Ad Auctions with Data

Hu Fu, Patrick Jordan, Mohammad Mahdian, Uri Nadav, Inbal Talgam-Cohen, Sergei Vassilvitskii*

***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, additional data leads to additional revenue. However in simpler auctions, namely the second price auction with reserve prices, there are instances in which additional data decreases the revenue, albeit by only a small constant factor.**

I. 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, 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.

*Hu Fu is with the Department of Computer Science, Cornell University, Ithaca, NY 14853 (hufu@cs.cornell.edu). Work done while author was at Yahoo!.

Patrick Jordan, Uri Nadav and Sergei Vassilvitskii are with Yahoo! Research, Sunnyvale, CA ({prjordan, urinadav, sergei}@yahoo-inc.com).

Mohammad Mahdian is with Google, Inc., Mountain View, CA (mahdian@google.com). Work done while author was at Yahoo!.

Inbal Talgam-Cohen is with the Department of Computer Science, Stanford University, Stanford, CA 94305 (italgam@cs.stanford.edu). Work done while author was at Yahoo!.

While concealing information can only decrease social efficiency, it may be advantageous in terms of revenue, since releasing information may decrease competition and lead to lower revenues. As a concrete 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 hedge and 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; on the other hand, if the gender is kept hidden, the auctioneer will have two bids of \$5 and will collect \$5 in a standard second price auction.[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 [4] 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.

Our Contribution: In this work we show that while revealing information can lead to a decrease in expected revenue, in the context of ad auctions adding reserve prices helps counteract that trend. Specifically, we propose a model for ad auctions with data, and show that in Myerson's optimal auction [15], the expected revenue is guaranteed to increase when information about the ad is revealed. We then ask whether this result requires the full generality of the optimal mechanism, or if simpler reserve-based approaches lead to a revenue increase as well. We show that simply introducing reserve prices is not enough, by demonstrating examples where anonymous or monopoly reserves lead to a decrease in expected revenue for the auctioneer. However, this decrease in revenue is upper bounded by a small multiplicative constant when the bidders' values are drawn from distributions that satisfy a standard regularity condition.

[1] A similar example shows how withholding information can decrease social welfare: if the first advertiser values males at \$2 and females at \$8 and the second advertiser values males at \$8 and females at \$3, when gender is hidden the second advertiser always wins, even though it is more efficient to allocate female users to the first advertiser.

A. Related Work

Auction theory has extensively studied the following scenario: The auctioneer has access to a private source of data about the item; she wishes to maximize her expected revenue by pre-committing to a policy of revealing or concealing data. Two separate effects of revealing data have been identified—the *linkage principle* by Milgrom and Weber [14], [11], and more recently Board's *allocation effect* [4]. The former states that in certain settings, revealing data increases the expected revenue from first price, second price or English auctions. The latter applies to a different but overlapping family of settings, and states that revealing data can either increase or decrease the expected revenue from second price auctions, depending on the number of participating bidders.

In more detail, the linkage principle is relevant to settings with *interdependent values*, in which bidders' estimates of the item's features are correlated, and their values for the item depend not only on their own estimate but on estimates of others as well. The linkage principle says that when bidders are symmetric and their estimates are *affiliated*, the auctioneer can increase her expected revenue by revealing her own estimate of the item. Intuitively, revelation reduces bidders' private information and their ability to transform it into rents. In contrast, the allocation effect and following it our result both apply to settings with *independent private values*. Despite being a special case of interdependent values, the linkage principle has no effect in these settings since revealing the auctioneer's information does not decrease private information. Instead, the allocation effect says that whenever information revelation alters the order of the bidders' values, the expected second price revenue may change - when there are two bidders the auctioneer will be worse off, and when the number of bidders grows asymptotically the effect is reversed (under certain conditions).

We study the effect of revealing information beyond the second price auction and with any number of bidders. Our focus on ad auctions justifies a specific model in which bidders' values are scaled by different constants determined by the given information—a special case of Board's general model in which information can have an arbitrary effect on values. Our analysis uses similar elements to Board's for the two-bidder case, but since we consider Myerson's mechanism, the revenue is determined by the maximum virtual value instead of the minimum value, thus leading to different outcomes.

