Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Incentive-Compatible Learning of Reserve Prices for Repeated Auctions

Yash Kanoria, Hamid Nazerzadeh

informs
http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 69, No. 2, March–April 2021, pp. 509–524
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Crosscutting Areas

Incentive-Compatible Learning of Reserve Prices for Repeated Auctions

Yash Kanoria,[a] **Hamid Nazerzadeh**[b]

[a] Graduate School of Business, Columbia University, New York, New York 10027; [b] Marshall School of Business, University of Southern California, Los Angeles, California 90089

Contact: ykanoria@columbia.edu, https://orcid.org/0000-0002-7221-357X (YK); hamidnz@marshall.usc.edu, https://orcid.org/0000-0002-0433-3318 (HN)

Received: December 10, 2014
Revised: March 20, 2017; January 26, 2019; October 30, 2019
Accepted: February 13, 2020
Published Online in Articles in Advance: February 11, 2021

Subject Classifications: games/group decisions: bidding/auctions
Area of Review: Games, Information, and Networks

https://doi.org/10.1287/opre.2020.2007

Abstract. Large fractions of online advertisements are sold via repeated second-price auctions. In these auctions, the reserve price is the main tool for the auctioneer to boost revenues. In this work, we investigate the following question: how can the auctioneer optimize reserve prices by learning from the previous bids while accounting for the long-term incentives and strategic behavior of the bidders? To this end, we consider a seller who repeatedly sells ex ante identical items via a second-price auction. Buyers' valuations for each item are drawn independently and identically from a distribution F that is unknown to the seller. We find that if the seller attempts to dynamically update a common reserve price based on the bidding history, this creates an incentive for buyers to shade their bids, which can hurt revenue. When there is more than one buyer, incentive compatibility can be restored by using *personalized* reserve prices, where the personal reserve price for each buyer is set using the historical bids of *other* buyers. Such a mechanism asymptotically achieves the expected revenue obtained under the static Myerson optimal auction for F. Further, if valuation distributions differ across bidders, the loss relative to the Myerson benchmark is only quadratic in the size of such differences. We extend our results to a contextual setting where the valuations of the buyers depend on observed features of the items. When up-front fees are permitted, we show how the seller can determine such payments based on the bids of others to obtain an approximately incentive-compatible mechanism that extracts nearly all the surplus.

Funding: This work was supported in part by Microsoft Research New England. The work of H. Nazerzadeh was supported by Google [Faculty Research Award].Y. Kanoria acknowledges the support of the National Science Foundation via grant [CMMI-1653477].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/opre.2020.2007.

Keywords: online advertising • repeated auctions • dynamic mechanism design • reserve price optimization

1. Introduction

Advertising is the main component of the monetization strategy of most internet companies. Large fractions of online advertisements are sold via auctions where advertisers bid in real time for a chance to show their ads to users. Examples of such auction platforms, called advertisement exchanges (Muthukrishnan 2009, McAfee 2011), include Google's Doubleclick (AdX), Facebook, AppNexus, and OpenX.

The second-price auction is a common mechanism used by advertisement exchanges. It is a simple mechanism that incentivizes advertisers to be truthful in a static setting. The second-price auction can maximize the social welfare (i.e., the value created in the system) by allocating the item to the highest bidder.

To maximize the revenue earned in a second-price auction, the auctioneer can set a reserve price and not make any allocations when the bids are low. In fact, under symmetry and regularity assumptions (see Section 2), the second-price auction with an appropriately chosen reserve price is optimal and maximizes revenue among all selling mechanisms (Myerson 1981, Riley and Samuelson 1981).

However, to set the reserve price effectively, the auctioneer requires information about the distribution of the valuations of the bidders. A natural idea, which is widely used in practice, is to construct these distributions using the history of the bids. This approach, although intuitive, raises a major concern with regard to the long-term (dynamic) incentives of the advertisers. Because the bid of an advertiser may determine the price she pays in future auctions, this approach may lead the advertisers to shade their bids and ultimately result in a loss in revenue for the auctioneer.

To understand the effects of changing reserve prices based on previous bids, we study a setting where the auctioneer sells impressions (advertisement space) via repeated second-price auctions. More specifically, in the main model we consider, the valuations of the bidders are drawn independently and identically from a distribution. The bidders are strategic and aim to maximize their cumulative utility. We demonstrate that the long-term incentives of advertisers play an important role in the performance of these repeated auctions.

We show that natural mechanisms that set a common reserve price using the history of the bids may create substantial incentives for the buyers to shade their bids. On the other hand, we propose an *incentive-compatible* mechanism that sets a personal reserve price for each agent based on the previous bids of *other* agents.[1] Our mechanism allocates the item to the highest bidder if his bid exceeds his personal reserve price. If the item is allocated, the price is equal to the maximum of the second-highest bid and the personal reserve price of the winner. This structure corresponds to mechanisms used in practice, as described in Paes Leme et al. (2016). By appropriately choosing the function that maps historical bids of others to a personal reserve price, we show that the expected revenue per round is asymptotically as large as that under the static Myerson optimal auction that a priori knows the distribution of the bids.[2]

We discussed earlier that only using the bids of other buyers has the "first-order effect" of preventing a bidder from lowering the reserve price she will see in the future by misreporting her valuation. However, we show that despite only using bids of other agents, there is room for a "second-order effect" under which a bidder could seek to benefit by affecting the future reserve prices of others and thus, indirectly herself. Hence, importantly, to prevent the second-order effect, our mechanism is "lazy" (see more on this in the section on related work), in that it allocates the item only to the highest bidder (if she exceeds her personal reserve price) and otherwise leaves it unallocated. An "eager" variant would allocate the item to the highest bidder among those who exceed their reserve price; in particular, the eager mechanism would allocate the item as long as *some* bidder exceeds her personal reserve price. The eager approach would create an incentive for agents to overbid so as to increase the personal reserve prices of other agents in the future, thereby increasing the likelihood that those agents are eliminated.[3]

As described earlier, our mechanism allocates the item to the highest bidder if her bid exceeds her personal reserve price. The personal reserve is chosen to maximize revenue for a distribution estimated using other agents' bids. A natural concern with such an approach is that if agents' valuation distributions differ from each other, it may lead to a lower personal reserve price for agents with a higher valuation distribution, and vice versa, thereby hurting revenue. We show that this issue is not significant when differences in valuation distributions are not too large (our notion of the distance between two distributions is the maximum absolute difference between their virtual value functions). In particular, we show that the loss relative to the Myerson benchmark is only quadratic in the size of such differences and supplement this theoretical result with numerical examples.

We also generalize our result along another dimension. Namely, we extend our results to a contextual setting with heterogeneous items that are represented by a feature vector of covariates. The valuations of the buyers are linear in the feature vectors (with a priori unknown coefficients) plus an idiosyncratic private component. We present a learning algorithm that determines the reserve price for each buyer using an ordinary least squares estimator for the vector of feature coefficients. We show that the loss of revenue is sublinear in the number of samples (previous auctions).

For the aforementioned results, we benchmarked the performance of the mechanisms with respect to the static Myerson optimal auction that knows the distribution of the bids in advance. However, we note that this static mechanism is not the optimal mechanism among the class of dynamic mechanisms. In fact, we present a mechanism that can extract (almost all of) the surplus of the agents. The basic idea is that using the bids of other agents, the seller can construct an estimate of the valuation distribution and hence, of the expected utility per round of each agent when individual items are allocated using second-price auctions. Based on this estimate, the mechanism charges a surplus-extracting up-front payment at the beginning of each round. Because agents can influence the up-front payments of other agents, they may have an incentive to overbid so as to eliminate competing agents from future auctions. We propose a solution that asymptotically removes the incentive for agents to deviate from truthfulness: the mechanism *simulates* agents who choose not to pay the entrance fee. We show that under our mechanism, truthfulness constitutes an approximate equilibrium.

In each of our proposed mechanisms, we overcome incentive issues using the same two key ideas: (i) we eliminate incentives for *underbidding* by individually choosing a pricing rule for each agent, based only on the bids of *other* agents, and (ii) we disincentivize *overbidding* by preventing an agent from benefiting from suppressing the participation of other agents by raising the prices they face; this has been achieved in our setting by allocating the item only to the highest bidder in our mechanisms that achieve the Myerson benchmark and by simulating nonparticipating agents in our surplus-extracting mechanism. In a setting where agents' valuation distributions are identical (or similar to each other), this approach enables the seller to obtain as much revenue as if she knew the valuation distribution F while maintaining incentive compatibility. We believe that these design principles should be broadly applicable to overcome the lack of knowledge of F when there is competition between strategic agents/buyers; see Section 9 for further discussion.

1.1. Related Work

In this section, we briefly discuss work closely related to ours along two dimensions, behavior-based (personalized) pricing and reserve pricing optimization for online advertising.

1.1.1. Behavior-Based Pricing. Our work is closely related to the literature on behavior-based pricing strategies where the seller changes the prices for *one* buyer (or a segment of the buyers) based on her previous behavior. For instance, the seller may increase the price after a purchase or reduce the price in the case of no purchase; see Fudenberg and Villas-Boas (2007) and Esteves (2009) for surveys.

The common insight from the literature is that the optimal pricing strategy is to commit to a single price over the length of the horizon (Stokey 1979, Hart and Tirole 1988, Salant 1989). In fact, when customers anticipate a future reduction in prices, dynamic pricing may hurt the seller's revenue (Taylor 2004, Villas-Boas 2004). Similar insights are obtained in environments where the goal is to sell a fixed initial inventory of products to unit-demand buyers who arrive over time (Aviv and Pazgal 2008, Dasu and Tong 2010, Correa et al. 2016, Aviv et al. 2019).

There has been renewed interest in behavior-based pricing strategies, mainly motivated by the development in e-commerce technologies that enable online retailers and other internet companies to determine the price for the buyer based on her previous purchases. Acquisti and Varian (2005) show that when sufficient proportions of customers are myopic or when the valuations of customers increase (by providing enhanced services), dynamic pricing may increase the revenue. Another setting where dynamic pricing can boost the revenue is when the seller is more patient than the buyer and discounts his utility over time at a lower rate than the buyer (Bikhchandani and McCardle 2012, Amin et al. 2013, Mohri and Medina 2014a, Chen and Wang 2016). See Taylor (2004) and Conitzer et al. (2012) for privacy issues and anonymization approaches in this context. In contrast with these works, our focus is on auction environments, and we study the role of competition among strategic bidders who remain in the system over a long horizon. We observe that when there is competition, there is value in personalizing prices, in particular when the valuations are drawn i.i.d. over time. In fact, the seller can extract nearly all the surplus.

