Mechanism Design for Data Science

Shuchi Chawla[*] Jason Hartline[†] Denis Nekipelov[‡]

August 12, 2015

Abstract

Good economic mechanisms depend on the preferences of participants in the mechanism. For example, the revenue-optimal auction for selling an item is parameterized by a reserve price, and the appropriate reserve price depends on how much the bidders are willing to pay. A mechanism designer can potentially learn about the participants' preferences by observing historical data from the mechanism; the designer could then update the mechanism in response to learned preferences to improve its performance. The challenge of such an approach is that the data corresponds to the actions of the participants and not their preferences. Preferences can potentially be inferred from actions but the degree of inference possible depends on the mechanism. In the optimal auction example, it is impossible to learn anything about preferences of bidders who are not willing to pay the reserve price. These bidders will not cast bids in the auction and, from historical bid data, the auctioneer could never learn that lowering the reserve price would give a higher revenue (even if it would). To address this impossibility, the auctioneer could sacrifice revenue optimality in the initial auction to obtain better inference properties so that the auction's parameters can be adapted to changing preferences in the future. This paper develops the theory for optimal mechanism design subject to good inferability.

1 Introduction

This paper develops prior-independent methods for revenue management of an auctioneer. The classical revenue optimal auction is prior dependent, i.e., requiring knowledge of the distribution over values of the bidders (cf. Myerson, 1981). We study a paradigmatic family of auctions and develop a method for counter-factual inference that allows the equilibrium revenue of one auction in the family to be estimated from bids in another. One application of this method is a framework for A/B testing of auctions, a.k.a., randomized controlled trials, wherein an auctioneer can compare the revenue of auctions A and B. Another application is in instrumented optimization where we identify sufficient properties of an auction so that the revenue of any counter-factual auction can be estimated and then optimize revenue over auctions with these properties.

The main technical contribution of this paper is a method for counter-factual revenue estimation: given two auctions we define an estimator for the equilibrium revenue of one from equilibrium bids of the other. Our estimator has a number of appealing properties in contrast to the standard econometric approach to inference in auctions. In the standard approach, first the value distribution

[*]University of Wisconsin - Madison
[†]Northwestern University
[‡]University of Virginia

is inferred from bids of the first auction using equilibrium analysis, and then the estimated value distribution is used to calculate the expected revenue of the second auction. To infer the value distribution, the standard approach employs estimates of the derivative of the bid function via an estimator that typically must be tuned to trade-off bias and variance. by an assumption on bid distribution. In contrast, our method estimates revenue directly from the bids and our estimator requires no distribution dependent tuning Our method is statistically efficient with estimation error proportional to one over the square root of the number of observed bids.

Our work applies to first-price and all-pay position auctions, a model popularized by studies of auctions for advertising on Internet search engines (cf. Varian, 2007, and Edelman et al., 2007).[1] A position auction is defined by a decreasing sequence of weights which correspond to allocation probabilities, bidders are assigned to weights assortatively by bid, and pay their bid if allocated (first-price) or always (all-pay). The configurable parameters in this family of auctions are the weights of the positions. Given two position auctions B and C, each defined by positions weights, and N samples from the Bayes-Nash equilibrium bid distribution from C, our estimator for the Bayes-Nash equilibrium revenue of B is a *weighted order statistic*. We apply a formula to the position weights in B and C to get a weight for each order statistic of the N bids, and then the estimator is the weighted sum of the order statistics. The error bounds for this estimator are proportional to $\sqrt{1/N}$ and a term derived from the position weights of B and C.

Our first application of this revenue estimator is to A/B testing of auctions. A/B testing, otherwise known as randomized controlled trials, is an industry standard method for evaluation, tuning, and optimizing Internet services and e-commerce systems. This form of online experimentation is happening all the time and the participants of the experiments are almost always unaware that the experiment is being conducted. Our framework for A/B testing of auctions is motivated – as we describe subsequently – by auction environments where *ideal A/B testing* is impossible.[2] In our framework bidders bid first and then the experimenter randomly selects and runs control auction A or treatment auction B on the bids. Importantly, the bidders are unaware of whether they are in the control A or treatment B, but have instead bid according to the Bayes-Nash equilibrium of auction C, the convex combination of auctions A and B. Our task of A/B testing of auctions is then to compare estimates of revenue of A and B given bids in C. Note, a convex combination of position auctions is a position auction with position weights given by the convex combination. Suppose the A/B test auction C runs the control auction A with probability $1 - \epsilon$ and the treatment auction B with probability ϵ. Our main result for A/B testing is that the revenue from B can be estimated from bids in C with error that depends on ϵ as $\log(1/\epsilon)$. This error bound improves exponentially over the $\sqrt{1/\epsilon}$ dependence on ϵ that would be obtained by ideal A/B testing.

Our second application of our revenue estimator is to instrumented optimization. Note that classical revenue optimization in auctions, e.g., by reserve prices and ironing (Myerson, 1981), is at odds with classical structural inference. Reserve pricing and ironing pool bidders with distinct values and, thereafter, no procedure for structural inference can distinguish them. The position auctions (with neither ironing nor reserve prices) of our study do not exhibit this stark behavior; nonetheless, the error in our revenue estimate for auction B from the bids in auction C does depend on auction C. (In particular on a relationship between the position weights in B and C; e.g., the

[1]First-price and all-pay position auctions generalize classical single-item and multi-unit auction models and are important auction form for theoretical study, see e.g., Chawla and Hartline (2013). Unfortunately, our methods cannot be directly applied to position auctions with the so-called "generalized second price" payment rule of Google's AdWords platform.

[2]In ideal A/B testing, the bids in A and B are respectively in equilibrium for A and B.

position weights are related by ϵ in the A/B testing application, above.) Our first result shows that there is universal treatment B such that in the A/B test mechanism C, the revenue of any other position auction can be estimated. A solution, then, to the instrumented optimization problem is to run the A/B test mechanism C that corresponds to the revenue optimal position auction A and this universal treatment B. Our second result incorporates a bound on the desired rate of estimation into the revenue optimization problem and derives the revenue optimal auction subject to good revenue inference. Our analysis gives a tradeoff between revenue bounds (relative to the optimal position auction) and the desired rate of inference. Finally, we show that the revenue optimal position auction (without reserve prices or ironing) approximates the revenue-optimal auction (with reserve prices and ironing); thus, there is little loss in revenue from restricting to the family of position auctions for which our inference methods are applicable.

1.1 Motivating Example: Auctions for Internet Search Advertising.

Our work is motivated by the auctions that sell advertisements on Internet search engines (see historical discussion by Fain and Pedersen, 2006). The first-price position auction we study in this paper was introduced in 1998 by the Internet listing company GoTo. This auction was adapted by Google in 2002 for their AdWords platform, modified to have a second-price-like payment rule, and is known as the *generalized second-price auction*. Early theoretical studies of equilibrium in the generalized second-price auction were conducted by Varian (2007) and Edelman et al. (2007); unlike the second-price auction for which it is named, the generalized second-price auction does not admit a truthtelling equilibrium.

Internet search advertising works as follows. A user looking for web pages on a given specifies *keywords* on which to search. The search engine then returns a listing of web pages that relate to these keywords. Alongside these search results, the search engine displays sponsored results. These results are conventionally explicitly labeled as sponsored and appear in the *mainline*, i.e., above the search results, or in the *sidebar*, i.e., to the right of the search results. The mainline typically contains up to four ads and the sidebar contains up to seven ads. The order of the ads is of importance as the Internet user is more likely to read and click on ads in higher positions on the page. In the classic model of Varian (2007) and Edelman et al. (2007) the user's click behavior is exogenously given by weights associated with the positions,[3] and the weights are decreasing in position. An advertiser only pays for the ad if the user clicks on it. Thus in the classic first-price position auction, advertisers are assigned to positions in order of their bids, and the advertisers on whose ads the user clicks each pay their bids.

As described above, the ads in the mainline have higher click rates than those in the the sidebar. The mainline, however, is not required to be filled to capacity (a maximum of four ads). In the first-price position auction described above, the choice of the number of ads to show in the mainline affects the revenue of the auction and, in the standard auction-theoretic model of Bayes-Nash equilibrium, this choice is ambiguous with respect to revenue ranking. For some distributions of advertiser values, showing more ads in the mainline gives more revenue, while for other distributions fewer ads gives more revenue.

The keywords of the user enable the advertisers to target users with distinct interests. For example, hotels in Hawaii may wish to show ads to a user searching for "best beaches," while a

[3] Endogenous click models have also been considered, e.g., Athey and Ellison (2011), but are less prevalent in the literature.

computer hardware company would prefer users searching for "laptop reviews." Thus, the search advertising is in fact a collection of many partially overlapping markets, with some high-volume high-demand keywords and a long tail of niche keywords. The conditions of each of these markets are distinct and thus, as per the discussion of the preceding paragraph, the number of ads to show in the mainline depends on the keywords of the search.

One empirical method for evaluating two alternatives, e.g., showing one or two mainline ads, is A/B testing. In the ideal setting of A/B testing, the auctions for a given keyword would be randomly divided into the A and B, and in part A the advertisers would bid in Bayes-Nash equilibrium for A and in part B they would bid in equilibrium for B. Unfortunately, because we need to test both A and B in each market, ideal A/B testing would require soliciting distinct bids for each variant of the auction. This approach is impractical, both from an engineering perspective and from a public relations perspective. In practice, A/B tests are run on these ad platforms all the time and without informing the advertisers. Of course, advertisers can observe any overall change in the mechanism and adapt their bids accordingly, i.e., they can be assumed to be in equilibrium. Our approach of A/B testing, that of assuming that bids are in equilibrium for auction C, i.e., the convex combination of A and B, is consistent with the Industry standard practice.

Our A/B testing framework is is motivated specifically by the goal of optimizing an auction to local characteristics of the market in which the auction is run. There is a third framework for A/B testing framework that can be used to evaluate global characteristics of collections of auctions. This framework randomly partitions markets (by keywords) with little overlap into the control group (where auction A is run) and treatment group (with auction B). From such an A/B test we can evaluate whether it is better to run A in every market or B in every market. It cannot be used, however, for our motivating application of determining the number of mainline ads to show, where the optimal number naturally varies across markets. The work of Ostrovsky and Schwarz (2011) on reserve pricing in Yahoo!'s ad auction demonstrates how such a global A/B test can be valuable. They first used a parametric estimator for the value distribution in each market to determine a good reserve price for that market. Then they did a global A/B test to determine whether the auction with their calculated reserve prices (the B mechanism) has higher revenue on average than the auction with the original reserve prices (the A mechanism). Our methods related to and can replace the first step of their analysis.

