenue of an optimal second price auction with anonymous reserve is $\frac{9}{64}$ more than $\frac{5}{12}$, whereas for a low type the revenue is $\frac{1}{6}$ less. On average, if we reveal the type, the expected revenue is strictly less than $\frac{5}{12}$.

5.3 Revenue Loss with Monopoly Reserves

This section presents an example in which announcing the item type decreases the revenue of the second price ad auction with monopoly reserve prices.

As in the previous section, we assume 2 bidders and 2 types, with F_U being uniform. Bidder 1's valuation is uniformly drawn from $[0, 8]$ for a "high" type, and uniformly from $[0, 4]$ for a "low" type, whereas bidder 2 is not sensitive to the item type and her valuation is uniformly drawn from $[0, 6]$ regardless of the type. When the type is not revealed, the optimal auction is a second price auction with reserve price 3, and the expected revenue is 2.5. When the type is revealed to be high, the monopoly reserves are 4 and 3, respectively. The expected revenue is:

$$\begin{aligned}&4\cdot\Pr(v_1\in[4,8],v_2\in[0,3])+3\cdot\Pr(v_1\in[0,4],v_2\in[3,6])+\\&4\cdot\Pr(v_1\in[4,8],v_2\in[3,4])+\frac{14}{3}\cdot\Pr(v_1,v_2\in[4,6])+\\&\qquad 5\cdot\Pr(v_1\in[6,8],v_2\in[4,6])=2.889.\end{aligned}$$

When the type is revealed to be low, the monopoly reserves are 2 and 3, respectively. The expected revenue is:

$$\begin{aligned}&2\cdot\Pr(v_1\in[2,4],v_2\in[0,3])+3\cdot\Pr(v_1\in[0,2],v_2\in[3,6])+\\&3\cdot\Pr(v_1\in[2,3],v_2\in[3,6])+\frac{7}{2}\cdot\Pr(v_1\in[3,4],v_2\in[4,6])+\\&\qquad\frac{10}{3}\cdot\Pr(v_1,v_2\in[3,4])=2.0556.\end{aligned}$$

Thus, the expected revenue when the type is revealed is $2.4722 < 2.5$.

6 Conclusion and Open Questions

Incorporating data into ad auctions raises many questions of practical importance to which our work may be applicable. We mention two open questions:

(1) In simple second price ad auctions, an intermediate revelation scheme may generate more revenue than both full revelation and no revelation. Can the auctioneer find such a scheme in a computationally efficient way? This question was studied by Emek et al. [9] in settings either more general or more restricted than ours, and remains open for the ad auctions model. (2) Can the auctioneer increase her revenue by *asymmetric* revelation of information to the bidders, perhaps charging them appropriate prices for the information? The answer will involve overcoming several challenges, some of which are studied in [1, 2, 3].

Bibliography

[1] Abraham, I., Athey, S., Babaioff, M., and Grubb, M. (2011). Peaches, lemons, and cookies: Designing auction markets with dispersed information. Manuscript.

[2] Alon, N., Feldman, M., Gamzu, I., and Tennenholtz, M. (2012). The asymmetric matrix partition problem. Manuscript.

[3] Babaioff, M., Kleinberg, R., and Leme, R. P. (2012). Optimal mechanisms for selling information. In *ACM Conference on Electronic Commerce*, pages 92–109.

[4] Beck, M. and Milgrom, P. (2011). Auctions, adverse selection and internet display advertising. Working paper.

[5] Board, S. (2009). Revealing information in auctions: the allocation effect. *Econ Theory*, 38:125–135.

[6] Bulow, J. and Roberts, J. (1989). The simple economics of optimal auctions. *The Journal of Political Economy*, 97(5):1060–1090.

[7] Celis, L. E., Lewis, G., Mobius, M., and Nazerzadeh, H. (2011). Buy-it-now or take-a-chance: A simple sequential screening mechanism. In *WWW*.

[8] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., and Zemel, R. (2012). Fairness through awareness. In *ITCS*.

[9] Emek, Y., Feldman, M., Gamzu, I., Leme, R. P., and Tennenholtz, M. (2012). Signaling schemes for revenue maximization. In *ACM Conf. on Electronic Commerce*.

[10] Ghosh, A., Mahdian, M., McAfee, R. P., and Vassilvitskii, S. (2012). To match or not to match: Economics of cookie matching in online advertising. In *ACM Conference on Electronic Commerce*.

[11] Hartline, J. D. and Karlin, A. R. (2007). Profit maximization in mechanism design. In Nisan, N., Roughgarden, T., Tardos, É., and Vazirani, V. V., editors, *Algorithmic Game Theory*, chapter 13. Cambridge University Press.

[12] Hartline, J. D. and Roughgarden, T. (2009). Simple versus optimal mechanisms. In *ACM Conference on Electronic Commerce*, pages 225–234.

[13] Levin, J. and Milgrom, P. (2010). Online advertising: Heterogeneity and conflation in market design. *American Economic Review*, 100(2):603–607.

[14] McAfee, R. P. (2012). When does improved targeting increase revenue? Working paper, private communication.

[15] Milgrom, P. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122.

[16] Miltersen, P. B. and Sheffet, O. (2012). Send mixed signals — earn more, work less. In *ACM Conference on Electronic Commerce*.

[17] Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

[18] Palfrey, T. R. (1983). Bundling decisions by a multiproduct monopolist with incomplete information. *Econometrica*, 51(2):463–483.