The problem of learning the distribution of valuations and optimal pricing has also been studied in the context of revenue management and pricing for markets where each (infinitesimal) buyer does not have an effect on future prices and the demand curve can be learned with near-optimal regret (Baliga and Vohra 2003, Segal 2003, Besbes and Zeevi 2009, Harrison et al. 2012, Wang et al. 2014); see den Boer (2015) for a survey. In this work, we consider a setting where the goal is to learn the optimal reserve price with a small number of strategic and forward-looking buyers with multiunit demand, where the action of each buyer can change the prices in the future.

1.1.2. Reserve Price Optimization. Several recent works have studied reserve price optimization. Most of them focused on algorithmic issues but ignored strategic aspects and incentive-compatibility issues: cf. Cesa-Bianchi et al. (2013, 2015), Mohri and Medina (2014b), Roughgarden and Wang (2016), and Golrezaei et al. (2019). Most closely related to our work is the work by Paes Leme et al. (2016), who compare different generalizations of the second-price auction with a personalized reserve. In their "lazy" version, the item is allocated only to the highest bidder. In their "eager" version, first all the bidders below their personal reserve are eliminated, and then, the item is allocated to the highest surviving bidder. From an optimization/learning perspective, they show that lazy reserves are easy to optimize and A/B test in production, whereas eager reserves lead to higher surplus; however, their optimization is NP-complete, and naive A/B testing leads to incorrect conclusions. The mechanism we propose corresponds to their lazy version. We show how this mechanism—a lazy second-price auction with personalized reserves—can be used to optimize reserve prices in an incentive-compatible way by appropriately learning from the previous bids (the eager version may create incentives to overbid).

Ostrovsky and Schwarz (2009) conducted a large-scale field experiment at Yahoo! and showed that choosing reserve prices guided by the theory of optimal auctions can significantly increase the revenue of sponsored search auctions. To mitigate the aforementioned incentive concerns, they drop the highest bid from each auction when estimating the distribution of the valuations. However, they do not formally discuss the consequences of this approach.

Another common solution offered to mitigate incentive constraints is to bundle a large number of impressions (or keywords) together so that the bid of each advertiser has little impact on the aggregate distribution learned from the history of bids. However, this approach may lead to significant estimation errors because a variety of different types of impressions falls into the same bundle, resulting in a suboptimal choice of reserve price: cf. Epasto et al. (2018). To the best of our knowledge, the present work is the first to rigorously study the long-term and dynamic incentive issues in repeated auctions with dynamic reserves.

1.2. Organization

The rest of the paper is organized as follows. We formally present our model in Section 2. In Section 3, we show that mechanisms that optimize a common reserve price suffer from incentive issues, and this may also hurt revenue. By contrast, in Section 4, we present truthful mechanisms with personal reserve prices, where the reserve prices are optimized based on earlier bids by competing agents. In Section 5, we show that our revenue guarantee is robust to differences in valuation distributions across buyers. Then, in Section 6, we generalize our result to the case of heterogeneous items. Finally, we present a truthful surplus-extracting mechanism in Section 7. Proofs are deferred to the online appendix.

2. Model and Preliminaries

A seller, using a second-price auction, sells items over time to $n \geq 1$ agents. The valuation of agent $i \in \{1, \ldots, n\}$ for an item at time t, denoted by v_{it}, is drawn independently and identically from distribution F. (Later, in Section 5, we will consider the case where different agents have different valuation distributions.) There is exactly one item for sale at each time $t = 1, 2, \cdots$. In Section 6, we extend our results to a contextual setting with heterogeneous items. For the sake of simplicity, we assume that the length of the horizon is infinite and that the seller and the agents aim to maximize their average long-term revenue and utility, respectively. This is a reasonable assumption, given the very large number of impressions sold in practice.

More specifically, the average per-round revenue of the seller, denoted by $\textsc{Rev}$, is equal to

$$\textsc{Rev} = \lim_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{n} p_{it}\right]\right), \tag{1}$$

where p_{it} denotes the payment of agent i at time t. Note that if the limit exists, then the average revenue is maximized. Otherwise, the seller aims to maximize the worst-case average revenue. Similarly, for the average per-round utility of buyer i, denoted by U_i, we have

$$U_i = \lim_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}\left[\sum_{t=1}^{T}(v_{it}q_{it} - p_{it})\right]\right), \tag{2}$$

where $q_{it} = 1$ if the item at time t is allocated to agent i, and otherwise, it is equal to zero. The expectations are with respect to the realizations of the valuations of the agents and any randomization in the mechanism and agent strategies. Each agent aims to maximize the worst-case average utility. The mechanisms we will introduce and the corresponding equilibria/strategies of agents will be stationary in time (after an initial transient), and hence, the aforementioned limits will exist for our mechanisms.

We assume that the valuation distribution F is unknown to the auctioneer/seller, who may not even have a prior on F. The valuation v_{it} of agent i is privately known to agent i. To simplify the presentation, we assume that the valuation distribution F is common knowledge among the agents. (We discuss our informational assumptions later in this section.) We assume that F is a monotone hazard rate (MHR) distribution; that is, the hazard rate $f(v)/(1 - F(v))$ is monotone nondecreasing in v. MHR distributions include all sufficiently light-tailed distributions, including uniform, exponential, and normal. For most of our results, we provide versions that apply to the larger class of *regular* distributions: that is, distributions for which the virtual value function $\phi(v) = v - (1 - F(v))/f(v)$ is monotone increasing in v. (For instance, log-normal distributions are regular but not MHR.)

Let us now consider the seller's problem. The seller aims to maximize his expected revenue via a repeated second-price auction, despite his lack of knowledge of F. He can attempt to do this by dynamically updating the reserve price based on the history of bids so far.

2.1. A "Generic" Dynamic Second-Price Mechanism

At time 0, the auctioneer announces the reserve price function $\Omega : \mathcal{H} \to \mathbb{R}^+$ that maps the history observed by the mechanism to a reserve price. The history observed by the mechanism up to time τ, denoted by $H_{\Omega,\tau} \in \mathcal{H}$, consists of the reserve price, the agents participating in round t and their bids, and the allocation and payments for each round $t < \tau$. More precisely,

$$H_{\Omega,\tau} \triangleq \left\langle (r_1, b_1, q_1, p_1), \cdots, (r_{\tau-1}, b_{\tau-1}, q_{\tau-1}, p_{\tau-1})\right\rangle,$$

where

- r_t is the reserve price at time t.
- $b_t = \langle b_{1t}, \cdots, b_{nt}\rangle$ where b_{it} denotes the bid of agent i at time t.

- q_t corresponds to the allocation vector. If all the bids are smaller than the reserve price r_t, the item is not allocated. Otherwise, the item is allocated to agent $i^\star = \arg\max_i\{b_{it}\}$, and we have $q_{i^\star t} = 1$; in the case of a tie, the item is allocated to a uniformly random agent among those who bid highest. For all the agents who do not receive the item, q_{it} is equal to zero.
- p_t is the vector of payments. If $q_{it} = 0$, then $p_{it} = 0$, and if $q_{it} = 1$, then

$$p_{it} = \max\left\{\max_{j\neq i}\{b_{jt}\}, r_t\right\}.$$

In our notation, Ω specifies a reserve price function for each period t. Note that the auctioneer commits beforehand to a reserve price function Ω. It is well known that in the absence of commitment, the seller earns less revenue (see, e.g., Devanur et al. 2014).

An important subclass of the mechanisms is *static* mechanisms where the reserve price does not depend on the history or time. Another important subclass is *window-based* mechanisms, with window length W, which use only the bids received in the previous W periods to determine the reserve price in the next period.[4] A window-based mechanism is *stationary* if the rule that maps bids in the last W periods to the reserve price in period t does not depend on t. When considering stationary window-based mechanisms, we call the function (a close cousin of Ω) that maps the history of bids in the last W periods to the reserve price in the next period the *reserve optimization function* (ROF).

The seller aims to choose a reserve price function Ω that maximizes the average revenue, defined in Equation (1), when the buyers play an equilibrium with respect to the choice of Ω. To define the utility of the agents and the information available to them, let $H_{i,\tau}$ denote the history observed by agent i up to time τ, consisting of her valuations, bids, allocations, and payments. Namely,

$$H_{i,\tau} = \langle (v_{i1}, b_{i1}, q_{i1}, p_{i1}), \cdots, (v_{i,\tau-1}, b_{i,\tau-1}, q_{i,\tau-1}, p_{i,\tau-1})\rangle.$$

We refer to $H_{i,\tau}$ as the *personal history* of agent i.

We assume that agents do not see the reserve price before they bid[5] but that they know the reserve price function Ω.

The bidding strategy $B_i : \mathcal{H}_i \times \mathbb{R} \to \mathbb{R}$ of agent i maps the valuation of the agent $v_{i\tau}$, the history $H_{i,\tau}$, and the reserve r_τ at time τ to a bid $b_{i\tau}$. Here, $\mathcal{H}_i$ is the set of possible histories observed by agent i.

Finally, we define the history of the game up to time τ as

$$H_\tau = \langle (r_1, v_1, b_1, q_1, p_1), \cdots, (r_{\tau-1}, v_{\tau-1}, b_{\tau-1}, q_{\tau-1}, p_{\tau-1})\rangle.$$

Note that compared with $H_{\Omega,\tau}$, which is the history observed by the seller, H_τ also includes the valuations of the agents.