2 Related Work

In the mechanism design literature, the problem of designing mechanisms to enable learning the parameters of a market has not been considered from a theoretical perspective previously. Several works have considered the problem of learning optimal pricing schemes in an online setting (e.g., Babaioff et al. (2012)). However, these works assume non-strategic behavior on part of the agents, which makes the inference much simpler. Other works consider the problem of learning click-through-rates in the context of a sponsored search auction (a generalization of the position environment we study) while simultaneously obtaining good revenue (e.g., Devanur and Kakade (2009); Babaioff et al. (2009); Gatti et al. (2012)), however, they restrict attention to truthful mechanisms, and again do not require inference.

Several works have considered the problem of empirically optimizing the reserve price of an auction in an online repeated auction setting (e.g., Reiley (2006); Brown and Morgan (2009); Ostrovsky and Schwarz (2011)). The most notable of these is the work of Ostrovsky and Schwarz

(2011). Ostrovsky and Schwarz (2011) adapt their mechanism over time to respond to empirical data by determining the optimal reserve price for the empirically observed distribution, and then setting a reserve price that is slightly smaller. This allows for inference around the optimal reserve price and ensures that the mechanism quickly adapts to changes in the distribution.

Finally, the theory that we develop for optimizing revenue over the class of iron by rank auctions is isomorphic to the theory of envy-free optimal pricing developed by Hartline and Yan (2011).

Our approach to data-driven mechanism design is based on evaluation of mechanisms using the data directly thus *bypassing the computation of equilibrium best responses of agents under the new mechanism.* That comes in contrtast with the approach frequently deployed for counterfactual inference in Industrial Organization. The traditional approach to counterfactual analysis (see Athey and Haile (2007), Doraszelski and Pakes (2007) and Paarsch and Hong (2006)) involves, first, using the observational data to infer the underlying preference parameters (such as valuations of bidders in auctions). Second, the inferred structural parameters are used as an input in the computation of the equilibrium outcome of the new proposed mechanism.

The traditional approach requires good properties from the first stage inference procedure given that the outcomes of this procedure is used as an input for counterfactual prediction for the new mechanism. Rich literature in Econometrics including Newey (1994), Ai and Chen (2003), Blundell and Powell (2004), Chernozhukov et al. (2013), and Chen and Pouzo (2012) require a careful control over the properties of the object estimated in the first stage to ensure good statistical properties of the object of interest (such as the counterfactual prediction). For instance, in case of first-price auctions in Guerre et al. (2000) such a control requires the balanced choice of the smoothing beandwidth for estimation of the distribution of bids which depends on the smoothness of the true (infeasible) distribution of bids.

Our approach is based on using the bids observed in the data directly to make inferences regarding the counterfactual mechanisms. This allows us to do inference without relying on assumed theoretical properties of the underlying distribution of valuations of players. Moreover, that also allows us to consider cases where teh object of interest is not a smooth functional of the distribution of the data without imposing additional constraints on the model or the data generating process.

3 Preliminaries

3.1 Auction Theory

A standard auction design problem is defined by a set $[n] = \{1, \ldots, n\}$ of $n \geq 2$ agents, each with a private value v_i for receiving a service. The values are bounded: $v_i \in [0, 1]$; They are independently and identically distributed according to a continuous distribution F. If x_i indicates the probability of service and p_i the expected payment required, agent i has linear utility $u_i = v_i x_i - p_i$. An auction elicits bids $\boldsymbol{b} = (b_1, \ldots, b_n)$ from the agents and maps the vector $\boldsymbol{b}$ of bids to an allocation $\tilde{\boldsymbol{x}}(\boldsymbol{b}) = (\tilde{x}_1(\boldsymbol{b}), \ldots, \tilde{x}_n(\boldsymbol{b}))$, specifying the probability with which each agent is served, and prices $\tilde{\boldsymbol{p}}(\boldsymbol{b}) = (\tilde{x}_1(\boldsymbol{b}), \ldots, \tilde{x}_n(\boldsymbol{b}))$, specifying the expected amount that each agent is required to pay. An auction is denoted by $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$. We focus on symmetric auctions, namely those for which the allocation and payment functions, $\tilde{x}_i$ and $\tilde{p}_i$, do not depend on the identities of the agents. That is, $\tilde{x}_i(\boldsymbol{b}) = \tilde{x}(\boldsymbol{b})$ and $\tilde{p}_i(\boldsymbol{b}) = \tilde{p}(\boldsymbol{b})$ for all agents i, all bids $\boldsymbol{b}$ and some functions $\tilde{x}$ and $\tilde{p}$.

Standard payment formats In this paper we study two standard payment formats. In a *first-price* format, each agent pays his bid upon winning, that is, $\tilde{p}_i(\boldsymbol{b}) = b_i\,\tilde{x}_i(\boldsymbol{b})$. In an *all-pay* format, each agent pays his bid regardless of whether or not he wins, that is, $\tilde{p}_i(\boldsymbol{b}) = b_i$.

Bayes-Nash equilibrium The values are independently and identically distributed according to a continuous distribution F. This distribution is common knowledge to the agents. A strategy s_i for agent i is a function that maps the value of the agent to a bid. The distribution of values F and a profile of strategies $\boldsymbol{s} = (s_1, \cdots, s_n)$ induces interim allocation and payment rules (as a function of bids) as follows for agent i with bid b_i.

$$\tilde{x}_i(b_i) = \mathbf{E}_{\boldsymbol{v}_{-i}\sim F}[\tilde{x}_i(b_i, \boldsymbol{s}_{-i}(\boldsymbol{v}_{-i}))] \text{ and}$$
$$\tilde{p}_i(b_i) = \mathbf{E}_{\boldsymbol{v}_{-i}\sim F}[\tilde{p}_i(b_i, \boldsymbol{s}_{-i}(\boldsymbol{v}_{-i}))]\,.$$

Agents have linear utility which can be expressed in the interm as:

$$\tilde{u}_i(v_i, b_i) = v_i\tilde{x}_i(b_i) - \tilde{p}_i(b_i).$$

The strategy profile forms a *Bayes-Nash equilibrium* (BNE) if for all agents i, values v_i, and alternative bids b_i, bidding $s_i(v_i)$ according to the strategy profile is at least as good as bidding b_i. I.e.,

$$\tilde{u}_i(v_i, s_i(v_i)) \geq \tilde{u}_i(v_i, b_i). \tag{1}$$

A symmetric equilibrium is one where all agents bid by the same strategy, i.e., $\boldsymbol{s}$ statisfies $s_i = s$ for all i and some s. For a symmetric equilibrium of a symmetric auction, the interim allocation and payment rules are also symmetric, i.e., $\tilde{x}_i = \tilde{x}$ and $s_i = s$ for all i. For implicit distribution F and symmetric equilibrium given by stratey s, a mechanism can be described by the pair $(\tilde{x}, \tilde{p})$. Chawla and Hartline (2013) show that the equilibrium of every auction in the class we consider is unique and symmetric.

The strategy profile allows the mechanism's outcome rules to be expressed in terms of the agents' values instead of their bids; the distribution of values allows them to be expressed in terms of the agents' values relative to the distribution. This latter representation exposes the geometry of the mechanism. Define the *quantile* q of an agent with value v to be the probability that v is larger than a random draw from the distribution F, i.e., $q = F(v)$. Denote the agent's value as a function of quantile as $v(q) = F^{-1}(q)$, and his bid as a function of quantile as $b(q) = s(v(q))$. The outcome rule of the mechanism in quantile space is the pair $(x(q), p(q)) = (\tilde{x}(b(q)), \tilde{p}(b(q)))$.

Revenue curves and auction revenue Myerson (1981) characterized Bayes-Nash equilibria and this characteriation enables writing the revenue of a mechanism as a weighted sum of revenues of single-agent posted pricings. Formally, the *revenue curve* $R(q)$ for a given value distribution specifies the revenue of the single-agent mechanism that serves an agent with value drawn from that distribution if and only if the agent's quantile exceeds q: $R(q) = v(q)\,(1 - q)$. $R(0)$ and $R(1)$ are defined as 0. Myerson's characterization of BNE then implies that the expected revenue of a mechanism at BNE from an agent facing an allocation rule $x(q)$ can be written as follows:

$$P_x = \mathbf{E}_q\big[R(q)x'(q)\big] = -\mathbf{E}_q\big[R'(q)x(q)\big] \tag{2}$$

where x' and R' denote the derivative of x and R with respect to q, respectively.

The expected revenue of an auction is the sum over the agents of its per-agent expected revenue; for auctions with symmetric equilibrium allocation rule x this revenue is $n\,P_x$.

Position environments and rank-based auctions A *position environment* expresses the feasibility constraint of the auction designer in terms of *position weights* $\boldsymbol{w}$ satisfying $1 \geq w_1 \geq w_2 \geq \cdots \geq w_n \geq 0$. A *position auction* assigns agents (potentially randomly) to positions 1 through n, and an agent assigned to position i gets allocated with probability w_i. The *rank-by-bid position auction* orders the agents by their bids, with ties broken randomly, and assigns agent i, with the ith largest bid, to position i, with allocation probability w_i. *Multi-unit environments* are a special case and are defined for k units as $w_j = 1$ for $j \in \{1, \ldots, k\}$ and $w_j = 0$ for $j \in \{k+1, \ldots, n\}$. The *highest-k-bids-win* multi-unit auction is the special case of the rank-by-bid position auction for the k-unit environment.

In our model with agent values drawn i.i.d. from a continuous distribution, rank-by-bid position auctions with either all-pay or first-price payment semantics have a unique Bayes-Nash equilibrium and this equilibrium is symmetric and efficient, i.e., in equilibrium, the agents' bids and values are in the same order (Chawla and Hartline, 2013).