Additional Related Work: Emek et al. [2] consider information revelation in second price auctions with a general information model. They demonstrate that even when the number of bidders grows asymptotically, in general it may be the case that neither full revelation nor full concealment maximize the expected revenue, and in fact the auctioneer can do better off by a non-constant factor by applying an intermediary revelation policy. They proceed to study the problem of finding the optimal revelation scheme and show that, while computationally hard in general, several cases of interest are polynomially solvable.

Levin and Milgrom [13] highlight the disadvantages of full information revelation from a market design point of view—with too much information every impression is essentially unique, leading to thinner markets that are harder to operate. Several proposed mechanisms address these issues [3], [6]. Dwork et al. [7] discuss important fairness concerns that arise from revealing user data. A separate body of work considers the case in which the bidders and not the auctioneer have private sources of information about the item, resulting in asymmetries among them; a recent example is Abraham et al. [1].

II. Preliminaries

We give a brief description of Myerson's optimal mechanism, which will be heavily used in Section IV. Myerson in his seminal paper [15] showed that, in a single-item, truthful auction where bidders' valuations are drawn independently from known distributions $D_1, \cdots, D_n$, the expected revenue (a.k.a. *virtual surplus*) collected by assigning the item to bidder i with valuation v_i is $\varphi_i(v_i) = v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$, where F_i is the cummulative distribution function of D_i, and f_i is its density function. $\varphi_i(v_i)$ is called the *virtual valuation* of v_i. When $\varphi_i(v_i)$ is a monotone non-decreasing function of v_i, the corresponding distribution is said to be *regular*; in the special case when $\frac{1-F_i(v_i)}{f_i(v_i)}$ is a monotone non-increasing function of v_i it is said to have *monotone hazard rate (MHR)*. When all distributions are regular, an optimal truthful mechanism is simply to assign the item to a bidder with the highest nonnegative virtual valuation. When the distributions are not regular, however, such an allocation is not truthful. For such cases, Myerson showed a procedure called *ironing*, which monotonizes virtual valuations and produces the so-called *ironed virtual valuations* $\{\tilde{\varphi}_i(v_i)\}$. The ironed virtual valuation $\tilde{\varphi}_i(v_i)$ is monotone non-decreasing with v_i for all distributions, and the optimal truthful auction allocates the item to the bidder with the highest nonnegative ironed virtual valuation. The optimal expected revenue is therefore $\mathbb{E}[\max\{0, \tilde{\varphi}(v_1), \ldots, \tilde{\varphi}(v_n)\}]$.

The (ironed) virtual value can be seen as the *marginal revenue*, as pointed out by Bulow and Roberts [5], and it is the derivative of the *(ironed) revenue curve*. Given a 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, as shown in Figure 1, and the ironed revenue curve is the concave hull of this curve, as indicated by the dashed curve in Figure 1. The ironed virtual valuation of v is then $\left.\frac{\mathrm{d}\tilde{R}(q)}{\mathrm{d}q}\right|_{q=1-F(v)}$.

III. Model

In this section, we describe our model for an ad auction with data. The players are the publisher who acts as an auctioneer, and n advertisers, bidding on an opportunity to show an ad to the specific user.