We say that an agent plays the always truthful strategy, or we simply call the agent *truthful*, if at every time t, we have $b_{it} = v_{it}$ irrespective of the history H_{it} and the reserve r_t. We now formalize our definition of incentive compatibility. We define the *inf-utility* and *sup-utility* of agent i when each agent i' plays strategy $B_{i'}$, respectively, as follows:

$$\underline{U}_i(B_i, B_{-i}) = \liminf_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}\left[\sum_{t=1}^{T} v_{it}q_{it} - p_{it}\right]\right),$$
$$\overline{U}_i(B_i, B_{-i}) = \limsup_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}\left[\sum_{t=1}^{T} v_{it}q_{it} - p_{it}\right]\right).$$

We say that a mechanism is *incentive compatible* (IC) if, for each agent i, other agents are always truthful; then, the inf-utility under the always truthful strategy (weakly) exceeds the sup-utility under any other strategy. Formally, we require

$$\underline{U}_i(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}}) \geq \overline{U}_i(B_i, B_{-i}^{\mathrm{TR}})$$

for any strategy B_i, where B_i^{TR} denotes the truthful strategy. Intuitively, a mechanism is IC if all agents using the always truthful strategy constitute a Nash equilibrium. We emphasize that because our environment and proposed mechanisms are time invariant (after an initial transient), and always truthful is also a time-invariant strategy, the right-hand side of the definition of utility (2) has a limiting value as $T \to \infty$ when all agents are always truthful. More precisely, $U_i(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}})$ is well defined and equal to $\underline{U}_i(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}})$.

The notion of incentive compatibility is static in the sense that the strategies that agents choose before the game starts define an equilibrium. We now define a stronger and dynamic notion. We say that a mechanism is *dynamic incentive compatible* or more precisely, *periodic ex post incentive compatible* if at every time τ, for every history H_τ, each agent i's best-response strategy to her personal history $H_{i,\tau}$ is to be truthful assuming that all the other agents will be truthful in the future (Bergemann and Välimäki 2010). More precisely, define the future inf-utility of an agent as

$$\underline{U}_{i,H_{i,\tau}}(B_i, B_{-i}^{\mathrm{TR}}) = \liminf_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}_{H_{i,\tau}}\left[\sum_{t=\tau}^{T} v_{it}q_{it} - p_{it}\right]\right); \quad (3)$$

that is, it is the (worst-case) future per-auction utility of agent i at time τ, assuming all other agents will be truthful and agent i plays strategy B_i. Again, this limit will exist for the mechanisms we consider when $B_i = B_i^{\mathrm{TR}}$. Also, define the future sup-utility as

$$\overline{U}_{i,H_{i,\tau}}(B_i, B_{-i}^{\mathrm{TR}}) = \limsup_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}_{H_{i,\tau}}\left[\sum_{t=\tau}^{T} v_{it}q_{it} - p_{it}\right]\right). \quad (4)$$

A mechanism is *dynamic incentive compatible* if, for each agent i, we have

$$\underline{U}_{i,H_{i,\tau}}(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}}) \geq \overline{U}_{i,H_{i,\tau}}(B_i, B_{-i}^{\mathrm{TR}})$$

for any time τ, any personal history $H_{i,\tau}$, and any strategy B_i where B_i^{TR} denotes the truthful strategy. As discussed earlier, when all agents follow a truthful strategy in our setting, we have

$$\underline{U}_{i,H_{i,\tau}}(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}}) = U_{i,H_{i,\tau}}\left(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}}\right) = \lim_{T\to\infty}\left(\frac{1}{T}\times \mathrm{E}_{H_{i,\tau}}\left[\sum_{t=\tau}^{T} v_{it}q_{it} - p_{it}\right]\right).$$

In Section 7, we present an approximate notion of dynamic incentive compatibility.

2.2. Discussion on Informational Assumptions

Our results are not sensitive to our informational assumptions. Our main results (Theorems 1, 2, and 4) are for incentive-compatible mechanisms, and hence, they hold even if the agents do not have perfect information regarding the valuation distribution(s) and/or the reserve price function Ω. Theorems 1–4 (and their proofs) remain valid if agents obtain information regarding past reserve prices and past bids, allocations, and payments of other agents.[6] Regarding the seller, our mechanisms use prior-free learning algorithms. Of course, the revenue guarantees remain valid if the seller *does* know something about the valuation distribution(s).

Finally, consider our negative results in Section 3 (Example 1 and Proposition 1 in the online appendix). Providing additional information to agents can only make things (weakly) worse. On the other hand, strategic bid shading by an agent does rely on knowledge of the valuation distribution; note that if an agent initially lacks this knowledge, she can acquire it over time.

2.3. Benchmark

In the first part of the paper, we restrict ourselves to dynamic second-price mechanisms. We use as a benchmark the average revenue that could have been achieved via the optimal static mechanism if F had been known to the seller (i.e., the average revenue per round under the static Myerson auction with the optimal reserve for distribution F). (Note that because F is an MHR distribution, Myerson's result says that the optimal static mechanism is, in fact, a second-price auction with a reserve price. This extends to the case where F is a regular distribution.) Let Rev_* denote the benchmark average revenue. We demonstrate an incentive-compatible second-price mechanism (with personal reserve prices) that asymptotically achieves the benchmark revenue (see Section 4). Later, in Section 7, we go beyond dynamic second-price mechanisms to allow additional mechanism features such as up-front payments. We show how, using a modification of the same ideas, the seller can approximately achieve the largest possible revenue, namely the revenue corresponding to full surplus extraction, while retaining (approximate) incentive compatibility.

3. Incentive Problems with Learning a Common Reserve Price

In this section, we argue that if the seller attempts to learn a common reserve price using historical bids, this leads to incentive issues; specifically, agents may shade their bids in order to reduce the reserve prices they face in the future, and such shading may in turn reduce the revenue earned by the seller.

For simplicity, we start by analyzing a simple reserve price optimization approach, which we call the histogram method, that is the basis of a lot of nonparametric approaches used in practice (see Nazerzadeh et al. 2016) and find significant issues. (In Online Appendix A.2, we argue that these issues are typical in mechanisms that attempt to learn a common reserve price from historical bids.) Throughout this section, we will consider stationary mechanisms and time-invariant strategies, and we look for a nontruthful agent strategy such that if other agents are always truthful, the limiting agent's utility (as defined in (2); note that the limit exists) is strictly larger than that resulting from being always truthful.

3.1. Histogram Method

For simplicity, we consider a setting with $n = 2$ bidders and demonstrate the issue with incentives and the resulting revenue impact. (The problem is even more acute when there is just one buyer/agent, in which case the agent can drive the reserve price, and hence, the seller's revenue, down to zero while still winning the item each time. We comment on this case later.)

Let $\hat{F}_t$ be the joint empirical distribution of all the bids submitted during the last W periods (we will consider the limit $W \to \infty$ in our analysis). The *histogram method* is a window-based stationary second-price mechanism with a very simple ROF. The reserve price at time t is chosen to be the one that maximizes expected revenue when the bid vector is from $\hat{F}_t$. Formally,

$$r_t = \arg\max_r \Big\{ \mathrm{E}_{(b_1,b_2)\sim\hat{F}_t}[\max\{(\max\{b_1,b_2\} - \max\{r, \min\{b_1,b_2\}\}), 0\}] \Big\}; \quad (5)$$

in the case of a tie, r_t is the smallest reserve price in the arg max.

As described in Section 2, the seller allocates the item to a buyer with the highest bid larger than the reserve price. If no bid is above r_t, the item is not allocated. In the case of a tie, the item is allocated at random to one of the highest bidders.

To convey intuition about the incentive issues associated with this approach to reserve price optimization, we start with a simple model. Assume that the valuations of the bidders are drawn i.i.d. from a Uniform$(0,1)$ distribution.

Let us see how an agent may react in response to this mechanism. Intuitively, an agent may want to shade her bid. We present a simple shading strategy where an agent shades her bid if her valuation is between two parameters $\underline{r}$ and $\overline{r}$ and bids truthfully otherwise. More specifically,

- If $0 \le v_t \le \underline{r}$ or $\overline{r} \le v_t \le 1$, then $b_t = v_t$.
- If $\underline{r} < v_t < \overline{r}$, then $b_t = \underline{r}$.

We observe that by playing the strategy, an agent can increase her utility by reducing the reserve, *even if the other agent is truthful*.[7] More importantly, shading can significantly increase the agent's utility. Further, such strategic shading by an agent reduces the revenue of the seller.[8]

Example 1 (Learning a Common Reserve Price Using the Histogram Method Is Not IC)**.** Assume that the valuations of the agents are i.i.d. Uniform$(0,1)$ and that the seller uses histogram-based reserve optimization, using bids from the last W rounds, and consider $W \to \infty$. If one of the agents follows the shading strategy for values of $\underline{r} = 0.378$ and $\overline{r} = 2/3 = 0.667$, whereas the other agent is always truthful, then the reserve price converges to[9] $\underline{r} = 0.378$, with the following consequences.

- Revenue. The limiting average revenue obtained by the seller is close to 0.383. By contrast, if both agents are truthful, the limiting average revenue is equal to $\frac{5}{12} = 0.417$. If the seller does not use any reserve price, the average revenue is equal to $\frac{1}{3} = 0.333$. Therefore, more than 40% of the benefit from reserve price optimization is lost even if one of the agents shades her bid strategically.
- Incentives. The limiting expected utility of the agent from the always truthful strategy is close to 0.083. On the other hand, the limiting expected utility from the aforementioned shading strategy is close to 0.109. Therefore, the agent can increase her utility by more than 30% via shading.

See Online Appendix A.1 for details. A little reflection immediately reveals that in the absence of competition between agents, the incentive issues associated with the histogram method are even more acute.

Remark 1. If, instead, there is only $n = 1$ agent, then the agent can employ the shading strategy with $\underline{r} = \epsilon \in (0, 1/2)$ and $\overline{r} = \infty$ (or equivalently, $\overline{r} = 1$). Under such an agent strategy, as $W \to \infty$, the seller's estimated $\hat{F}_t$ has an atom of mass exceeding 1/2 at ϵ, leading to $\lim_{W\to\infty} r_t = \epsilon$ for all $t > W$. By choosing ϵ close enough to zero, the agent can win the item in almost all rounds while making arbitrarily small payments; thus, this is a best response for the agent as $\epsilon \to 0^+$. The result is that the seller's revenue is vanishing when she uses the histogram method when selling to a single strategic agent.

We note that our example extends to general valuation distributions F.