Rank-by-bid position auctions can be viewed as convex combinations of highest-bids-win multi-unit auctions. The *marginal weights* of a position environment are $\boldsymbol{w}' = (w'_1, \ldots, w'_n)$ with $w'_k = w_k - w_{k+1}$. Define $w'_0 = 1 - w_1$ and note that the marginal weights $\boldsymbol{w}'$ can be interpreted as a probability distribution over $\{0, \ldots, n\}$. As rank-by-bid position auctions are efficient, the rank-by-bid position auction with weights $\boldsymbol{w}$ has the exact same allocation rule as the mechanism that draws a number of units k from the distribution given by $\boldsymbol{w}'$ and runs the highest-k-bids-win auction.

Denote the highest-k-bids-win allocation rule as $x^{(k:n)}$ and its revenue as $P_k = P_{x^{(k:n)}} = \mathbf{E}_q\big[-R'(q)\, x^{(k:n)}(q)\big]$. This allocation rule is precisely the probability an agent with quantile q has one of the highest k quantiles of n agents, or at most $k-1$ of the $n-1$ remaining agents have quantiles greater than q. Formulaically,

$$x^{(k:n)}(q) = \sum_{i=0}^{k-1} \tbinom{n-1}{i} q^{n-1-i}(1-q)^i.$$

Importantly, the allocation rule of a rank-by-bid position auction does not depend on the distribution at all. The allocation rule x of the rank-by-bid position auction with weights $\boldsymbol{w}$ is:

$$x(q) = \textstyle\sum_k w'_k\, x^{(k:n)}(q).$$

By revenue equivalence (Myerson, 1981), the per-agent revenue of the rank-by-bid position auction with weights $\boldsymbol{w}$ is:

$$P_x = \textstyle\sum_k w'_k\, P_k.$$

Of course, $P_0 = P_n = 0$ as always serving or never serving the agents gives zero revenue.

A *rank-based* auction is one where the probability that an agent is served is a function only of the rank of the agent's bid among the other bids and not the magnitudes of the bids. Any rank-based auction induces a position environment where $\bar{w}_k$ denotes the probability that the agent with the kth ranked bid is served. This auction is equivalent to the rank-by-bid position auction with these weights $\bar{\boldsymbol{w}}$. In a position auction with weights $\boldsymbol{w}$, the following lemma characterizes the weights $\bar{\boldsymbol{w}}$ that are induced by rank-based auctions.

Lemma 3.1 (e.g., Devanur et al., 2013)**.** *There is a rank-based auction with induced position weights $\bar{\boldsymbol{w}}$ for a position environment with weights $\boldsymbol{w}$ if and only if their cumulative weights satisfy $\sum_{j=0}^{k} \bar{w}_j \leq \sum_{j=0}^{k} w_j$ for all k.*

3.2 Inference

As we discussed this above, the traditional structural inference in the auction settings is based on inferring distribution of values, which is unobserved, can be inferred from the distribution of bids, which is observed. Once the value distribution is inferred, other properties of the value distribution such as its corresponding revenue curve, which is fundamental for optimizing revenue, can be obtained. In this section we briefly overview this approach.

The key idea behind the inference of the value distribution from the bid distribution is that the strategy which maps values to bids is a best response, by equation (1), to the distribution of bids. As the distribution of bids is observed, and given suitable continuity assumptions, this best response function can be inverted.

The value distribution can be equivalently specified by distribution function $F(\cdot)$ or value function $v(\cdot)$; the bid distribution can similarly be specified by the bid function $b(\cdot)$. For rank-based auctions (as considered by this paper) the allocation rule $x(\cdot)$ in quantile space is known precisely (i.e. it does not depend on the value function $v(\cdot)$). Assume these functions are monotone, continuously differentiable, and invertible.

Inference for first-price auctions Consider a first-price rank-based auction with a symmetric bid function $b(q)$ and allocation rule $x(q)$ in BNE. To invert the bid function we solve for the bid that the agent with any quantile would make. Continuity of this bid function implies that its inverse is well defined. Applying this inverse to the bid distribution gives the value distribution.

The utility of an agent with quantile q as a function of his bid z is

$$u(q, z) = (v(q) - z)\, x(b^{-1}(z)). \tag{3}$$

Differentiating with respect to z we get:

$$\tfrac{d}{dz} u(q, z) = -x(b^{-1}(z)) + \big(v(q) - z\big)\, x'(b^{-1}(z))\, \tfrac{d}{dz} b^{-1}(z).$$

Here x' is the derivative of x with respect to the quantile q. Because $b(\cdot)$ is in BNE, the derivative $\frac{d}{dz}u(q, z)$ is 0 at $z = b(q)$. Rarranging, we obtain:

$$v(q) = b(q) + \tfrac{x(q)\, b'(q)}{x'(q)} \tag{4}$$

Inference for all-pay auctions We repeat the calculation above for rank-based all-pay auctions; the starting equation (3) is replaced with the analogous equation for all-pay auctions:

$$u(q, z) = v(q)\, x(b^{-1}(z)) - z. \tag{5}$$

Differentiating with respect to z we obtain:

$$\tfrac{d}{dz} u(q, z) = v(q)\, x'(b^{-1}(z))\, \frac{d}{dz} b^{-1}(z) - 1,$$

Again the first-order condition of BNE implies that this expression is zero at $z = b(q)$; therefore,

$$v(q) = \tfrac{b'(q)}{x'(q)}. \tag{6}$$

Known and observed quantities Recall that the functions $x(q)$ and $x'(q)$ are known precisely: these are determined by the rank-based auction definition. The functions $b(q)$ and $b'(q)$ are observed. The calculations above hold in the limit as the number of samples from the bid distribution goes to infinity, at which point these obserations are precise.

3.3 Statistical Model and Methods

Our framework for counterfactual predictions is based on directly using the distribution of bids for inference. The main error in estimation of the bid distribution is the *sampling error* due to drawing only a finite number of samples from the bid distribution.

The analyst obtains N samples from the bid distribution. Each sample is the corresponding agent's best response to the *true* bid distribution. We assume that the number of samples N is roughly polynomial in n, the number of agents in a single auction.

We can estimate the quantile function of the equilibrium bid distribution $b(q)$ as follows. Let $\hat{b}_1, \cdots, \hat{b}_N$ denote the N samples drawn from the bid distribution. Sort the bids so that $\hat{b}_1 \leq \hat{b}_2 \leq \cdots \leq \hat{b}_N$ and define the *estimated bid distribution* $\hat{b}(\cdot)$ as

$$\hat{b}(q) = \hat{b}_i \qquad\qquad \forall i \in N, q \in [i-1, i)/N \tag{7}$$

Definition 1. *For function $b(\cdot)$ and estimator $\hat{b}(\cdot)$, the* uniform mean absolute error *is defined as*

$$\mathbf{E}_{\hat{b}}\left[\sup_q \left|b(q) - \hat{b}(q)\right|\right].$$

The main object that will arise in our further analysis will be the weighted quantile function of the bid distribution where the weights are determined by the allocation rule of the auction under consideration. Our statistical results stem from the previous work on the uniform convergence of quantile processes and weighted quantile processes in Csorgo and Revesz (1978), Csörgö (1983), Cheng and Parzen (1997). It turns out that the our main object of interest for inference is the weighted quantile function of the bid distribution. The weight is proportional to the inverse derivative of the allocation rule. This feature leads to highly desirable proprties of the $\sqrt{N}$-normalized mean absolute error, making it bounded by a universal constant.

Lemma 3.2. *The uniform mean absolute error of the empirical quantile function $\hat{b}(\cdot)$ weighted by its derivative is bounded as $N \to \infty$*

$$\mathbf{E}\left[\sup_q \left|\sqrt{N}(b'(q))^{-1}(b(q) - \hat{b}(q))\right|\right] < \frac{1}{2}.$$

The derivative of the quantile function of the bid distribution is determined by the derivative of the allocation rule and the quantile function of the value distribution. Recalling that $v(q) \leq 1$ for all quantiles q, we obtain the following expressions for weighted uniform absolute error in terms of the allocation function.

(i) In the first-price auction the allocation rule-weighted uniform absolute error in the quantile function of bids bids can be evaluated as $\mathbf{E}\left[\sup_q \left|\sqrt{N}\frac{x(q)}{x'(q)}\left(b(q) - \hat{b}(q)\right)\right|\right] \leq \frac{1}{2}$.

(ii) In the all-pay auction the allocation rule-weighted uniform absolute error in the quantile function of bids can be evaluated as $\mathbf{E}\left[\sup_q \left|\sqrt{N}(x'(q))^{-1}\left(b(q) - \hat{b}(q)\right)\right|\right] \leq \frac{1}{2}$.

Equations (4) and (6) enable the value function, or equivalently, the value distribution, to be estimated from the estimated bid function and an estimator for the derivative of the bid function, or equivalently, the density of the bid distribution. Estimation of densities is standard; however, it requires assumptions on the distribution, e.g., continuity, and the convergence rates in most cases will be slower. Our main results do not take this standard approach.

4 Inference methodology and error bounds for all-pay auctions

We will now develop a methodology and error bounds for estimating the revenue of one rank-based auction using bids from another rank-based auction. There are two reasons behind our assumption that the auction that we run (that generates the observed bids) is a rank-based auction. First, the allocation rule (in quantile space) of a rank based auction is independent of the bid and value distribion; therefore, it is known and does not need to be estimated. Second, the allocation rules that result from rank-based auctions are well behaved, in particular their slopes are bounded, and our error analysis makes use of this property.

Recall from Section 3.1 that the revenue of any rank-based auction can be expressed as a linear combination of the multi-unit revenues $P_1, \ldots, P_n$ with P_k equal to the per-agent revenue of the highest-k-bids-win auction. Therefore, in order to estimate the revenue of a rank-based auction, it suffices to estimate the P_ks accurately.

In the following section we describe a function mapping the observed bids to the revenue estimate. In Sections 4.2 and 4.3 we develop error bounds on our estimate for the revenue of a multi-unit auction and a general position auction respectively.