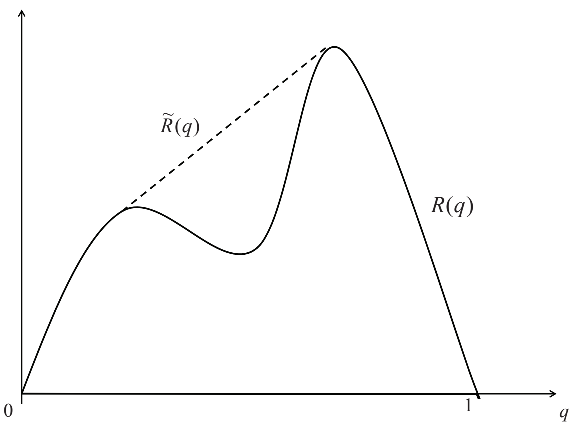

Fig. 1. (Ironed) Revenue Curve

We assume there are m different types of users and that the distribution over user types is publicly known. Formally, every user is characterized by a discrete random variable U drawn from a publicly known distribution F_U on support $\{1, \ldots, m\}$. The auctioneer has access to information about the type of user viewing the impression, i.e., she knows the realization u of U (u is also called the auctioneer's *signal*). In contrast, unless the auctioneer decides to reveal u, the advertisers only know the distribution F_U, and in this case it will be convenient to say that they know the user is of a fictitious *average* type $\bar{u}$.

How do the advertisers value the impression offered to them by the auctioneer? Every advertiser gains a private utility s_i from the event that the user clicks on his ad (s_i is also called advertiser i's signal).[2] We study the Bayesian setting in which s_i is drawn independently at random from a publicly known distribution F_i with density f_i. Advertiser i's value v_i for the impression is thus s_i times the probability of the user clicking on the ad, called the *click through rate (CTR)*.

While the advertiser's value per click, s_i, does not depend on the type of user who clicks on the ad, the CTR is completely determined by the user's type. For every $u \in \{1, \ldots, m\} \cup \{\bar{u}\}$, we denote by $p_{i,u}$ the probability that a user of type u clicks on advertiser i's ad. By definition, $p_{i,\bar{u}} = \mathbb{E}_{u \sim F_U}[p_{i,u}]$. We assume that the auctioneer knows the value of $p_{i,u}$ for every i, u. We can now write advertiser i's value for the impression as $v_i = p_{i,u} s_i$.

Without loss of generality, we focus our attention on direct revelation mechanisms where the bidders directly report their private signals. We require incentive compatibility and individual rationality (IR). The auctions we consider have the following form:

1) The auctioneer commits to whether or not she will use its knowledge of the user's type during the auction.
2) The auctioneer learns the user's type.
3) The advertisers report their values per click $\{s_i\}$.
4) For every i, the auctioneer calculates the value per impression $v_i = p_{i,u} s_i$, where u is either the user's known type or $\bar{u}$ if the type information is ignored.
5) The auctioneer runs a truthful IR mechanism on $\{v_i\}$.

Note that if the advertisers know the CTRs $\{p_{i,u}\}$, the above is equivalent to first letting the auctioneer decide whether or not she will reveal the user's type to the advertisers, and then (after realization and possible revelation of the type) having the advertisers calculate and bid their values v_i and running a truthful IR mechanism. The above mechanism is thus truthful and IR in expectation. In what follows, we say that the auctioneer *reveals the user type*, if she uses it in the auction to calculate the values $\{v_i\}$.

Remark III.1 *We assume the auctioneer uses the true type of the user, and do not address here the interesting but separate question of incentivizing the auctioneer to be honest (e.g., by designing appropriate reputation mechanisms, etc.).*

A. Discussion

Abraham et al. [1] point out that advertisers have their own private data about users, which they store in cookies on the users' computers. However, a prerequisite for using this data is that the advertisers know the user's identity. Our model captures the realistic situation in which the auctioneer has exclusive knowledge of the user's identity. In this situation, the auctioneer can opt not to reveal the user's identity to the advertisers until after the auction has taken place (note that this is a separate decision from whether or not to reveal the user's *type* in the auction).

Modeling the value per impression as the value per click times a type-dependent CTR is compatible with the *pay per click* approach common in practice—the winning advertiser pays only when his ad is clicked, indicating that this is the event to which he attaches value. Our model assumes that the value per click is not affected by the user type and that the auctioneer knows the CTRs; this matches standard assumptions in the context of sponsored search auctions [12]. Note that just like in sponsored search, the auctioneer is in excellent position to learn the CTR values over time by accumulating auction statistics. The assumption that the auctioneer knows the user's type is also reasonable—she has access to demographic information about the user, and can purchase further information such as search history from third-party data providers.