Remark 2. Although we fixed F to Uniform$(0,1)$ in Example 1, the idea easily extends to general regular F with continuous density f, when the seller sets the reserve using the histogram method. Let r_* be the Myerson optimal reserve price, and consider the static mechanism that uses common reserve price r_* in each round. (As the window length $W \to \infty$, under truthful bidding by all agents, the reserve price set by the seller converges to r_*.) Suppose that agents other than i are always truthful. Then, for sufficiently small ϵ, a shading strategy based on $\underline{r} = r_* - \epsilon$ and $\overline{r} = r_* + 1.1\epsilon$ constitutes a profitable deviation for agent i, by causing the seller to set a reserve price of $\underline{r} = r_* - \epsilon$ with high probability in steady state. Such shading leads to a myopic loss of $O(\epsilon^2)$ in expected utility from the current round—because of losing the item, although i would have won it under truthful bidding—which occurs with probability $O(\epsilon)$ and causes a loss of $O(\epsilon)$ in utility in each case. However, there is a (larger) $\Omega(\epsilon)$ increase in expected utility because of the reserve price being lower by ϵ because of bid shading in the past. This bid shading by agent i causes a loss of $\Omega(\epsilon)$ in revenue for the seller.

In Online Appendix A.2, we show that incentive concerns apply not just to the histogram method but to a broad class of dynamic reserve price mechanisms that set a common reserve price based on historical bids.

4. Incentive-Compatible Optimization of Personal Reserve Prices

In the previous section, we identified significant incentive concerns associated with optimizing the reserve price when all agents face exactly the same reserve price and bidders are strategic. Specifically, natural mechanisms for optimizing the reserve price that are based on historical bids encourage bidders to shade their bids, which in turn, reduces the revenue earned by the seller.

In this section, we present a mechanism that eliminates incentives for agents to misreport their valuations. As mentioned earlier, we overcome incentive issues using two key ideas: (i) we *personalize* reserve prices by choosing a pricing rule[10] for each agent, based only on the bids of *other* agents (hence, agents do not benefit from underbidding); and (ii) we do so by allocating the item only to the highest bidder (hence, agents do not benefit from overbidding so as to prevent others from participating in future auctions).

4.1. Highest-Only Self-Excluding Reserve Price Mechanisms

A second-price auction with personal reserve prices is a highest-only self-excluding reserve price (HO-SERP) mechanism if it satisfies the following two properties.

• Highest only: The mechanism allocates the item only to the highest bidder. If the highest bidder i does not meet his reserve price ($b_{it} < r_{it}$), then the item is not allocated. If $b_{it} \geq r_{it}$, the highest bidder i is charged a price equal to $\max\{r_{it}, \max_{j \neq i}\{b_{jt}\}\}$.

• Self-excluding reserve price (SERP): The reserve price for agent i is determined using only the bids of other bidders and does not depend on the bids of agent i herself.[11] Let $\hat{F}_{-i}$ be the empirical distribution of the bids by agents other than i in the relevant rounds (a window-based mechanism will consider the last W rounds).[12] Then, the personal reserve price r_{it} of agent i is set based on $\hat{F}_{-i}$.

Personal reserve prices may appear more complex than a mechanism with a common reserve price. However, we note that they have been widely used in practice (see, e.g., see Paes Leme et al. 2016). Moreover, we establish strong incentive properties for HO-SERP mechanisms.

To (approximately) maximize the revenue earned, we set r_{it} to be the optimal monopoly price for costless goods when buyers have this valuation distribution: that is,[13]

$$r_{it} = \arg\max_{r} \; r\left(1 - \hat{F}_{-i}(r)\right). \qquad (6)$$

This allows us to approximately achieve the revenue benchmark. The latter is proved using convergence rate bounds from Dhangwatnotai et al. (2015); other related papers on learning the optimal reserve price from samples include Cole and Roughgarden (2014), Huang et al. (2015), and Devanur et al. (2016).

Theorem 1. *Any HO-SERP mechanism is periodic ex post incentive compatible. In particular, all agents following the always truthful strategy constitute an equilibrium. Further, there exists $C < \infty$ that does not depend on the valuation distribution F, such that for any F that is MHR and any $\epsilon \in (0,1)$, the HO-SERP mechanism with window length $W \geq C\log(1/\epsilon)/\epsilon^2$ and personal reserve prices set as per (6) achieves an average per-round revenue that is at least $(1-\epsilon)\textsc{Rev}_*$, where $\textsc{Rev}_*$ is the expected revenue under the optimal static mechanism (i.e., the second-price auction with a Myerson-optimal reserve price).*

Theorem 1 is proved in Online Appendix B. The rapid decay of revenue loss with window length W suggests that our approach should do well with as few as thousands of items/impressions. We remark that a similar result can be established under the weaker requirement of a *regular* valuation distribution F, for a window length bounded as[14] $W \geq C\log(1/\epsilon)/\epsilon^3$.

In Online Appendix B, we provide a finite horizon version of Theorem 2 (Corollary 1 in the online appendix), showing that the revenue loss under our HO-SERP mechanism (using all samples so far) is $O(\sqrt{T}\log T)$ over a horizon of length T for MHR F. We further show that the revenue loss under our mechanism is lower bounded by $\Omega(T^{1/3-\epsilon})$ (Theorem 5 in the online appendix) for a standard (exponential) distribution. The lower bound in our key supporting lemma (Lemma 4 in the online appendix, leading to Theorem 5 in the online appendix) contributes to the agenda pursued in Dhangwatnotai et al. (2015) regarding choosing a price to optimize revenue based on a limited number of samples from the valuation distribution and may be of independent interest.

Note that the HO-SERP mechanism makes use of all bids by agents other than i, including those that do not exceed that agent's own reserve price. However, unlike static settings, using other agents' bids to determine the payments may not be enough to yield *robust* incentive compatibility. Whereas truthfulness is a best response when agent i's valuation $v_{it} < r_{it}$, it is also a best response to submit any other bid $b_{it} \in [0, r_{it})$. In order to make truthfulness the unique best-response strategy, we can tweak the mechanism such that with a small probability γ in each round, all the reserve prices are set to zero, i.i.d. across rounds.[15] Agents are told of this tweak, but they do not know if the reserve prices are zero in the *current* round at the time they submit their bids. This makes truthfulness the unique dominant strategy best response to other agents following the always truthful strategy, a more robust form of incentive compatibility. The loss in expected revenue because of occasionally setting the reserve prices to zero is bounded from above by a γ fraction of the benchmark.

A seeming disadvantage of the HO-SERP mechanism is that if the highest bidder's valuation does not exceed her reserve price, the item goes unallocated even though there may be other bidders who exceed their reserve prices. (The reserve prices may differ from each other because of statistical variation and/or differences in valuation distributions across bidders.) An "eager" variation of self-excluding reserve price is the following: allocate the item to the highest bidder among all the agents whose bids "survive" by being above their personal reserve price,[16] and charge her the larger of her personal reserve price and the second-highest surviving bid. Unfortunately, this variation of the SERP auction creates an incentive to deviate from truthfulness. The intuition is that an agent can benefit from increasing the likelihood that a competing agent is eliminated (because of bidding below her personal reserve price), and this creates an incentive to overbid so as to raise the personal reserve price faced by competing agents in the future. The example illustrates this phenomenon.

Example 2 ("Eager" SERP Is Not IC). Let us consider a setting with two agents whose valuations are drawn i.i.d. from uniform distribution $[0, 1]$. The item is allocated as follows. First, remove all the agents whose bid is less than their personal reserve price, set as per (6). If no agents remain, the item will not be allocated. If only one agent survives, the item will be allocated at a price equal to her personal reserve price. If two agents remain, the item will be allocated to the highest agent at a price equal to the maximum of her personal reserve price and the other bid.
Suppose that the first agent is truthful. We present a profitable deviation for the second agent as follows:

$$b_{it} = \begin{cases} v_{it} & 0 \le v_{it} < \frac{1}{2} \\ 1 & \frac{1}{2} \le v_{it} \le 1. \end{cases}$$

Note that the second agent overbids if her valuation is larger than $\frac{1}{2}$ and is truthful otherwise. Hence, the limiting reserve price for the first agent is equal to one. Therefore, the first agent would be eliminated from all the auctions. In the online appendix, we present a family of profitable deviation strategies including this one and show that the expected per-round utility of the second agent will be increased by $\frac{1}{24}$ under the strategy. In other words, the second agent can increase her utility by 50% because her utility under the truthful strategy is equal to $\frac{1}{12}$.

In the next section, we will show the robustness of the near optimality of the HO-SERP mechanism to small differences in the valuation distributions across agents. In particular, this will imply that revenue losses caused by the item going unallocated, even though some agents exceed their reserve price, are small in expectation when valuation distributions are similar across agents.

5. Robustness to Asymmetry Among Bidders

We have so far assumed that the agents have the same distributions of valuations. In this section, we discuss the robustness of our results with respect to the asymmetry among bidders. We first would like to note that when the valuations are heterogeneous, then the second-price auction, even with optimized personal reserve prices, may not be the optimal static mechanism and that the revenue-maximizing Myerson auction takes a somewhat more complicated form when the item is allocated to an agent with the highest virtual value.[17] Nevertheless, we show that when agents have different valuation distributions, the loss in limiting revenue per round of HO-SERP compared with the static Myerson optimal auction can be bounded.

Consider a case with two agents i and j. Suppose that agent i has a higher valuation distribution than that of j, in the sense that the optimal monopoly price for F_i is larger than the optimal monopoly price for F_j. Then, an HO-SERP mechanism with personal reserve prices set as per (6) sets $r_j > r_i$ instead. As a result, losses are incurred for two reasons. (i) The reserve price of each agent is not suitable for the valuation distribution of that agent. (Further, the static Myerson optimal auction allocates to the bidder with the highest *virtual* value,[18] which HO-SERP does not do.) (ii) The fact that the reserve prices of the two agents are different from each other means that the realized pair of valuations in a round could be such that $r_j > v_j > v_i > r_i$. If this occurs, the item is not allocated because the highest valuation agent (agent j) fell short of her reserve price, although a different agent (agent i) did exceed his reserve price. In this section, we will show that if the valuation distributions are not too different from each other, the loss in revenue under our mechanism relative to the static Myerson optimal auction is small (specifically, it is quadratic in the size of the difference between valuation distributions).