4.1 Inference equation

Consider estimating the revenue of an auction with allocation rule y from the bids of an all-pay position auction. The per-agent revenue of the allocation rule y is given by:

$$P_y = \mathbf{E}_q\big[y'(q)R(q)\big] = \mathbf{E}_q\big[y'(q)v(q)(1-q)\big]$$

Let x denote the allocation rule of the auction that we run, and b denote the bid distribution in BNE of this auction. Recall that for an all-pay auction format, we can convert the bid distribution into the value distribution as follows: $v(q) = b'(q)/x'(q)$. Substituting this into the expression for P_y above we get

$$P_y = \mathbf{E}_q\left[y'(q)(1-q)\frac{b'(q)}{x'(q)}\right] = \mathbf{E}_q\big[Z_y(q)b'(q)\big]$$

where $Z_y(q) = (1-q)\frac{y'(q)}{x'(q)}$.

This expression allows us to derive the revenue using the empirical bid distribution. However, it leads to a large error because the derivative of the bid distribution, $b'(\cdot)$, cannot be estimated with a good convergence rate. We get around this issue by modifying the expression so it becomes a simple weighted average of the empirical bid distribution.

Specifically, writing the expectation as an integral and integrating by parts we obtain the following lemma. Here we note that $b(0) = 0$ and $Z_k(1) = 0$.

Lemma 4.1. *The per-agent revenue of a rank-based auction with allocation rule y can be written as a linear combination of the bids in an all-pay auction:*

$$P_y = \mathbf{E}_q\left[-Z_y'(q)b(q)\right]$$

where $Z_y(q) = (1-q)\frac{y'(q)}{x'(q)}$ depends on the allocation rule x of the mechanism and is known precisely.

This formulation allows us to estimate P_y directly as a weighted average of the observed bids. Specifically, recalling that $\hat{b}(\cdot)$ denotes the estimated bid distribution, we can write the estimate of P_y as:

$$\hat{P}_y = \mathbf{E}_q\left[-Z_y'(q)\hat{b}(q)\right] \tag{8}$$

and, using that $\hat{b}$ is a step function,

$$= \sum_{i=1}^{N} \hat{Z}_{y,i}\hat{b}_i \tag{9}$$

where $\hat{b}_i$ is the ith smallest of the N bids obtained from auction x, and,

$$\hat{Z}_{y,i} = \left(1 - \frac{i-1}{N}\right)\frac{y'(\frac{i-1}{N})}{x'(\frac{i-1}{N})} - \left(1 - \frac{i}{N}\right)\frac{y'(\frac{i}{N})}{x'(\frac{i}{N})}.$$

A consequence of (9) is that our estimator for P_y is linear in y', which, if y is a position auction, in turn depends linearly on the position weights. Therefore a relative error in estimating the position weights in y translates into the same relative error in estimating P_y, in addition to the error bounds given below.

4.2 Error bound for the multi-unit revenues

We will now develop an error bound for the estimator for Lemma 4.1 for the case of the multi-unit revenues. In the following, denote the allocation rule of the highest-k-bids-win auction as x_k (for an implicit number n of agents), and let $Z_k = Z_{x_k}$. We can therefore express the error in the estimation of P_k in terms of the error in estimating the bid distribution.

$$|\hat{P}_k - P_k| \quad = \quad \left|\mathbf{E}_q\left[-Z_k'(q)(\widehat{b}(q) - b(q))\right]\right| \tag{10}$$

A weak upper bound on the error. We first give a simple but weak analysis of the error. The following inequality separates the error bound into two components, each of which we can bound separately.

$$|\hat{P}_k - P_k| \quad = \quad \left|\mathbf{E}_q\left[-Z_k'(q)(\widehat{b}(q) - b(q))\right]\right| \leq \mathbf{E}_q\left[|Z_k'(q)|\right] \sup_q |\widehat{b}(q) - b(q)|$$

So, the mean absolute error in P_k can be written as:

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_k - P_k|\right] \quad \leq \quad \mathbf{E}_q\left[|Z_k'(q)|\right] \, \mathbf{E}_{\hat{b}}\left[\sup_q |\hat{b}(q) - b(q)|\right]$$

Lemma 3.2 gives the following bound on the mean absolute error for the bid function with the second inequality following from the fact that $b'(q) = v(q)x'(q)$ and $v(q) \leq 1$:

$$\mathbf{E}_{\hat{b}}\left[\sup_q |\hat{b}(q) - b(q)|\right] \quad \leq \quad \frac{\sup_q b'(q)}{\sqrt{2N}} \leq \frac{\sup_q x'(q)}{\sqrt{2N}}$$

We now proceed to bound $\mathbf{E}_q[|Z'_k(q)|]$. We recall that x is a convex combination over the allocation rules of the multi-unit highest-bids-win auctions, and use this to prove that Z_k has a single local maximum (see the appendix for a proof).

Lemma 4.2. *Let x_k denote the allocation function of the k-highest-bids-win auction and x be any convex combination over the allocation functions of the multi-unit auctions. Then the function $Z_k(q) = (1-q)\frac{x'_k(q)}{x'(q)}$ achieves a single local maximum for $q \in [0,1]$.*

Let $Z_k^* = \sup_q Z_k(q)$. Then, we can bound $\mathbf{E}_q[|Z'_k(q)|]$ by $2Z_k^*$. We get the following theorem:

Proposition 4.3. *Let x_k denote the allocation function of the k-highest-bids-win auction and x be the allocation function of any rank-based auction. Then for all k, the mean absolute error in estimating P_k from N samples from the bid distribution for an all-pay auction with allocation rule x is bounded by:*

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_k - P_k|\right] \leq \sqrt{\frac{2}{N}}\ \sup_q\{x'(q)\}\ \sup_q\left\{\frac{x'_k(q)}{x'(q)}\right\}$$

To understand this error bound, we note that the maximum slope of the multi-unit allocation rules x_k, and therefore also that of any rank-based auction, is always bounded by n, the number of agents in the auction (summarized as Fact 4.4, below). This bounds the $\sup_q x'(q)$ term.

Fact 4.4. *The maximum slope of the allocation rule x of any n-agent position auction is bounded by n: $\sup_q x'(q) \leq n$. The maximum slope of the allocation rule x_k for the n-agent highest-k-bids-win auction is bounded by*

$$\sup_q x'_k(q) \in \left[\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\right] \frac{n-1}{\sqrt{\min\{k-1, n-k\}}} = \Theta\left(\frac{n}{\sqrt{\min\{k, n-k\}}}\right).$$

The quantity $\sup_q\left\{\frac{x'_k(q)}{x'(q)}\right\}$ on the other hand may be rather large, even unbounded, if $x'(\cdot)$ is near zero at some q. In Section 6.1 we take the approach of explicitly designing x with position weight $w'_k > \epsilon$ (i.e., to mix x_k into x) to ensure this latter term is bounded.

A stronger bound on the error. We will now develop a stronger bound on the error in P_k with better dependence on the quantity $\sup_q\left\{\frac{x'_k(q)}{x'(q)}\right\}$. We start with Equation (10) and partition the expectation on the right hand side of the equation as follows for some $\alpha > 0$:

$$\begin{aligned} |\hat{P}_k - P_k| &= \left|\mathbf{E}_q\left[-Z'_k(q)(\widehat{b}(q) - b(q))\right]\right| \\ &\leq \mathbf{E}_q\left[\frac{(\log(1+Z_k(q)))^\alpha}{Z_k(q)}|Z'_k(q)|\right] \sup_q\left|\frac{Z_k(q)}{(\log(1+Z_k(q)))^\alpha}(\hat{b}(q) - b(q))\right| \end{aligned}$$

An appropriate choice of α gives us the following theorem. We defer the proof to the appendix.

Theorem 4.5. *Let x_k denote the allocation function of the k-highest-bids-win auction and x be the allocation function of any rank-based auction. Then for all k, the mean absolute error in estimating P_k from N samples from the bid distribution for an all-pay auction with allocation rule x is bounded by:*

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_k - P_k|\right] \leq \frac{20}{\sqrt{N}} \sup_q\{x_k'(q)\} \log \max\left\{\sup_{q:x_k'(q)\geq 1} \frac{x'(q)}{x_k'(q)}, \sup_q \frac{x_k'(q)}{x'(q)}\right\}.$$

Invoking Fact 4.4, we note that the error in P_k given by Theorem 4.5 can be bounded by

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_k - P_k|\right] = O\left(\frac{n}{\sqrt{\min\{k, n-k\}}}\right) \frac{1}{\sqrt{N}} \log \max\left\{\sup_{q:x_k'(q)\geq 1} \frac{x'(q)}{x_k'(q)}, \sup_q \frac{x_k'(q)}{x'(q)}\right\}.$$

To understand the above error bound better, we make the following observations:

- When $x = x_k$, we get an error bound of $O(\sup_q x_k'(q)/\sqrt{N})$, which is the same (within constant factors) as the statistical error in bids.
- We also get a good error bound when x and x_k are close enough without being identical: when $\epsilon x_k' \leq x' \leq x_k'/\epsilon$, we get a bound of $O(\log(1/\epsilon)\sup_q x_k'(q)/\sqrt{N})$.
- Finally, as long as $x' \geq \epsilon x_k'$, that is, the highest-k-bids-win auction is mixed in with ϵ probability into x, we observe via Fact 4.4 that $\sup_{q:x_k'(q)\geq 1} x'(q)/x_k'(q) \leq \sup_q x'(q) \leq n$, and obtain an error bound of $O(\log(n/\epsilon)\sup_q x_k'(q)/\sqrt{N})$.

4.3 Error bound for arbitrary rank-based revenues

We now develop an error bound for our estimator for the revenue, P_y, of an arbitrary position auction with allocation rule y from the bids of another position auction x. Let us write y as a position auction with weights $\boldsymbol{w}$:

$$\begin{aligned} y &= \textstyle\sum_k w_k' \, x_k \\ P_y &= \textstyle\sum_k w_k' \, P_k \end{aligned}$$

Accordingly, the error in P_y is bounded by a weighted sum of the error in P_k:

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_y - P_y|\right] \leq \textstyle\sum_k w_k' \, \mathbf{E}_{\hat{b}}\left[|\hat{P}_k - P_k|\right]$$

applying Theorem 4.5,

$$\begin{aligned} &\leq \frac{20}{\sqrt{N}} \textstyle\sum_k w_k' \sup_q\{x_k'(q)\} \left(\log n + \log \frac{1}{w_k'} + \log \sup_q \frac{y'(q)}{x'(q)}\right) \\ &= O\left(\frac{n \log n}{\sqrt{N}} \log \sup_q \frac{y'(q)}{x'(q)}\right) \end{aligned} \tag{11}$$

Unfortunately, the above bound can be quite loose, as the following simple example demonstrates. Suppose that $x = y$ and $w_k' = 1/n$ for all k. Then the above approach (via a tighter bound on the sum over k) leads to an error bound of $O(\sqrt{n}\log n/\sqrt{N})$, whereas, the true error bound should be $O(1/\sqrt{N})$, arising due to the statistical error in bids. Furthermore, it is desirable to obtain an error bound that depends directly on $\sup_q y'(q)$, rather than on the constituent $\sup_q x'_k(q)$; The latter can be much larger than the former. Below, we analyze the error in P_y directly, leading to a slightly tighter bound.