At first it may seem that the dependence of the advertisers' values $\{v_i\}$ on a common random value—the auctioneer's signal indicating the user's type—makes the values themselves common to some extent. However, our model falls within the setting of private and not interdependent values, since the user's type is either made public or not used at all (cf. [4]). Note that this would not be the case if the auctioneer were to use the information about the user asymmetrically, that is, only in the calculation of the values for a subset of the advertisers.

[2] Alternatively, s_i can express advertiser i's value for a *conversion*, the event that the user who clicks on the ad proceeds to make a purchase.

IV. Revenue Monotonicity of the Optimal Mechanism

In this section we show that when Myerson's optimal mechanism is applied to our model, the expected revenue (weakly) increases when the auctioneer fully reveals its information regarding the user's type.

Recall that applying Myerson's mechanism to our setting means that the auctioneer calculates the ironed virtual valuation $\tilde{\varphi}_i(v_i)$ for each bidder given his bid $\{v_i\}$ and then allocates the item to the bidder with the highest non-negative ironed virtual valuation.

Observation IV.1 *Let $u \in \{1, \ldots, m\} \cup \{\bar{u}\}$, $v_i = p_{i,u}s_i$. The distribution of v_i is $F_{i,u}(x) = F_i\left(\frac{x}{p_{i,u}}\right)$ and the corresponding ironed virtual value function is $\tilde{\varphi}_{i,u}(x) = p_{i,u}\tilde{\varphi}_i\left(\frac{x}{p_{i,u}}\right)$.*

The last equality follows by looking at the revenue curves R_i and $R_{i,u}$ corresponding to distributions F_i and $F_{i,u}$ respectively:

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

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

The next observation will be useful in analyzing the expected revenue from Myerson's mechanism with and without revealing information about the user type. By Observation IV.1 and the definition of $p_{i,\bar{u}}$, for every value per click s_i, the ironed virtual value for the impression when the user's type is not used is the expected ironed virtual value when the type is used.

Observation IV.2

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

We now state our main result.

Proposition IV.3 (Revenue Monotonicity) *The expected revenue from Myerson's mechanism when the user's type is revealed is at least as high as the expected revenue when the user's type is not revealed.*

Proof: Myerson proved that the expected revenue of any truthful mechanism is equal to its expected ironed virtual surplus [15] (see also [9, Theorem 13.10]). We use Myerson's result to prove Proposition IV.3 pointwise, i.e., we show that it holds for every fixed profile of values per click $(s_1, \ldots, s_n)$. Taking expectation over the profiles completes the proof.

Fix $(s_1, \ldots, s_n)$ and let $u \in \{1, \ldots, m\}$ be the user's known type. The virtual surplus of Myerson's mechanism when u is revealed is

$$\max\{0, \tilde{\varphi}_{1,u}(p_{1,u}s_1), \ldots, \tilde{\varphi}_{n,u}(p_{n,u}s_n)\}$$

Taking expectation over u gives the expected virtual surplus when the user's type is revealed

$$\mathbb{E}_{u \sim F_U}[\max\{0, \tilde{\varphi}_{1,u}(p_{1,u}s_1), \ldots, \tilde{\varphi}_{n,u}(p_{n,u}s_n)\}] \quad (1)$$

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

$$\max\{0, \tilde{\varphi}_{1,\bar{u}}(p_{1,\bar{u}}s_1), \ldots, \tilde{\varphi}_{n,\bar{u}}(p_{n,\bar{u}}s_n)\}$$

By Observation IV.2, this is equal to

$$\max\ \{0\} \cup \{\mathbb{E}_{u \sim F_U}[\tilde{\varphi}_{i,u}(p_{i,u}s_i)]\}_{i=1}^n. \quad (2)$$

Since $\max$ is a convex function, by Jensen's inequality (1) $\geq$ (2), so revealing the user's type does not reduce the expected revenue. ■

A. Strategic Revelation

Up until now we've considered only two possibilities for the auctioneer—to fully reveal the user's type or to conceal it. However, following Milgrom and Weber [14, Theorem 9] and Emek et al. [8], there are also many intermediate possibilities. Let $r : \{1, \ldots, m\} \to 2^{\{1,\ldots,m\}}$ be a *revelation strategy*, which takes the real user type $u \in \{1, \ldots, m\}$ and outputs a (possibly random) subset of user types $r(u)$.[3] Possible strategies include $r(u) = u$ (full revelation), $r(u) = \{1, \ldots, m\}$ (no revelation), $r(u) \ni u$ (partial revelation), and noisy revelation in which $r(u)$ may not even contain the real type u.