Consider a setting with two agents whose valuation distributions are δ different from each other. (We formally define a notion of distance.) We claim that the loss in revenue, relative to repeating the Myerson optimal mechanism for known valuation distributions, is typically $O(\delta^2)$. The rough reason is that each of the two problems causes a loss of this order. Having a reserve price for each agent that is wrong by $O(\delta)$, or related to this, *not* mapping the reported valuation appropriately to a virtual value, causes a loss of order $O(\delta^2)$ because we are at a distance $O(\delta)$

from the global maximum of a well-behaved optimization problem. The chance that i exceeds his reserve price but the item is not allocated to anyone is bounded by

$$\Pr(r_2 > v_2 > v_1 > r_1) < \Pr(v_1 \in (r_1, r_2) \text{ AND } v_2 \in (r_1, r_2)) = O(\delta) \cdot O(\delta) = O(\delta^2).$$

Hence, this issue also causes a loss of order $O(\delta^2)$.

Let us begin with an example before we make this rigorous.

Example 3. Suppose that agent 1 has a Uniform$(0, 1)$ valuation distribution, whereas agent 2 has a Uniform $(\delta, 1+\delta)$ valuation distribution for some (small) $\delta > 0$. Then, the mechanism we introduced sets $r_1 = (1+\delta)/2$ and $r_2 = 1/2$. The expected revenue it earns is

$$\begin{aligned} &E\big[\text{Revenue from agent 1}\big] + E\big[\text{Revenue from agent 2}\big] \\ &\quad = \frac{5 - 6\delta - 3\delta^3 + 4\delta^3}{24} + \frac{5 + 15\delta}{24} \\ &\quad = \frac{10 + 9\delta - 3\delta^2 + 4\delta^3}{24}. \end{aligned}$$

On the other hand, consider the Myerson optimal mechanism that uses virtual values $\phi_1(v_1) = 2v_1 - 1$ and $\phi_2(v_2) = 2v_2 - 1 - \delta$, allocates the item to the agent with the highest virtual value, if it is positive, and charges that agent the smallest bid/valuation for which she would still have been awarded the item. This mechanism produces revenue $\frac{10+9\delta+3\delta^2+3\delta^3}{24}$. It follows that the revenue under the Myerson optimal mechanism is $\frac{6\delta^2-\delta^3}{24} = O(\delta^2)$ more than that under our mechanism. For $\delta = 0.1$, the revenue loss is just 0.0025 or 0.54%; for $\delta = 0.2$, the revenue loss is just 0.0097 or 1.9%, and even for large $\delta = 0.3$, the revenue loss is 0.021 or 3.9%.

We now formalize this. Let agent i have valuation distribution F_i, which is once again assumed to be MHR (i.e., to have an increasing hazard rate).[19] We define the distance $\|F_i - F_j\|$ between distributions F_i and F_j as

$$\|F_i - F_j\| = \max_v \; |\phi_i(v) - \phi_j(v)|, \tag{7}$$

where $\phi_i(v) = v - (1 - F_i(v))/f_i(v)$ is the virtual value function.[20]

Theorem 2. *Consider a setting with n agents where agent i's valuation distribution is F_i. Again, any* HO-SERP *mechanism is periodic ex post incentive compatible. Suppose that for each agent i, the valuation distribution F_i is MHR and has density bounded above by $f_{\max}$. Suppose also that $\|F_i - F_j\| = \delta$ for all pairs of agents i and j, for some $\delta < \infty$. We have that the limiting average per-round revenue under HO-SERP with personal reserve prices set as per (6) is at least* $\text{REV}_* - 2(n-1)f_{\max}\delta^2$ *as $W \to \infty$, where* REV_* *is the expected revenue achieved by the static Myerson optimal auction. Equivalently, HO-SERP with these reserve prices achieves a fraction $(1 - 2(n-1)f_{\max}\delta^2/\text{REV}_*)$ of the benchmark revenue in the limit $W \to \infty$.*

Thus, if agent valuation distributions are not too different from each other, our proposed mechanism approximately achieves the benchmark revenue. The proof (see Online Appendix C) formalizes the intuition by using Myerson's lemma (Myerson 1981), which says that the expected revenue of a truthful mechanism is equal to the expected virtual value of the winning bidder (defined as zero if the item is not allocated). The revenue-maximizing static mechanism allocates to the bidder with the largest virtual value, if this virtual value is nonnegative. We show that our mechanism deviates from this allocation with probability no more than $2(n-1)f_{\max}\delta = O(\delta)$ and further chooses an allocation that is within δ of the ideal allocation in terms of virtual value in cases where it allocates wrongly. These bounds then enable us to obtain a $2(n-1)f_{\max}\delta^2 = O(\delta^2)$ bound on the loss in expected revenue.

As an illustration, we can apply this result to the setting in Example 3. We have $n = 2$, $\bar{f} = 1$, and $\|F_1 - F_2\| = \delta$, and so, we obtain from Theorem 2 that the revenue loss relative to the Myerson benchmark is bounded above by $2\delta^2$. The actual loss turns out to be $\frac{6\delta^2-\delta^3}{24}$.

We conclude this section with a discussion of a byproduct of our results that could be of independent interest. Hartline and Roughgarden (2009) show that where the valuation of each agent is drawn independently from a different regular distribution, second-price auctions with personalized reserve prices obtain a $\frac{1}{2}$ approximation of the optimal revenue. As a corollary of the analysis leading to Theorem 2, we obtain a complementary result: namely, that using a second-price auction in the asymmetric valuations case, the seller can obtain an expected revenue within $O(\delta^2)$ of the optimal, where δ is the maximum "distance" between valuation distributions; see Remark 3 in Online Appendix C for details.

6. Heterogeneous Items

In this section, we provide guidance on how heterogeneity between items can be incorporated into our proposed HO-SERP mechanism described in Theorem 1. (A similar approach can be used to extend the surplus-extracting self-excluding (SESE) mechanism from Section 7 to a heterogeneous items setting.) Our model of valuations may be interpreted as one way to incorporate correlation between agents' valuations for an item; see McAfee and Vincent (1992).

We generalize the model in Section 2 as follows. Each item has m attributes, where m is a fixed constant.

We denote attributes of the tth item by $x_t = (x_{t1}, x_{t2}, \ldots, x_{tm})^T$ and henceforth, call x_t the *context* in period t. We model the valuation v_i of each agent i for the tth item as

$$v_i = \beta^T x_t + \tilde{v}_i, \tag{8}$$

where $\tilde{v}_i \sim F$ is drawn independently across agents and items, and $\beta \in \mathbb{R}^m$ is the vector of context coefficients (common across agents and items). Thus, the context causes an additive translation in valuations; the amount of translation has a linear functional form in the attributes and is common across agents. We assume that the contexts $(x_t)_{t=1}^{\infty}$ are drawn i.i.d. from some distribution G. Our technical development in this section draws upon the work of Golrezaei et al. (2019) on contextual auctions. Two key high-level differences from that paper are as follows. (i) Our agents are patient, and hence, to obtain good incentive properties, we stay with our proposal to choose a personal price for agent i based on the past bids of *other* agents. In the aforementioned paper, agents are impatient, and so, the mechanism is able to set a personal price for i based on the past bids of agent i herself. (ii) We assume that valuation distribution F is time invariant and obtain revenue guarantees for any F in a class $\mathcal{F}$, whereas the aforementioned paper solves a robust optimization problem where F can vary arbitrarily over time within $\mathcal{F}$.

6.1. Assumptions on F, G, and β

We assume that F is MHR as before but in addition, assume that F has bounded support $(-B_F, B_F)$ for some $B_F < \infty$. We absorb the mean of distribution F into $\beta^T x_\tau$ (we can include an attribute that always takes the value 1 so that its coefficient will be the intercept that includes the mean of F) and hence, assume that $\mathbb{E}_{\tilde{v}\sim F}[\tilde{v}] = 0$. (This implies $\mathrm{E}[v_{i'\tau}] = \beta^T x_\tau$.) We assume that distribution G has bounded support; without loss of generality, we assume that it is supported on $\{x : \|x\| \le 1\}$ (we use the Euclidean norm throughout). We further assume that G has a second-moment matrix $\Sigma = \mathbb{E}_{x\sim G}[xx^T]$ that is strictly positive definite with a smallest eigenvalue at least $1/B_\Sigma$ for some $B_\Sigma < \infty$. We also assume that $\|\beta\| \le B_\beta$ for some $B_\beta < \infty$.

6.2. Auctioneer's Knowledge

The auctioneer observes the context x_t before each period t and knows the bounds B_F, B_Σ, and B_β but does not know F, G, or β beforehand. As before, the auctioneer wants to set personal reserve prices to maximize long-run average revenue (1) while accounting for the strategic response of the agents who aim to maximize the average per-round utility (2). (All expectations now include expectation over the contexts $x_t \sim G$ i.i.d. across periods t.)

The reserve price function $\Omega : \mathcal{H} \to \mathbb{R}^n$ maps the history observed by the mechanism up to t, including the current context x_t to reserve prices $(r_{it})_{i=1}^n$. The history observed by the mechanism, and by each agent, up to time t now includes the contexts in each period so far including period t:

$$H_{\Omega,t} = \langle (x_1, r_1, b_1, q_1, p_1), \cdots, (x_{t-1}, r_{t-1}, b_{t-1}, q_{t-1}, p_{t-1}), x_t \rangle, \tag{9}$$

$$H_{i,t} = \langle (x_1, v_{i1}, b_{i1}, q_{i1}, p_{i1}), \cdots, (x_{t-1}, v_{i,t-1}, b_{i,t-1}, q_{i,t-1}, p_{i,t-1}), x_t \rangle. \tag{10}$$

We include x_t in $H_{i,t}$ to clarify that the agents know the current context x_t before they submit their bids b_t. The auctioneer commits beforehand to Ω. The definition of dynamic incentive compatibility remains as before. As in Theorem 1, we will provide a stationary window-based Ω with good properties and average revenue approaching that under the static Myerson auction. Note that the reserve price of the benchmark static Myerson auction will be dependent on the context, and correspondingly, the reserve price set by our mechanism in period t will account for x_t.

Let $\mathrm{Rev}_*(x)$ be, for context $x \in \mathbb{R}^m$, the expected revenue under the Myerson optimal auction for agents with valuations drawn i.i.d. from the contextual valuation distribution F^x given by

$$F^x(v) \triangleq F(v - \beta^T x). \tag{11}$$

In an effort to obtain a revenue close to $\mathrm{Rev}_*(x_t)$ in period t, but while retaining incentive compatibility, our proposed mechanism proceeds as follows. Fix the window length W. For every agent i, the mechanism sets the personal price r_{it} using (i) the current context x_t and (ii) the bids of other agents in the last W periods (treating those bids as truthful) and the contexts in those periods.