Theorem 4.6. *The expected absolute error in estimating the revenue of a position auction with allocation rule y using N samples from the bid distribution for an all-pay position auction with allocation rule x is bounded as below; Here n is the number of positions in the two position auctions.*

$$\begin{aligned}\mathbf{E}_{\hat{b}}\left[|\hat{P}_y - P_y|\right] \le \frac{20}{\sqrt{N}}\sqrt{n\log n}\ \sup_q\{y'(q)\}\ \log\max\left\{\sup_{q:y'(q)\ge 1}\frac{x'(q)}{y'(q)}, \sup_q\frac{y'(q)}{x'(q)}\right\}\\ + \frac{O(1)}{N}\sup_q\{x'(q)\}\ \sup_q\left\{\frac{y'(q)}{x'(q)}\right\}\end{aligned}$$

Note that the first term in the error bound in Theorem 4.6 dominates, and this term is identical to the bound in Theorem 4.5, except for an extra $\sqrt{n\log n}$ term. Moreover, when $\sup_q y'(q) < \sqrt{n}$, Theorem 4.6 gives us a tighter error bound than Equation (11). We will now prove the theorem. Here we provide an outline for the proof; The complete argument can be found in the appendix.

Proof Sketch: As for the multi-unit revenues,

$$|\hat{P}_y - P_y| = \left|\mathbf{E}_q\left[Z_y(q)(\widehat{b}'(q) - b'(q))\right]\right| = \left|\mathbf{E}_q\left[Z'_y(q)(\widehat{b}(q) - b(q))\right]\right|$$

In order to simplify our analysis of the error in P_y, we will break up the error into two components: the bias in the estimator $\hat{P}_y$ and the deviation of $\hat{P}_y$ from its mean.

$$\begin{aligned}|\hat{P}_y - P_y| &\le \left|\hat{P}_y - \mathbf{E}\left[\hat{P}_y\right]\right| + \left|\mathbf{E}\left[\hat{P}_y\right] - P_y\right|\\ &= \left|\mathbf{E}_q\left[Z'_y(q)(\widehat{b}(q) - \tilde{b}(q))\right]\right| + \left|\mathbf{E}_q\left[Z'_y(q)(\tilde{b}(q) - b(q))\right]\right| \qquad (12)\end{aligned}$$

Here, $\tilde{b}$ is a step function that equals the expectation of the empirical bid function $\widehat{b}$:

$$\tilde{b}(q) = \mathbf{E}\left[\widehat{b}(q)\right]$$

The bias of the estimator $\hat{P}$ (i.e. the second term in (12)) is easy to bound. We prove in the appendix that $\tilde{b}(q) - b(q)$ is at most $O(1)/N$ times $\sup_q\{x'(q)\}$. This implies the following lemma.

Lemma 4.7. *With $\tilde{b}$ defined as above,*

$$\left|\mathbf{E}\left[\hat{P}_y\right] - P_y\right| = \left|\mathbf{E}_q\left[Z'_y(q)(\tilde{b}(q) - b(q))\right]\right| = \frac{O(1)}{N}\sup_q\{x'(q)\}\ \sup_q\left\{\frac{y'(q)}{x'(q)}\right\}.$$

We now focus on the first term in (12), namely the integral over the quantile axis of $Z'_y(q)(\widehat{b}(q) - \tilde{b}(q))$. The approach of Section 4.2 does not provide a good upper bound on this quantity, because a counterpart of Lemma 4.2 fails to hold for Z_y. Instead, we will express the integral as a sum over several independent terms, and show that it is small in expectation.

To this end, we first identify the set of quantiles at which the function $\widehat{b}$ "crosses" the function $\tilde{b}$ from below. This set is defined inductively. Define $i_0 = 0$. Then, inductively, let i_ℓ be the smallest integer strictly greater than $i_{\ell-1}$ such that

$$\widehat{b}\left(\frac{i_\ell - 1}{N}\right) \leq \tilde{b}\left(\frac{i_\ell - 1}{N}\right) \text{ and } \widehat{b}\left(\frac{i_\ell}{N}\right) > \tilde{b}\left(\frac{i_\ell}{N}\right).$$

Let i_{k-1} be the last integer so defined, and let $i_k = N$. Let I denote the set of indices $\{i_0, \cdots, i_k\}$. Let $\mathbf{T}_{i,j}$ denote the following integral:

$$\mathbf{T}_{i,j} = \int_{q=i/N}^{q=j/N} Z'_y(q)(\widehat{b}(q) - \tilde{b}(q))\,\mathrm{d}q$$

Then, our goal is to bound the quantity $\mathbf{E}_{\hat{b}}[|\mathbf{T}_{0,N}|]$ where $\mathbf{T}_{0,N}$ can be written as the sum:

$$\mathbf{T}_{0,N} = \sum_{\ell=0}^{\ell=k-1} \mathbf{T}_{i_\ell, i_{\ell+1}}.$$

We now claim that conditioned on I and the maximum bid error, this is a sum over independent random variables. In the following, let G denote the bid distribution, and $\hat{G}$ the empirical bid distribution.

Lemma 4.8. *Conditioned on the set of indices I and $\Delta = \sup_q |\hat{G}(b(q)) - G(b(q))|$, over the randomness in the bid sample, the random variables $\mathbf{T}_{i_\ell, i_{\ell+1}}$ are mutually independent.*

Then we apply Chernoff-Hoeffding bounds, coupled with the approach from Section 4.2 to bound each individual $\mathbf{T}_{i_\ell, i_{\ell+1}}$, to obtain a bound on the proability that $\mathbf{E}_{\hat{b}}[|\mathbf{T}_{0,N}| \,|\, I, \Delta]$ exceeds some value $a > 0$.

Lemma 4.9. *With $I = \{i_0, \cdots, i_k\}$ and $T_{i,j}$ defined as above, for any $a > 0$,*

$$\Pr\left[|\mathbf{T}_{0,N}| \geq a \,|\, I, \Delta\right] \leq exp\left(-\frac{a^2}{n(40\Delta C)^2}\right), \quad \textit{where } C = \sup_q\{y'(q)\} \log \max\left\{\sup_{q: y'(q) \geq 1} \frac{x'(q)}{y'(q)}, \sup_q \frac{y'(q)}{x'(q)}\right\}.$$

Combining this lemma with an absolute bound on $|\mathbf{T}_{0,N}|$, and removing the conditioning on I and Δ, we obtain the Theorem 4.6. $\square$

5 Simulation evidence

We now present some simulation evidence to support our theoretical results. Our focus will be the inference for the revenues from an auction with a particular allocation rule $y(\cdot)$ from the observations of bidding in an auction with another allocation rule $x(\cdot)$ which is constructed as mixture of the auction with the allocation rule $y(\cdot)$ with weight $0 < \epsilon < 1$ and another allocation rule with weight $1 - epsilon$. We consider the following three designs.

- **Design 1:** $y(q) = q$ (uniform allocation rule), $x(q) = (1-\epsilon)\,q^{n-1} + \epsilon\, q$ (a mixture of the 1 unit auction and the uniform allocation rule).

- **Design 2:** $y(q) = q^{n-1}$ (1 unit auction allocation rule), $x(q) = (1-\epsilon)q + \epsilon\, q^{n-1}$ (a mixture of the 1 unit auction and the uniform allocation rule).

- **Design 3:** $y(q) = q^{n-1}$ (1 unit auction allocation rule), $x(q) = (1-\epsilon)\,(1-(1-q)^{n-1}) + \epsilon\, q^{n-1}$ (a mixture of the 1 unit auction and the $(n-1)$-unit auction allocation rule).

Our focus in this exercise is the inference for P_y for the allocation rule defined by each of th e three designs. The simulation is designed in the following way. We consider the grid over quantiles with the cell size Δ such that for each grid point p $q_{(p)} = \Delta\, p$. For a given number of players n we compute the allocation functions for each allocation rule $x(\cdot)$ and $y(\cdot)$ on the grid yielding $x_{(p)} = x(q_{(p)})$ and $y_{(p)} = y(q_{(p)})$ for each grid point.

The derivatives of allocation rules $x(\cdot)$ and $y(\cdot)$ are computed analytically. Function $Z_y(\cdot)$ and its first derivative are approximated on the same grid as

$$Z_{y(p)} = (1-\Delta\, p)\frac{y_p'}{x_p'} \text{ and } Z_{y(p)}' = \frac{Z_{y(p+1)} - Z_{y(p-1)}}{2\Delta}.$$

We take the distribution of values to be the beta distribution with parameters $\alpha = \beta = 2$. This distribution of values is supported on $[0,1]$, it is unimodal with the mode and the mean at $1/2$ and it is symmetric about the mean.

To compute the equilibrium distribution of bids, we compute the quantile function of the value distribution on the grid as

$$v_{(p)} = v(\Delta\, p).$$

Then for the all-pay auction with the allocation rule $x(\cdot)$ we compute the approximated quantile function of the bid distribution $b(\cdot)$ as

$$b_{(p)} = \sum_{l \leq p} v_{(l)}(x_{(l)} - x_{(l-1)}) \approx \int_0^{\Delta\, p} v(q)x'(q)\, dq.$$

Using the same grid summation technique we compute the "true" revenue from the k-unit auction of interest as

$$P_y = \sum (1-\Delta\, p) v_{(p)} y_p'.$$

To perform simulations, at each simulation round, we generate a sample of bids from the equilibrium distribution of bids by drawing uniformly a vector of N integers between 1 and $[1/\Delta]$ and for each such integer s the corresponding simulated draw of the bid is $b_{(s)}$. In each N-sample $\{b_i\}_{i=1}^N$ we order the bids and then compute the estimator for P_k as

$$\widehat{P}_y = -\frac{1}{N}\sum_{i=1}^{N} Z_{y(i)}' b^{(i:N)}.$$

For each choice of the number of players n and the size of the sample used N we re-sample the bids from the bid distribution $N_s = 1000$ times and then compute the average mean absolute deviation of the estimated $\widehat{P}_y$ from the computed P_y across simulations.