The auction now proceeds as follows. The auctioneer publicly commits (before learning the user's type u) to a revelation strategy r. This strategy, together with the realized subset $r(u)$ and the type distribution F_U, induces a new *ex post* distribution $\tilde{F}_U$ on the user types. In order to maintain incentive compatibility, the auctioneer sets v_i according to this distribution as $\mathbb{E}_{u \sim \tilde{F}_U}[p_{i,u}s_i]$ (equivalently, the auctioneer reveals $r(u)$ to the advertisers and they report their values $\{v_i\}$ where $v_i = \mathbb{E}_{u \sim \tilde{F}_U}[p_{i,u}s_i]$).

A direct corollary of Proposition IV.3 is that the full revelation strategy yields the highest expected revenue of all revelation strategies.

Corollary IV.4 (Full Revelation is Optimal) *For every revelation strategy r, the expected revenue from Myerson's mechanism is upper bounded by the expected revenue when the user's type is fully revealed.*

Proof: Condition on the revealed subset $r(u)$. Together with r and F_U it induces the distribution $\tilde{F}_U$ on the user

[3] We use this definition for concreteness. $r(u)$ can be defined more generally to include reports beyond subsets of types (such as summary statistics etc.) and the result in this subsection holds.

types. We can now apply Proposition IV.3 to conclude that the expected revenue from full revelation of u is at least as high as the expected revenue from revealing $r(u)$. ■

V. Simple Auctions with Reserve Prices

In this section, we present examples of simple auctions with reserve prices in which data release strictly decreases the expected revenue. We look at two types of such auctions: the second price auction with an anonymous reserve price, and the second price auction with monopoly reserve prices. In the former, a single reserve price is applied to all bidders, and those who bid above the reserve price compete in the second price auction. In the latter, a *monopoly reserve price* is applied to each bidder, and bidders who bid above their respective reserves enter the second price auction. A monopoly reserve price for a bidder is the optimal reserve price set in an auction with this bidder alone. Equivalently, it is equal to the value v whose corresponding ironed virtual value $\tilde{\varphi}(v)$ is 0.

Our examples show that releasing data can decrease the expected revenue even when the values are drawn from MHR distributions (a special case of regular distributions). However, for all regular distributions, the simple auctions we consider are guaranteed to give at least a constant factor of the optimal expected revenue [10]. We use this fact to show that since releasing data does not hurt the optimal expected revenue, the loss in expected revenue from data revelation in simple auctions is bounded by a constant factor.

A. Second Price Auctions with Anonymous Reserve

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

The example has two bidders and $m = 2$, with F_U being uniform between 1 and 2. Bidder 1's valuation for a "high" type is uniformly drawn from $[0, 2]$, and for a "low" type is constantly 0. Bidder 2 is not sentitve to the types and her valuation is drawn uniformly from $[0, 1]$ regardless of the item 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 item is a low type, the optimal auction is a second price auction with a reserve price $1/2$, and the revenue is $1/4$. We now compute the optimal anonymous reserve price for a high type and the revenue it generates. When setting a reserve price to be $x \in [0, 1]$, the revenue is

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

To maximize this, we set x to be $3/4$, and the revenue is $\frac{9}{64}+\frac{5}{12}$. Setting a reserve price in $[1, 2]$ does no better (the optimal reserve price in that interval is 1, which generates a revenue of 0.5).

Therefore, for a high type, the revenue of an optimal second price auction with an anonymous reserve price 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 $\frac{5}{12}$.