To properly account for the context x_t in the choice of reserve prices r_{it}, our mechanism needs to learn the coefficient vector β. We proceed as follows. Recall that the expected valuation of item τ is $\mathrm{E}[v_{i'\tau}] = \beta^T x_\tau$, allowing us to treat each period τ bid $b_{i'\tau} = v_{i'\tau}$ by agent $i' \neq i$ as a noisy observation of $\beta^T x_\tau$, corrupted by zero mean "noise" $\tilde{v}_{i'\tau} \sim F$ that is i.i.d. across agents and periods. We use these observed bids to obtain an ordinary least squares estimate of β:

$$\hat{\beta}_{-i} \triangleq \operatorname*{argmin}_{\tilde{\beta}:\|\tilde{\beta}\|\le B_\beta} \mathcal{L}_{-i}(\tilde{\beta}), \quad \text{where}$$

$$\mathcal{L}_{-i}(\tilde{\beta}) \triangleq \frac{1}{(n-1)W} \sum_{i'\neq i} \sum_{\tau=t-W}^{t-1} \left(b_{i'\tau} - \tilde{\beta}^T x_\tau\right)^2.$$

This estimate converges rapidly to the true β.

Operations Research, 2021, vol. 69, no. 2, pp. 509–524, © 2021 INFORMS

Lemma 1. *Fix the constants m, B_F, B_Σ, and B_β. There exists a constant $C_1 = C_1(m, B_F, B_\Sigma, B_\beta) < \infty$ such that, for any F, G, and β (satisfying B_F, B_Σ, and B_β, respectively), for any window length $W > 1$, and each agent i, the estimated coefficients are close to the true ones: with probability $1 - 1/W$, we have*

$$\|\hat{\beta}_{-i} - \beta\| \le C_1\sqrt{\frac{\log W}{W}}. \tag{12}$$

We then deploy this estimate to "translate" the past bids to the current context x_t: a bid of $b_{i'\tau}$ submitted under context x_τ maps to the translated bid $\tilde{b}_{i'\tau,-i} \triangleq b_{i'\tau} + \hat{\beta}^T_{-i}(x_t - x_\tau)$. The empirical distribution $\hat{F}^{x_t}_{-i}$ of translated bids serves as an estimate of the true contextual distribution

$$F^{x_t}(v) \triangleq F\big(v - \beta^T x_t\big). \tag{13}$$

We need to be careful here because our estimate of $\hat{F}^{x_t}_{-i}(v)$ is imperfect for two reasons. First, as in Section 4, it is based on a finite number of samples. Second (and this is an issue we did not encounter before), our estimate $\hat{\beta}_{-i}$ is imperfect. As a result, the samples upon which $\hat{F}^{x_t}_{-i}$ is based are not drawn from F^{x_t} itself: instead, the samples based on bids in period τ correspond to sampling from F^{x_t} and then adding $(\hat{\beta}_{-i} - \beta)^T(x_t - x_\tau)$ to the realization. To ensure that this additional source of error does not inadvertently lead to a large reduction in the probability of selling (e.g., this could happen if F^{x_t} has an atom at the Myerson optimal reserve price), our mechanism sets the price by making a small reduction to the estimated optimal reserve price. Accordingly, we set the personal reserve price as per the following modification of (6):

$$r_{it} = -\delta + \arg\max_r \; r\big(1 - \hat{F}^{x_t}_{-i}(r)\big). \tag{14}$$

Here, we set $\delta = 2C_1\sqrt{\log W/W}$, where C_1 is the constant in Lemma 1. Informally, this is to ensure that with probability $1 - 1/W$, errors in bid translation do not cause us to unintentionally price out an agent. As a result of this adjustment to the mechanism, we now obtain an additive approximation to the revenue instead of a multiplicative approximation.[21]

Theorem 3. *Consider the setting with item attributes described, with constants n, m, B_F, B_Σ, and B_β. Any HO-SERP mechanism is periodic ex post incentive compatible. In particular, all agents following the always truthful strategy constitute an equilibrium. Further, there exists $C = C(n, m, B_F, B_\Sigma, B_\beta) < \infty$, such that for any F that is MHR, G, and β (satisfying B_F, B_Σ, and B_β, respectively), any $\epsilon \in (0,1)$, and any context $x_t : \|x_t\| \le 1$, the HO-SERP mechanism with window length $W \ge C\log(1/\epsilon)/\epsilon^2$ and personal reserve prices set as per (14) achieves an expected[22] revenue in period t that is at least $\text{Rev}_*(x_t) - \epsilon$, where $\text{Rev}_*(x_t)$ is the expected revenue under the optimal static mechanism (a second-price auction with the Myerson optimal reserve price) for the true bid distribution F^{x_t} given by (13).*

The proofs for this section are presented in Online Appendix D.

7. Incentive-Compatible Surplus Extraction

Although the second-price auction can be revenue maximizing in static settings, it may not be the optimal mechanism in dynamic environments. To convey intuition, let us first consider a setting with n agents and a horizon of length T where the seller *knows the distribution of the valuations* of agents. Consider the following mechanism. (i) The mechanism charges each agent i an *up-front payment* equal to $\sum_{t=1}^T \mathrm{E}[u_{it}]$, where u_{it} denotes the random variable corresponding to the utility of agent i at time t: namely,

$$u_{it} = \max\Big\{v_{it} - \max_{j\ne i}\{b_{jt}\}, 0\Big\}. \tag{15}$$

The expectation is calculated assuming that all agents are truthful. (ii) The mechanism runs a second-price auction (with no reserve) in each of the T rounds. Notice that Equation (15) is consistent with this design.

Note that by using the up-front payments, the mechanism extracts the whole surplus of the buyers and obtains an average revenue of $\mathrm{E}[\max_j\{v_{jt}\}]$. Assuming only individual rationality on the part of the agents, this is the maximum-achievable average revenue per round for any mechanism. This mechanism, although revenue optimal, is not directly applicable to the current online ad markets because it charges an up-front payment; see Mirrokni and Nazerzadeh (2017).[23] However, ignoring this practical consideration, we show how the ideas can be used to design an essentially optimal mechanism in our setting.

The surplus-extracting mechanism can also be implemented as follows (when the distribution of the valuations, F, is known): see Arrow (1979), d'Aspremont and Gérard-Varet (1979), Baron and Besanko (1984), and Ëso and Szentes (2007). In each round t, the mechanism charges an entrance fee of

$$\mu_i = \mathrm{E}_F[u_i] = \mathrm{E}_F\bigg[\max\Big\{v_{it} - \max_{j\ne i}\{v_{jt}\}, 0\Big\}\bigg]. \tag{16}$$

The agent may accept the entrance fee. Agents who pay the entrance fee then learn their valuation v_i and can bid in the auction. The item is allocated via a second-price auction with no reserve, and therefore, the agents will bid truthfully. Note that in the desired equilibrium, the agents are indifferent between participating or leaving, but the mechanism can always nudge the agents to participate by slightly reducing the entrance fee. Building on these ideas, we propose the following mechanism.

7.1. SESE Mechanism

The mechanism consists of two phases.

- In the first phase, which lasts for N rounds (where N is a parameter chosen by the seller), the item is allocated via a second-price auction with no reserve. At the end of the first phase, for each agent i, define $\hat{\mu}_i$ as follows:

$$\hat{\mu}_i = \frac{1}{n}\frac{1}{N/2}\sum_{k=1}^{N/2} z_k, \tag{17}$$

and z_k's for $1 \le k \le N/2$ are constructed as follows. We repeatedly sample *without replacement* n bids from the set of bids in the first phase from all the bidders except agent i. Let Z_k be the kth sampled set, and let z_k be the difference between the highest and the second-highest bid in Z_k. Note that because $n \ge 2$, the total number of sampled bids is $nN/2 \le (n-1)N$, ensuring feasibility.

- In each round $t > N$ in the second phase, the seller offers an entrance fee of $(\hat{\mu}_i - \sqrt{\frac{2\log N}{N}})$ to agent i. Note that the entrance fee is determined using the *other* agents' bids in the first phase.

The item is allocated using a second-price auction with no reserve. Let S be the set of agents who pay the entrance fee (and subsequently learn their valuation v_{it}), and let $\bar{S}$ represent the set of agents who refuse to participate in this round. The mechanism *simulates* the agents in $\bar{S}$. More specifically, the mechanism randomly chooses a round $\tau < N$ and uses the bids in that round for each agent $j \in \bar{S}$. At time t, if a simulated bid is the highest, the item will not be allocated. Otherwise, it will go to the highest bidder at the price equal to the second-highest bid among agents in *S and $\bar{S}$*.

Here is the intuition behind the mechanism. Observe that by the definition, when all the agents are truthful and have the same valuation distribution, we have

$$\mathrm{E}[z_k] = \mathrm{E}\left[\sum_{i=1}^{n} u_{it}\right] = \sum_{i=1}^{n} \mu_i = n\mu_i, \tag{18}$$

where μ_i denotes $\mathrm{E}[u_{it}]$ for agent i; see Equation (16). Hence, we have $\mathrm{E}[\hat{\mu}_i] = \mu_i$.

Our mechanism achieves (approximate) incentive compatibility by leveraging the same two key ideas that led to Theorems 1 and 3. (i) The entrance fee charged to each agent (in the second phase) depends only on the bids of the *other* agents in the first phase; thus, an agent's bids do not affect the entrance fee that the agent herself faces. We further deduce that the agents bid truthfully in the second phase because their bids have no future impact whatsoever. Hence, they would pay the entrance fee if[24] $\mathrm{E}[u_i] \ge \hat{\mu}_i - \sqrt{\frac{2\log N}{N}}$. (ii) Using simulated bids, we bound the gain from overbidding for the agents: note that the bids of the agents in the first phase can influence the outcomes in the second phase. More specifically, agents can overbid and inflate the entrance fee of other agents, which may result in the latter's refusal to participate in the auctions in the second phase. Our mechanism that simulates nonparticipating agents' bids significantly lessens the benefit that may be obtained from such deviations.