The simulation results for various choices of the number of players and the number of auctions are presented in Table 5 for different designs.

Table 1: Mean absolute deviation for $\widehat{P}_y$ across Monte-Carlo simulations normalized by $\sqrt{N/n}$

$y(q) = q, \quad x(q) = (1-\epsilon)x^{(1:n)}(q) + \epsilon\, y(q)$							
$n =$	$N =$						Theorem 4.5 upper bound
	2	10^1	10^2	10^3	10^4	10^5	
2^2	1.2534	0.3874	0.3625	0.3573	0.3535	0.3481	2.7111
2^3	0.3915	0.4336	0.4829	0.4994	0.5041	0.4918	2.3479
2^4	0.3025	0.3634	0.3982	0.4045	0.4088	0.4187	1.9170
2^5	0.1831	0.2690	0.2825	0.2821	0.2798	0.2901	1.5155
2^6	0.1437	0.2369	0.1929	0.1917	0.1898	0.1890	1.1739
2^7	0.1316	0.1902	0.159	0.1432	0.1407	0.1441	0.9431
2^8	0.1276	0.1462	0.149	0.1248	0.1226	0.1198	0.8155
2^9	0.1247	0.1267	0.1488	0.1207	0.1100	0.11500	0.6884
2^{10}	0.1215	0.1176	0.1616	0.1165	0.1114	0.1124	0.5702
$y(q) = x^{(1:n)}(q), \quad x(q) = (1-\epsilon)\, q + \epsilon\, y(q)$							
$n =$	$N =$						Theorem 4.5 upper bound
	2	10^1	10^2	10^3	10^4	10^5	
2^2	0.1416	0.0716	0.0586	0.0589	0.0578	0.059	3.4539
2^3	0.1281	0.0789	0.0637	0.0643	0.0670	0.0624	4.0295
2^4	0.079	0.0821	0.0596	0.0584	0.0591	0.0583	4.3173
2^5	0.0427	0.0687	0.0488	0.0472	0.0452	0.0490	4.4613
2^6	0.0224	0.0483	0.0369	0.0337	0.0337	0.0355	4.5332
2^7	0.0116	0.0259	0.0290	0.0237	0.0232	0.0236	4.5692
2^8	0.0059	0.0132	0.0230	0.0162	0.0157	0.0156	4.5872
2^9	0.003	0.0067	0.0186	0.0115	0.0107	0.0104	4.5962
2^{10}	0.0015	0.0034	0.0107	0.0094	0.0084	0.0081	4.6007
$y(q) = x^{(1:n)}(q), \quad x(q) = (1-\epsilon)\, x^{(n-1:n)}(q) + \epsilon\, y(q)$							
$n =$	$N =$						Theorem 4.5 upper bound
	2	10^1	10^2	10^3	10^4	10^5	
2^2	0.1522	0.1978	0.2127	0.2104	0.2113	0.2101	3.4539
2^3	0.1177	0.1200	0.1061	0.1056	0.1064	0.0998	4.0295
2^4	0.0759	0.1116	0.1067	0.1046	0.1019	0.1029	4.3173
2^5	0.0424	0.0866	0.0817	0.0798	0.0809	0.08	4.4613
2^6	0.0224	0.0499	0.0604	0.0602	0.0604	0.0581	4.5332
2^7	0.0116	0.0259	0.0474	0.0443	0.0434	0.0447	4.8063
2^8	0.0059	0.0132	0.0356	0.0323	0.0324	0.0322	5.5196
2^9	0.0030	0.0067	0.0209	0.0226	0.0226	0.0235	6.2242
2^10	0.0015	0.0034	0.0107	0.0171	0.0168	0.0161	6.9237

The values in Table 5 are normalized by the factor $\sqrt{N/n}$ that reflects the dependence of the median absolute error in the estimation of P_y from the overall sample size. By replicating the Monte carlo sampling we ensured that the Monte Carlo sample size leads to the relative error of at most 6%. From the table we note that for a given row (corresponding to a fixed number of bidders in the auction n) the normalized mean absolute deviation does not significantly vary across different numbers of samples N. However, there is a visible dependence of the normalized mean absolute deviation from the number of players n. For the design where the auction of interest is the uniform allocation rule demonstrates an initial increase in the normalized mean squared error which then stabilizes to a constant for large n. The designs where the allocation rule is the 1-unit n-player auction allocation rule, the normalized mean absolute error decreases with n reflecting the convergence of the revenue from that allocation rule to the upper boundary of support of the distribution of values. The table also contrasts the empirical performance of our estimator with the theoretical upper bound given in Theorem 4.5. The bound exceeds the measured mean squared error, suggesting that it can be tightened further as a function of the number of bidders in each auction if we elaborate the actual bound for the specific auction designs.

We also illustrate the dependence of the estimation error on the choice of the mixture weight ϵ for the three considered designs. We fix the number of players $n = 32$ and the sample size $N = 1000$ and choose the grid over ϵ varying it from 0 to 1 excluding the end points. On Figure 1 we demonstrate the dependence of the median absolute error computed as the ratio of the median absolute error of estimated revenue over the course of 1000 Monte Carlo replications and the theoretical revenue computed using the numerical integration. The figure demonstrates that for the third design where the auction used for inference is the mixture of 1 and $n-1$-unit auctions yields an approximately equal relative error over different values of ϵ. At the same time, design 1 (where a mixture of a 1-unit auction and an auction with uniform allocation rule is used to infer the revenue from the auction with the uniform allocation rule) demonstrates an increase in the error with the growth of ϵ, while the design 2 leads to an opposite tendency.

6 Applications to A/B testing

We now discuss applications of the inference approach we developed in Section 4.

6.1 Estimating revenues of novel mechanisms

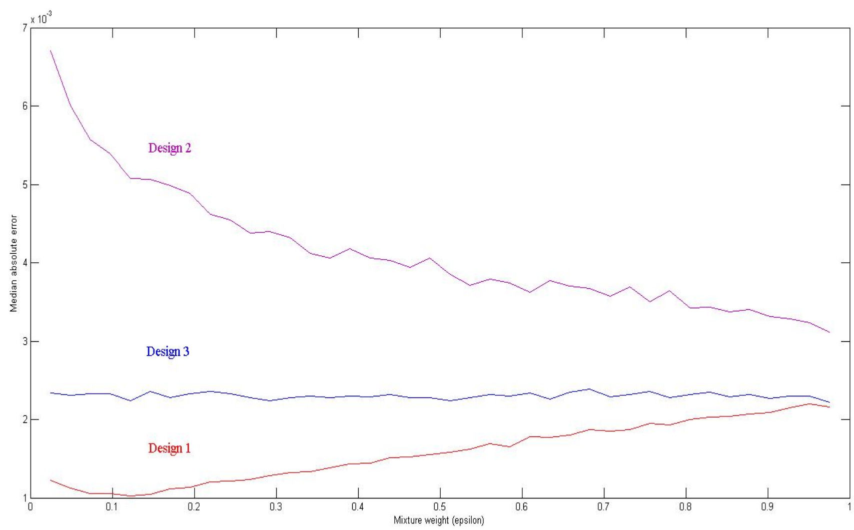
Let us consider the setup described in the introduction where an auction house running auction A would like to determine the revenue of a novel mechanism B. The typical approach for doing so is to run the auction B with some probability $\epsilon > 0$ and A with the remaining probability. Ideally, if in doing so, the auction house obtains ϵN bids in response to the auction B out of a total of N bids, the revenue of B can be estimated within an error bound of

$$\Theta\left(\frac{1}{\sqrt{\epsilon}}\right)\frac{\sup_q\{x'_B(q)\}}{\sqrt{N}} \tag{13}$$

where x_B denotes the allocation rule corresponding to B.

In practice, however, instead of obtaining bids in response to B alone, the seller obtains bids in response to the aggregate mechanism $C = (1-\epsilon)A + \epsilon B$. We can then use (9) to estimate the revenue of B. As a consequence of Theorem 4.6, and noting that for positions auctions with

Figure 1: Dependence of the relative median absolute error from the mixture weight ϵ



The graph shows the dependence of the median absolute error in estimation of the revenue of auction with allocation rule $y(\cdot)$ from the sample generated by the auction with allocation rule $x(\cdot)$ for one of the three designs. The sample size is fixed at 1000 while the number of bidders per auction is chosen to be 32. Each value on the graph is the ratio of the median absolute error and the theoretical revenue. The values are obtained using 1000 independent draws from the distribution of valuations.

n positions, $x'_C(q) \leq n$ and $x'_B(q)/x'_C(q) \leq 1/\epsilon$ for all quantiles q, we obtain the following error bound

Corollary 6.1. *The revenue of a rank based mechanism B can be estimated from N bids of a rank-based mechanism $C = (1-\epsilon)A + \epsilon B$ with absolute error bounded by*

$$O(1)\sqrt{n \log n}\log(n/\epsilon)\ \frac{\sup_q\{x'_B(q)\}}{\sqrt{N}}. \tag{14}$$

Relative to the ideal situation described above, our error bound has a better dependence on ϵ and a worse dependence on n. Note that when ϵ is very small, our error bound (14) may be smaller than the ideal bound in (13). This is not surprising: the ideal bound ignores information that we can learn about the revenue of B from the $(1-\epsilon)N$ bids obtained when B is not run.

When B is a multi-unit auction, we obtain a better error bound which is closer to the ideal bound in (13).

Corollary 6.2. *The revenue of the highest-k-bids-win mechanism B can be estimated from N bids of a rank-based mechanism $C = (1-\epsilon)A + \epsilon B$ with absolute error bounded by*

$$O(1)\log(n/\epsilon)\ \frac{\sup_q\{x'_B(q)\}}{\sqrt{N}}. \tag{15}$$

6.2 Comparing revenues

We have considered the case where the empirical task was to recover the revenues for one mechanism (y) using the sample of bids responding to another mechanism (x). In many practical situations the empirical task is simply the verification of whether the revenue from a given mechanism is higher than the revenue from another mechanism. Or, equivalently, the task could be to verify whether one mechanism provides revenue which is a certain percentage above that of another mechanism. We now demonstrate that this is a much easier empirical task in terms of accuracy than the task of inferring the revenue.