B. Second Price Auction with Monopoly Reserves

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

As in the previous section, the example is on two bidders and m is set to 2, with F_U being uniform between 1 and 2. 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 type of the item and her valuation is uniformly drawn from $[0, 6]$ regardless of the item type.

When the type is not announced, the optimal auction is a second price auction with reserve price 3, and the expected revenue is 2.5.

When the item is of high type, the monopoly reserves are 4 and 3, respectively. The expected revenue is:

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

When the item is of low type, the monopoly reserves are 2 and 3, respectively. The expected revenue is:

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

Thus when the type is announced, the expected revenue is 2.4722, which is less than 2.5.

C. Upper Bound on Revenue Loss

Theorem V.1 (Hartline and Roughgarden) *For every single-item setting with values drawn independently from regular distributions:*

1) *[10, Theorem 5.1] There is an anonymous reserve price such that the expected revenue of the second price auction with this reserve is a 4-approximation to the optimal expected revenue.*
2) *[10, Theorem 3.7] The expected revenue of the second price auction with monopoly reserves is a 2-approximation to the optimal expected revenue.*

Corollary V.2 *The expected revenue from the second price auction with anonymous reserve (resp., monopoly reserves)* *when the user's type is revealed is a 4-approximation (resp., 2-approximation) to the expected revenue when the user's type is not revealed.*

Proof: By Theorem V.1, the expected revenue from the second price auction with anonymous reserve (resp., monopoly reserves) when the user's type is revealed is a 4-approximation (resp., 2-approximation) to the *optimal* expected revenue when the user's type is revealed, which by Proposition IV.3 is as high as the optimal expected revenue when the type is not revealed. ■

VI. Open Questions

Incorporating data in ad auctions raises many questions of practical importance to which our work may be applicable. We mention two open questions that follow directly from our work. (1) In simple auctions, can an intermediate revelation strategy lead to revenue increase, and if so can the auctioneer efficiently find the optimal such strategy? Note that this question was studied by [8] in a more general setting, however the answer in our restricted model may be different. (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 the challenges associated with analysis in the interdependent model (see [1]). Additional open questions arise from generalizations of our model, e.g., allowing the user type to affect the value per click in addition to the CTR.

References

[1] Ittai Abraham, Susan Athey, Moshe Babaioff, and Michael Grubb. Peaches, lemons, and cookies: Designing auction markets with dispersed information. Manuscript, 2011.

[2] Noga Alon, Michal Feldman, Iftah Gamzu, and Moshe Tennenholtz. Signaling in take-it-or-leave-it sales. Manuscript, 2011.

[3] Marissa Beck and Paul Milgrom. Auctions, adverse selection and internet display advertising. Working paper, 2011.

[4] Simon Board. Revealing information in auctions: the allocation effect. *Econ Theory*, 38:125–135, 2009.

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

[6] L. Elisa Celis, Gregory Lewis, Markus Mobius, and Hamid Nazerzadeh. Buy-it-now or take-a-chance: A simple sequential screening mechanism. In *WWW*, 2011.

[7] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Rich Zemel. Fairness through awareness. In *ITCS*, 2012.

[8] Yuval Emek, Michal Feldman, Iftah Gamzu, and Moshe Tennenholtz. Signaling schemes for revenue maximization. Manuscript, 2011.

[9] Jason D. Hartline and Anna R. Karlin. Profit maximization in mechanism design. In Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors, *Algorithmic Game Theory*, chapter 13. Cambridge University Press, 2007.

[10] Jason D. Hartline and Tim Roughgarden. Simple versus optimal mechanisms. In *ACM Conference on Electronic Commerce*, pages 225–234, 2009.

[11] Vijay Krishna. *Auction Theory*, chapter 7. Elsevier, 2010.

[12] Sebastien Lahaie, David M. Pennock, Amin Saberi, and Rakesh V. Vohra. Sponsored search auctions. In Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors, *Algorithmic Game Theory*, chapter 28. Cambridge University Press, 2007.

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

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

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