Note that our mechanism that simulates nonparticipating agents does not entirely eliminate the incentive to deviate. For example, suppose that there are two agents, and during the first (learning) phase, the first agent's bids are lower than usual. In this case, the second agent may prefer to compete against the "simulated version" of the first agent and can ensure this by overbidding to force the first agent out of the auction. In addition, an agent may be eliminated by mistake. Revisiting the scenario with two agents, suppose that in the first phase, the first agent's bids are higher than usual. This may result in a high entrance fee for the second bidder and may lead to elimination of the second bidder from all the subsequent auctions. We include a small slack in the chosen entrance fees to ensure that the likelihood of such mistaken elimination is small.

We can now state the main result of this section. Note that we do not need F to be an MHR or even a regular distribution. A bounded support suffices; any other conditions under which a Hoeffding-type bound holds uniformly would serve just as well (Hoeffding 1963).

Theorem 4 (Surplus-Extracting Mechanism). *Suppose that the valuations of all agents are drawn i.i.d. from distribution F over $[0,1]$. Distribution F is a priori unknown to the seller, but it is known to the agents. If all the agents are truthful, the SESE mechanism with an exploration phase of length N obtains an expected per-auction revenue*[25] *of* $\mathrm{E}[\max_j\{v_{jt}\}] - O(\sqrt{\log N/N})$.
In addition, under this mechanism, for any agent i and time t, if all the other agents are always truthful, then with probability $1 - O(N^{-2})$, the increase in per-auction utility that can be obtained by deviating from the truthful strategy is bounded by $O(\sqrt{\log N/N})$.

Note that the loss decreases as the length of the first phase increases. However, the mechanism loses revenue in the first phase. The theorem shows that SESE is approximately incentive compatible. In the proof presented in Online Appendix E, we show that for any strategy B and every period τ, with probability $1 - O(N^{-2})$ when all agents are always truthful, the personal history $H_{i,\tau}$ seen by agent i so far is such that

$$\underline{U}_{i,H_{i,\tau}}(B_i^{\mathrm{TR}}, B_{-i}^{\mathrm{TR}}) + O\left(\sqrt{\log N/N}\right) \geq \overline{U}_{i,H_{i,\tau}}(B_i, B_{-i}^{\mathrm{TR}});$$

see (3) and (4). With the remaining probability, $O(N^{-2})$, the benefit from deviating might be larger but is nevertheless bounded by one. Hence, the expected benefit of deviating from truthfulness is $O(\sqrt{\log N/N})$. In other words, truthfulness is an approximate best response to the other agents being always truthful. The notion of approximate incentive compatibility implies that agents do not deviate from the truthful strategy when the benefit from such a deviation is insignificant. The notion of approximate incentive compatibility is appealing when characterizing or computing the best-response strategy is challenging, and several works moreover use an additive notion of approximate IC similar to ours (Schummer 2004, McSherry and Talwar 2007, Daskalakis et al. 2009, Nazerzadeh et al. 2013). In online ad auctions, finding profitable deviation strategies requires solving complicated dynamic programs in a highly uncertain environment. Thus, agents can plausibly be expected to bid truthfully under an approximately incentive-compatible mechanism.

We remark that our notion of approximate incentive compatibility is additive in the sense that the absolute increase in utility from a deviation is small. An alternative definition would be multiplicative approximate incentive compatibility where the relative gain from a deviation is small. Note that these two notions differ when the utility of a bidder is small (close to zero).[26]

The first and second phases can be interpreted as exploration and exploitation phases, respectively. In an environment where valuations may change slightly over time, the seller can continue to explore occasionally in order to adjust for the change in valuations. For instance, with a small probability, any round $t > N$ can be designated an exploration round, and the entrance fees can be set to zero. Stale exploration data can be discarded as new data are generated. (This will also ensure that the long-run average revenue converges to the ex ante expected value with probability 1.)

8. Conclusion

Designing data-driven incentive-compatible mechanisms has become an important research agenda, motivated in part by the rapid growth of online marketplaces. In this work, we showed that the revenue of repeated auctions can be optimized when the valuations of each bidder can be estimated from the valuations of other bidders. The main goal of the paper was to study the tension between learning and incentive properties. The model is set up to study the hardest case of this tension, namely when all the bidders participate in all the auctions. If some bidders do not participate in an auction, their previous bids can be used to learn and set prices without causing any incentive issues, in addition to previous bids by bidders who are participating. Even though we have not explicitly modeled participation, our results would extend to such environments because we proposed mechanisms based on the following two principles: (i) the personal price for each agent should be based only on the historical bids of other agents, and (ii) an agent should not benefit from preventing other agents from participating by raising the prices they face.

We showed that our work can be practically useful by showing that there is only a small revenue loss in case of limited heterogeneity in bidder valuation distributions and by extending our ideas to a contextual setting with heterogeneous items that allows for correlation between valuations of buyers. A natural research direction is to explore the optimal trade-off between incentive compatibility and learning, as a function of heterogeneity among bidders; see Golrezaei et al. (2018). Another interesting direction would be the case where the auctions are connected via budget constraints; see Balseiro and Gur (2019).

Furthermore, we believe that the ideas developed here can be applied to other repeated auction mechanisms that were designed under the assumption that the valuation distributions are known. For instance, Balseiro et al. (2018) propose a repeated auction mechanism that is a hybrid of first-price and second-price auctions and can extract almost the entire surplus of the buyers. We believe that similar incentive-compatible approximately surplus-extracting mechanisms can be constructed for an *unknown* distribution using our approach.

Acknowledgments

The authors thank Tim Roughgarden and anonymous referees for their insightful comments and suggestions, along with seminar participants at The Web Conference 2019, the INFORMS Annual Meetings 2018 and 2019, Google, ACM Economics and Computation Conference 2017, and the Marketplace Innovation workshop 2017.

Endnotes

[1] In the case of unlimited supply, incentive compatibility directly follows if the price of each buyer depends only on the previous bids of other buyers; see Balcan et al. (2008). With limited supply, obtaining incentive compatibility is more challenging because of “competition” among buyers.

[2] From a technical perspective, we build on prior work that investigates how samples from a distribution can be used to set a near-optimal reserve price; see Dhangwatnotai et al. (2015).

[3] The first paper on the topic was an earlier conference paper by us (Kanoria and Nazerzadeh 2014), which studied a different model, namely one in which each bidder draws her valuation just once and

retains that valuation for all rounds (time periods), and introduced the idea of exploiting competition to manage bidder incentives in repeated auctions. Subsequently, Immorlica et al. (2017) studied a repeated sales setting and developed a mechanism that is similar in spirit in that it exploits competition to manage buyer incentives.

[4] We allow such a mechanism to use only the bids and not the reserve prices (nor the allocations and payments) because the entire history can be "encoded" in the decimal representation of reserve price r_τ, with vanishing impact on revenues, and this would defeat the purpose of defining window length W.

[5] This is similar to the common practice in ad exchanges, where the bidder may not see the reserve. Often, the exchange communicates a (possibly lower) reserve price, which may be different from the reserve price that is applied to the payments.

[6] This informational robustness is in contrast to repeated first-price auction settings (see, e.g., Bergemann and Horner 2010) where information revelation can significantly change the outcome.

[7] If both agents shade, the resulting equilibrium (or limit cycle) may involve further loss in revenue for the seller.

[8] A more general class of strategies involves bidding some $r_0 \in [\underline{r}, \bar{r}]$ for all valuations in $[\underline{r}, \bar{r}]$ and bidding truthfully otherwise. We expect that a best response in this class would yield a larger benefit from deviation while still hurting the revenue earned by the seller.

[9] The numbers in this example are rounded to three decimal points; see Online Appendix A.1 for details.

[10] Formally, the reserve price function Ω now outputs an n vector of reserve prices, one for each agent.

[11] Goldberg et al. (2001) and follow-up works broadly inspired this approach, although the setting and results are quite different; there, a digital good (which can be reproduced costlessly) is sold simultaneously to multiple buyers, and the seller does not know the valuation distribution.

[12] To clarify this definition, suppose that there are three bidders i, j, and k. Then, the bids $b_{j\tau}$ and $b_{k\tau}$ for relevant $\tau < t$ are regarded as two separate, scalar data points in the definition of $\hat{F}_{-i}$. Thus, if window length W is used, the empirical distribution is based on $(n-1)W$ data points/bids by other bidders during the last W rounds.

[13] We adopt the definition $F(r) = \Pr(v < r)$ with a strict inequality so that the arg max exists.

[14] In this case, the mechanism should compute the so-called "guarded empirical reserve" from the empirical distribution of historical bids, which eliminates the largest bids from consideration as potential reserve prices; see Dhangwatnotai et al. (2015, equation 12 and lemma 4.1).

[15] Because F is an MHR distribution, it has positive density everywhere in the support, making truthful bidding the unique myopic best response whenever there are two or more bidders.

[16] These variations are sometimes called *lazy* and *eager*; see Dhangwatnotai et al. (2015) and Paes Leme et al. (2016).

[17] See Golrezaei et al. (2018) for a discussion on the challenges of implementing the Myerson auction in practical settings.

[18] The virtual value of agent i is $\phi_i(v_i) = v_i - (1 - F_i(v_i))/f_i(v_i)$.

[19] We should be able to extend our analysis to α-strongly regular distributions (Cole and Roughgarden 2014), where the virtual value functions increase at rate at least α everywhere in the support. The lower bound α on the rate of increase (we have $\alpha = 1$ for MHR distributions) will be a part of the upper bound on revenue loss.

[20] In fact, in definition (7), we can ignore values of v below $\min(r_1, r_2)$ (the smaller of the Myerson optimal reserve prices for F_i and F_j). Theorem 2 still holds, and the proof is unaffected.

[21] Note that a multiplicative approximation would be a stronger result: given our boundedness assumptions, a multiplicative approximation implies an additive approximation but not vice versa. However, as a result of estimation errors in learning β, we obtain only an additive approximation here.

[22] The expectation is over the past contexts, past valuations, and period t valuations.

[23] Reservation (guaranteed delivery) contracts for selling display advertising specify the number of impressions to be allocated under the contract in advance. The allocation is determined by the publisher and not by an auction.

[24] To simplify the presentation, we assume that the agents know the distribution of valuations because agents may learn the distributions over time. Note that incentive compatibility clearly continues to hold even if agents do not know the distributions of valuations.

[25] The limiting revenue (1) as well as the limiting per-round utility (2) are well defined under SESE when agents are always truthful.