Suppose that we want to compare the revenues of mechanisms B_1 and B_2 by mixing them in to an incumbent mechanism A, and running the composite mechanism $M = \epsilon B_1 + \epsilon B_2 + (1-2\epsilon)A$. Specifically, we would like to determine whether $P_{B_1} > \alpha P_{B_2}$ for some $\alpha > 0$. Consider a binary classifier $\widehat{\gamma}$ which is equal to 1 when $P_{B_1} > \alpha P_{B_2}$ and 0 otherwise. Let $\gamma = \mathbf{1}\{P_{B_1} - \alpha\, P_{B_2} > 0\}$ be the corresponding "ideal" classifier for the case where the distribution of bids from mechanism M is known precisely. To evaluate the accuracy of the classifier, we need to evaluate the probability $\Pr(\widehat{\gamma} = 1|\gamma = 0)$, and likewise, $\Pr(\widehat{\gamma} = 0|\gamma = 1)$. The classifier will give the wrong output if the sampling noise in estimating $\hat{P}_{B_1} - \alpha\, \hat{P}_{B_2}$ is greater than $|P_{B_1} - \alpha\, P_{B_2}|$.

Our main result of this section says that keeping α, the difference $|P_{B_1} - \alpha\, P_{B_2}|$, and the number of positions n constant, the probability of incorrect output decreases exponentially with the number of bids N.

Theorem 6.3. *Suppose that N bids from a mechanism $M = \epsilon B_1 + \epsilon B_2 + (1-2\epsilon)A$ for arbitrary rank-based mechanisms B_1, B_2, and A, are used to estimate the classifier $\gamma = \mathbf{1}\{P_{B_1} - \alpha\, P_{B_2} > 0\}$ that establishes whether the revenue of mechanism B_1 exceeds α times the revenue of mechanism B_2. Then the error rate of the binary classifier is bounded from above by*

$$\exp\left(-O\left(\frac{Na^2}{\alpha^2 n^3 \log(n/\epsilon)}\right)\right),$$

where $a = |P_{B_1} - \alpha\, P_{B_2}|$. In other words, once the number of samples is polynomially large in n, the error rate decreases exponentially with the number of samples.

Proof. We need to bound the probability that the error in estimating $\hat{P}_{B_1} - \alpha\hat{P}_{B_2}$ is greater than $|P_{B_1} - \alpha\, P_{B_2}|$. This error can in turn be decomposed into the error in estimating P_{B_1} and that in estimating P_{B_2}. Denote $a = |P_{B_1} - \alpha\, P_{B_2}| > 0$. Then,

$$\Pr\left(|(\widehat{P}_{B_1} - \alpha\,\widehat{P}_{B_2}) - (P_{B_1} - \alpha\, P_{B_2})| > a\right) \le \Pr\left(|\widehat{P}_{B_1} - P_{B_1}| > a/2\right) + \Pr\left(|\widehat{P}_{B_2} - P_{B_2}| > a/2\alpha\right).$$

Let x denote the allocation rule of the mechanism M that we are running, and let b be the corresponding bid function. Now, recall that for

$$\Delta = \sup_q |\widehat{G}(b(q)) - G(b(q))|, \text{ and } C = \sup_q\{x'_{B_1}(q)\} \log\max\left\{\sup_{q:x'_{B_1}(q)\ge 1}\frac{x'(q)}{x'_{B_1}(q)}, \sup_q \frac{x'_{B_1}(q)}{x'(q)}\right\},$$

Equation (12), Lemma 4.7, and Lemma 4.9 together imply that (conditional on Δ)

$$\Pr\left(|\widehat{P}_{B_1} - P_{B_1}| > a/2\right) \le 2\exp\left(-\frac{1}{n(40\,\Delta\, C)^2}\left(\frac{a}{2} - O\left(\frac{n}{\epsilon N}\right)\right)^2\right).$$

Now recall that $C < n\log(n/\epsilon)$, and $\Delta < \text{constant}/\sqrt{N}$ with high probability (Lemma 3.2). As a result, we establish that for $a = \omega(n/\epsilon N)$,

$$\Pr\left(|\widehat{P}_{B_1} - P_{B_1}| > a/2\right) \le \exp\left(-O\left(\frac{Na^2}{n^3\log(n/\epsilon)}\right)\right).$$

Likewise,

$$\Pr\left(|\widehat{P}_{B_2} - P_{B_2}| > a/2\alpha\right) \le \exp\left(-O\left(\frac{Na^2}{\alpha^2 n^3\log(n/\epsilon)}\right)\right).$$

□

We obtain a similar error bound when our goal is to estimate which of r different novel mechanisms obtains the most revenue, for any $r > 1$:

Corollary 6.4. *Suppose that our goal is to determine which of r position auctions, $B_1, B_2, \cdots, B_r$, obtains the most revenue while running incumbent mechanism A, by running each of the novel mechanisms with probability ϵ/r. Then the error probability of the corresponding classifier constructed using N bids from composite mechanism $M = \sum_{i=1}^{r}\epsilon/rB_i + (1-\epsilon)A$ is bounded from above by*

$$r\exp\left(-O\left(\frac{Na^2}{n^3\log(rn/\epsilon)}\right)\right),$$

where a is the absolute difference between the revenue obtained by the best two of the r mechanisms.

7 Applications to optimization

In the previous sections we discussed the econometric properties of rank-based auctions, showing that the revenue of a rank-based auction can be estimated with small error using bids from another rank-based auction. In this section we shift our focus to instrumented optimization, asking: how can we use inference to optimize for the seller's revenue. In Section 7.1 we develop a theory for optimizing revenue over the class of all rank-based auctions that resembles Myerson's theory for optimal auction design. Our optimization requires knowing all of the multi-unit revenues P_k. Where Myerson's theory employs ironing by value and value reserves, our approach analogously employs ironing by rank and rank reserves. To implement this approach, in Section 7.2 we extend the approach of Section 4 to develop a "universal B-test" that can be used to estimate all of the multi-unit revenues P_k simultaneously. Finally, in Section 7.3 we show that the revenue of an optimal rank-based auction approximates the revenue of the optimal auction. Putting these together gives a simple uniform procedure for revenue optimization subject to inference.

We begin by recalling our framework for position environments and rank-based auctions. In a rank-based auction the allocation to an agent depends solely on the ordinal rank of his bid among other agents' bids, and not on the cardinal value of the bid. For a position environment, a rank-based auction assigns agents (potentially randomly) to positions based on their ranks. Consider a position environment given by non-increasing weights $\boldsymbol{w} = (w_1, \ldots, w_n)$. For notational convenience, define $w_{n+1} = 0$. Define the cumulative position weights $\boldsymbol{W} = (W_1, \ldots, W_n)$ as $W_k = \sum_{j=1}^{k} w_j$, and $W_0 = 0$. We can view the cumulative weights as defining a piece-wise linear, monotone, concave function given by connecting the point set $(0, W_0), \ldots, (n, W_n)$.

Multi-unit highest-bids-win auctions form a basis for position auctions. Consider the marginal position weights $\boldsymbol{w}' = (w'_1, \ldots, w'_n)$ defined by $w'_k = w_k - w_{k+1}$. The allocation rule induced by the position auction with weights $\boldsymbol{w}$ is identical to the allocation rule induced by the convex combination of multi-unit auctions where the k-unit auction is run with probability w'_k.

A randomized assignment of agents to positions based on their ranks induces an expected weight to which agents of each rank are assigned, e.g., $\bar{w}_k$ for the kth ranked agent. These expected weights can be interpreted as a position auction environment themselves with weights $\bar{\boldsymbol{w}}$. As for the original weights, we can define the cumulative position weights $\bar{\boldsymbol{W}}$ as $\bar{W}_k = \sum_{j=1}^{k} w_j$. An important issue for optimization of rank-based auctions is to characterize the inducible class of position weights.

Lemma 7.1 (e.g., Devanur et al., 2013)**.** *There is a rank-based auction with induced position weights $\bar{\boldsymbol{w}}$ for position environment with weights $\boldsymbol{w}$ if and only if their cumulative weights satisfy $\bar{W}_k \leq W_k$ for all k, denoted $\bar{\boldsymbol{W}} \leq \boldsymbol{W}$.*

Any feasible weights $\bar{\boldsymbol{w}}$ can be constructed from $\boldsymbol{w}$ by a sequence of the following two operations.

rank reserve For a given rank k, all agents with ranks between $k+1$ and n are rejected. The resulting weights $\bar{\boldsymbol{w}}$ are equal to $\boldsymbol{w}$ except $\bar{w}_{k'} = 0$ for $k' > k$.

iron by rank Given ranks $k' < k''$, the ironing-by-rank operation corresponds to, when agents are ranked, assigning the agents ranked in an interval $\{k', \ldots, k''\}$ uniformly at random to these same positions. The ironed position weights $\bar{\boldsymbol{w}}$ are equal to $\boldsymbol{w}$ except the weights on the ironed interval of positions are averaged. The cumulative ironed position weights $\bar{\boldsymbol{W}}$ are equal to $\boldsymbol{W}$ (viewed as a concave function) except that a straight line connects $(k'-1, \bar{W}_{k'-1})$ to $(k'', \bar{W}_{k''})$. Notice that concavity of $\boldsymbol{W}$ (as a function) and this perspective of the ironing procedure as replacing an interval with a line segment connecting the endpoints of the interval implies that $\boldsymbol{W} \geq \bar{\boldsymbol{W}}$ coordinate-wise, i.e., $W_k \geq \bar{W}_k$ for all k.

7.1 Optimal rank-based auctions

In this section we describe how to optimize for expected revenue over the class of rank-based auctions. Recall that rank-based auctions are linear combinations over k-unit auctions. The characterization of Bayes-Nash equilibrium, cf. equation (2), shows that revenue is a linear function of the allocation rule. Therefore, the revenue of a position auction can be calculated as the convex combination of the revenue P_k from the k-highest-bids-win auction for $k \in 1 \ldots n-1$.