[26] However, note that technically, the mechanism can share some of the surplus with the bidders.

References

Acquisti A, Varian HR (2005) Conditioning prices on purchase history. *Marketing Sci.* 24(3):367–381.

Amin K, Rostamizadeh A, Syed U (2013) Learning prices for repeated auctions with strategic buyers. Burges CJC, Bottou L, Ghahramani Z, Weinberger KQ, eds. *NIPS'13: Proc. 26th Intern. Conf. Neural Inform. Processing Systems vol 1*, (Curran Associates Inc., Red Hook, NY)1169–1177.

Arrow KJ (1979) The property rights doctrine and demand revelation under incomplete information. *Economics and Human Welfare* (Elsevier), 23–39.

Aviv Y, Pazgal A (2008) Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing Service Oper. Management* 10(3):339–359.

Aviv Y, Wei MM, Zhang F (2019) Responsive pricing of fashion products: The effects of demand learning and strategic consumer behavior. *Management Sci.* 65(7):2982–3000.

Balcan M-F, Blum A, Hartline JD, Mansour Y (2008) Reducing mechanism design to algorithm design via machine learning. *J. Comput. System Sci.* 74(8):1245–1270.

Baliga S, Vohra R (2003) Market research and market design. *Adv. Theoret. Econom.* 3(1):1059.

Balseiro SR, Gur Y (2019) Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Sci.* 65(9): 3952–3968.

Balseiro SR, Mirrokni VS, Paes Leme R (2018) Dynamic mechanisms with martingale utilities. *Management Sci.* 64(11):5062–5082.

Baron DP, Besanko D (1984) Regulation and information in a continuing relationship. *Inform Econom. Policy* 1(3):267–302.

Bergemann D, Horner J (2010) Should auctions be transparent? Cowles Foundation Discussion Paper No. 1764, Yale University, New Haven, CT.

Bergemann D, Välimäki J (2010) The dynamic pivot mechanism. *Econometrica* 78(2):771–789.

Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Oper. Res.* 57(6):1407–1420.

Bikhchandani S, McCardle K (2012) Behavior-based price discrimination by a patient seller. *B. E. J. Theoret. Econom.* 12(1):1–18.

Cesa-Bianchi N, Gentile C, Mansour Y (2013) Regret minimization for reserve prices in second-price auctions. *Proc. 24th Annual ACM-SIAM Sympos. Discrete Algorithms (SODA)*, 1190–1204.

Cesa-Bianchi N, Gentile C, Mansour Y (2015) Regret minimization for reserve prices in second-price auctions. *IEEE Trans. Inform. Theory* 61(1):549–564.

Chen X, Wang Z (2016) Bayesian dynamic learning and pricing with strategic customers. Preprint, submitted August 16.

Cole R, Roughgarden T (2014) The sample complexity of revenue maximization. *Proc. 46th Annual ACM Sympos. Theory Comput.* (ACM, New York), 243–252.

Conitzer V, Taylor CR, Wagman L (2012) Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Sci.* 31(2):277–292.

Correa J, Montoya R, Thraves C (2016) Contingent preannounced pricing policies with strategic consumers. *Oper. Res.* 64(1):251–272.

Daskalakis C, Mehta A, Papadimitriou CH (2009) A note on approximate Nash equilibria. *Theoret. Comput. Sci.* 410(17):1581–1588.

d'Aspremont C, Gérard-Varet L-A (1979) Incentives and incomplete information. *J. Public Econom.* 11(1):25–45.

Dasu S, Tong C (2010) Dynamic pricing when consumers are strategic: Analysis of posted and contingent pricing schemes. *Eur. J. Oper. Res.* 204(3):662–671.

den Boer A (2015) Dynamic pricing and learning: Historical origins, current research, and new directions. *Surveys Oper. Res. Management Sci.* 20(1):1–18.

Devanur NR, Huang Z, Psomas C-A (2016) The sample complexity of auctions with side information. *Proc. 48th Annual ACM Sympos. Theory Comput.*, 426–439.

Devanur NR, Peres Y, Sivan B (2014) Perfect Bayesian equilibria in repeated sales. *Proc. 26th Annual ACM-SIAM Sympos. Discrete Algorithms* (SIAM, Philadelphia, PA), 983–1002.

Dhangwatnotai P, Roughgarden T, Yan Q (2015) Revenue maximization with a single sample. *Games Econom. Behav.* 91:318–333.

Epasto A, Mahdian M, Mirrokni V, Zuo S (2018) Incentive-aware learning for large markets. *Proc. 2018 World Wide Web Conf.* (International World Wide Web Conferences Steering Committee, Geneva, Switzerland), 1369–1378.

Ëso P, Szentes B (2007) Optimal information disclosure in auctions and the handicap auction. *Rev. Econom. Stud.* 74(3):705–731.

Esteves RB (2009) A survey on the economics of behaviour-based price discrimination. NIPE Working Paper 5/2009, NIPE–Universidade do Minho, Braga, Protugal.

Fudenberg D, Miguel Villas-Boas J (2007) *Behavior-Based Price Discrimination and Customer Recognition* (Elsevier Science, Oxford, UK).

Goldberg A, Hartline JD, Wright A (2001) Competitive auctions and digital goods. *Proc. 12th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics), 735–744.

Golrezaei N, Javanmard A, Mirrokni V (2019) Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Advances in Neural Information Processing Systems*, 9756–9766.

Golrezaei N, Lin M, Mirrokni V, Nazerzadeh H (2018) Boosted second-price auctions for heterogeneous bidders. Preprint, submited August 10, https://dx.doi.org/10.2139/ssrn.3016465.

Harrison JM, Keskin NB, Zeevi A (2012) Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Management Sci.* 58(3):570–586.

Hart OD, Tirole J (1988) Contract renegotiation and Coasian dynamics. *Rev. Econom. Stud.* 55(4):509–540.

Hartline JD, Roughgarden T (2009) Simple vs. optimal mechanisms. *Proc. 10th ACM Conf. Electronic Commerce*, 225–234.

Hoeffding W (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* 58(301):13–30.

Huang Z, Mansour Y, Roughgarden T (2015) Making the most of your samples. *Proc. 2016 ACM Conf. Econom. Comput.* (ACM), 45–60.

Immorlica N, Lucier B, Pountourakis E, Taggart S, (2017) Repeated sales with multiple strategic buyers. *Proc. 2017 ACM Conf. Econom. Comput.* (ACM, New York), 167–168.

Kanoria Y, Nazerzadeh H (2014) Dynamic reserve prices for repeated auctions: Learning from bids. *Internat. Conf. Web Internet Econom.* (Springer, Cham), 232.

McAfee P (2011) The design of advertising exchanges. *Rev. Indust. Organ.* 39(3):169–185.

McAfee R, Vincent D (1992) Updating the reserve price in common-value auctions. *Amer. Econom. Rev.* 82(2):512–518.

McSherry F, Talwar K (2007) Mechanism design via differential privacy. *48th Annual IEEE Sympos. Foundations Comput. Sci. 2007* (IEEE), 94–103.

Mirrokni V, Nazerzadeh H (2017) Deals or no deals: Contract design for online advertising. *Proc. 25th Internat. Conf. World Wide Web* (International World Wide Web Conferences Steering Committee).

Mohri M, Medina AM (2014a) Optimal regret minimization in posted-price auctions with strategic buyers. *Proc. 27th Internat. Conf. Neural Inform. Processing Systems Vol. 2* (MIT Press, Cambridge, MA), 1871–1879.

Mohri M, Medina AM (2014b) Learning theory and algorithms for revenue optimization in second price auctions with reserve. *Proc. 31th Internat. Conf. Machine Learning*, 262–270.

Muthukrishnan S (2009) Ad exchanges: Research issues. *5th Internat. Workshop (WINE) Internet Network Econom.*, 1–12.

Myerson RB (1981) Optimal auction design. *Math. Oper. Res.* 6(1):58–73.

Nazerzadeh H, Saberi A, Vohra R (2013) Dynamic cost-per-action mechanisms and applications to online advertising. *Oper. Res.* 61(1):98–111.

Nazerzadeh H, Leme RP, Rostamizadeh A, Syed U (2016) Where to sell: Simulating auctions from learning algorithms. *Proc. 2016 ACM Conf. Econom. Comput.* (ACM, New York), 597–598.

Ostrovsky M, Schwarz M (2009) Reserve prices in internet advertising auctions: A field experiment. Research Paper No. 2054, Stanford University Graduate School of Business, Stanford, CA.

Paes Leme R, Pál M, Vassilvitskii S (2016) A field guide to personalized reserve prices. *Proc. 25th Internat. Conf. World Wide Web* (International World Wide Web Conferences Steering Committee), 1093–1102.

Riley JG, Samuelson WF (1981) Optimal auctions. *Amer. Econom. Rev.* 71(3):381–392.

Roughgarden T, Wang JR (2016) Minimizing regret with multiple reserves. *Proc. 2016 ACM Conf. Econom. Comput.* (ACM), 601–616.

Salant SW (1989) When is inducing self-selection suboptimal for a monopolist? *Quart. J. Econom.* 104(2):391–397.

Schummer J (2004) Almost-dominant strategy implementation: Exchange economies. *Games Econom. Behav.* 48(1):154–170.

Segal I (2003) Optimal pricing mechanisms with unknown demand. *Amer. Econom. Rev.* 93(3):509–529.

Stokey NL (1979) Intertemporal price discrimination. *Quart. J. Econom.* 93(3):355–371.

Taylor CR (2004) Consumer privacy and the market for customer information. *RAND J. Econom.* 35(4):631–650.

Villas-Boas JM (2004) Price cycles in markets with customer recognition. *RAND J. Econom.* 35(3):486–501.

Wang Z, Deng S, Ye Y (2014) Close the gaps: A learning-while-doing algorithm for single-product revenue management problems. *Oper. Res.* 62(2):318–331.

Yash Kanoria is the Sidney Taurel Associate Professor of Business at Columbia University, Graduate School of Business. His research focuses on the design and operations of marketplaces, especially matching markets and online platforms.

Hamid Nazerzadeh is a Dean's Associate Professor in Business Administration at University of Southern California, Marshall School of Business. His research focuses on mechanism design and optimization algorithms and their applications in operations and monetization of online markets. He works extensively with tech companies, including Microsoft, Google, and Uber.