Given these multi-unit revenues, $\boldsymbol{P} = (P_0, \ldots, P_n)$, the problem of designing the optimal rank-based auction is well defined: given a position environment with weights $\boldsymbol{w}$, find the weights $\bar{\boldsymbol{w}}$ for an rank-based auction with cummulative weights $\bar{\boldsymbol{W}} \leq \boldsymbol{W}$ maximizing the sum $\sum_k (\bar{w}_k - \bar{w}_{k+1}) P_k$. This optimization problem is isomorphic to the theory of envy-free optimal pricing developed by Hartline and Yan (2011). We summarize this theory below; a complete derivation can be found in Appendix A.

Define the *multi-unit revenue curve* as the piece-wise linear function connecting the points $(0, P_0), \ldots, (n, P_n)$. This function may or may not be concave. Define the *ironed multi-unit revenues* as $\bar{\boldsymbol{P}} = (\bar{P}_0, \ldots, \bar{P}_n)$ according to the smallest concave function that upper bounds the multi-unit revenue curve. Define the multi-unit marginal revenues, $\boldsymbol{P}' = P'_1, \ldots, P'_n$ and $\bar{\boldsymbol{P}}' = \bar{P}'_1, \ldots, \bar{P}'_n$, as the left slope of the multi-unit and ironed multi-unit revenue curves, respectively. I.e., $P'_k = P_k - P_{k-1}$ and $\bar{P}'_k = \bar{P}_k - \bar{P}_{k-1}$.

Theorem 7.2. *Given a position environment with weights $\boldsymbol{w}$, the revenue-optimal rank-based auction is defined by position weights $\bar{\boldsymbol{w}}$ that are equal to $\boldsymbol{w}$, except ironed on the same intervals as $\boldsymbol{P}$ is ironed to obtain $\bar{\boldsymbol{P}}$, and set to 0 at positions k for which $\bar{P}'_k$ is negative.*

As is evident from this description of the optimal rank-based auction, the only quantities that need to be ascertained to run this auction is the multi-unit revenue curve defined by $\boldsymbol{P}$. Therefore, an econometric analysis for optimizing rank-based auctions need not estimate the entire value distribution; estimation of the multi-unit revenues is sufficient.

Optimal rank-based auctions with strict monotonicity

Position auctions, by definition, have non-increasing position weights $\boldsymbol{w}$. The ironing in the iron-by-rank optimization of the preceding section converted the problem of optimizing multi-unit marginal revenue subject to non-increasing position weight, to a simpler problem of optimizing multi-unit marginal revenue without any constraints. In this section, we describe the optimization of rank-based auctions (i.e., ones for which position weights can be shifted only downwards or discarded) subject to *strictly decreasing* position weights. This strictness insures good inference properties, the details of which are formalized in Section 6.1 (see, e.g., Corollary 6.2).

As described by Lemma 7.1, position weights $\bar{\boldsymbol{w}}$ are feasible as a rank-based auction in position environment $\boldsymbol{w}$ if the cumulative position weights satisfy $W_k \geq \bar{W}_k$ for all k. Suppose we would like to optimize $\bar{\boldsymbol{w}}$ subject to strict monotonicity constraints: $\bar{w}'_k = \bar{w}_k - \bar{w}_{k+1} \geq \epsilon_k$ for all k, for some given $\epsilon_1, \cdots, \epsilon_n \geq 0$. We call an allocation rule satisfying these monotonicity constraints an $\boldsymbol{\epsilon}$-strictly-monotone allocation rule, where $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_n)$. As non-trivial ironing by rank always results in consecutive positions with the same weight, i.e., $\bar{w}'_k = 0$ for some k, the optimal rank-based mechanism will require overlapping ironed intervals.

To our knowledge, performance optimization subject to a strict monotonicity constraint has not previously been considered in the literature. At a high level our approach is the following. We start with $\boldsymbol{w}$ which induces the cumulative position weights $\boldsymbol{W}$ which constrain the resulting position weights $\bar{\boldsymbol{w}}$ of any feasible rank-based auction via its cumulative $\bar{\boldsymbol{W}}$. We view $\bar{\boldsymbol{w}}$ as the combination of two position auctions. The first has weakly monotone weights $\bar{\boldsymbol{y}} = (\bar{y}_1, \dots, \bar{y}_n)$; the second has strictly monotone weights $(E_1, E_2, \dots, E_n)$, where $E_i = \sum_{j \geq i} \epsilon_j$ for $1 \leq i < n$ and $E_n = 0$; and the combination has weights $\bar{w}_k = \bar{y}_k + E_k$ for all k. The revenue of the combined position auction is the sum of the revenues of the two component position auctions. Since the second auction has fixed position weights, its revenue is fixed. Since the first position auction is weakly monotone and the second is strictly, the combined position auction is strictly monotone and satisfies the constraint that $\bar{w}'_k \geq \epsilon_k$ for all k.

This construction focuses attention on optimization of $\bar{\boldsymbol{y}}$ subject to the induced constraint imposed by $\boldsymbol{w}$ and after the removal of the $\boldsymbol{\epsilon}$-strictly-monotone allocation rule. I.e., $\bar{\boldsymbol{w}}$ must be feasible for $\boldsymbol{w}$. The suggested feasibility constraint for optimization of $\bar{\boldsymbol{y}}$ is given by position weights $\boldsymbol{y}$ defined as $y_k = w_k - E_k$. Notice that, in this definition of $\boldsymbol{y}$, a lesser amount is subtracted from successive positions. Consequently, monotonicity of $\boldsymbol{w}$ does not imply monotonicity of $\boldsymbol{y}$.

To obtain $\bar{\boldsymbol{y}}$ from $\boldsymbol{y}$ we may need to iron for two reasons, (a) to make $\bar{\boldsymbol{y}}$ monotone and (b) to make the multi-unit revenue curve monotone. In fact, both of these ironings are good for revenue. The ironing construction for monotonizing $\boldsymbol{y}$ constructs the concave hull of the cumulative position weights $\boldsymbol{Y}$. This concave hull is strictly higher than the curve given by $\boldsymbol{Y}$ (i.e., connecting $(0, Y_0), \dots, (n, Y_n)$). Similarly the ironed multi-unit revenue curve given by $\bar{\boldsymbol{P}}$ is the concave hull of the multi-unit revenue curve given by $\boldsymbol{P}$. The correct order in which to apply these ironing procedures is to first (a) iron the position weights $\boldsymbol{y}$ to make it monotone, and second (b) iron the multi-unit revenue curve $\boldsymbol{P}$ to make it concave. This order is important as the revenue of the position auction with weights $\bar{\boldsymbol{y}}$ is only given by the ironed revenue curve $\bar{\boldsymbol{P}}$ when the $\bar{\boldsymbol{y}}' = 0$ on the ironed intervals of $\bar{\boldsymbol{P}}$.

Theorem 7.3. *The optimal $\boldsymbol{\epsilon}$-strictly-monotone rank-based auction for position weights $\boldsymbol{w}$ has position weights $\bar{\boldsymbol{w}}$ constructed by*

1. *defining $\boldsymbol{y}$ by $y_k = w_k - E_k$ for all k, where $E_i = \sum_{j \geq i} \epsilon_j$ for $1 \leq i < n$.*

2. *averaging position weights of $\boldsymbol{y}$ on intervals where $\boldsymbol{y}$ should be ironed to be monotone.*

3. *averaging the resulting position weights on intervals where $\boldsymbol{P}$ should be ironed to be concave to get $\bar{\boldsymbol{y}}$*

4. *setting $\bar{\boldsymbol{w}}$ as $\bar{w}_k = \bar{y}_k + E_k$.*

Proof. The proof of this theorem follows directly by the construction and its correctness. □

The rank-based auction given by $\bar{\boldsymbol{w}}$ in position environment given by $\boldsymbol{w}$ can be implemented by a sequence of iron-by-rank and rank-reserve operations. Such a sequence of operations can be found, e.g., via an approach of Alaei et al. (2012) or Hardy et al. (1929).

7.2 Universal B test

In Section 6.1 we discussed how to estimate the revenue of a single auction B from the bids of an auction C. We now consider the problem of estimating all of the multi-unit revenues P_k simultaneously from the bids of a single auction. What properties should the auction C have in order to enable this? We first note that, as a simple consequence of Corollary 6.2, it suffices to mix the k-unit auction for every k into C with some small probability. This gives us the following result.

Corollary 7.4. *N bids from a mechanism C with $x_C = (1-\epsilon)x_A + \sum_{k=1}^{n} \frac{\epsilon}{n} x_k$ can be used to simultaneously estimate* all *of the multi-unit revenues, and consequently all position auction revenues with absolute error bounded by*

$$O(1)\frac{n\log(n/\epsilon)}{\sqrt{N}}.$$

Next we observe that in fact we can get similar results by mixing in just a few of the multi-unit auctions. In particular, in order to estimate P_k accurately, it suffices to mix in a multi-unit auction with no more than k units, and another one with no less than k units. This gives us a more efficient universal B test for simultaneously inferring all of the multi-unit revenues (see Corollary 7.6).

Lemma 7.5. *The revenue of the highest-k-bids-win mechanism B can be estimated from N bids of a rank-based all-pay auction $C = (1-2\epsilon)A + \epsilon B_1 + \epsilon B_2$ where A is an arbitrary rank-based auction, and B_1 and B_2 are the highest-k_1-bids-win and highest-k_2-bids-win auctions respectively, with $k_1 \le k \le k_2$. The absolute error of the estimate is bounded by*

$$20(n + \log(1/\epsilon))\ \frac{\sup_q\{x'_k(q)\}}{\sqrt{N}}.$$

Proof. We begin by noting that for any j and k, $k \le j$,

$$\frac{x'_k(q)}{x'_j(q)} = \frac{\binom{j-1}{k-1}}{\binom{n-1-k}{n-1-j}} \left(\frac{q}{1-q}\right)^{j-k}$$

When $k \le j$ and $q \ge 1/2$, this ratio is less than 2^n. Likewise, we can show that when $k \ge j$ and $q \le 1/2$, the ratio $\frac{x'_k(q)}{x'_j(q)}$ is less than 2^n. Therefore, for any q, and $C = (1-2\epsilon)A + \epsilon B_1 + \epsilon B_2$ where B_1 and B_2 are the highest-k_1-bids-win and highest-k_2-bids-win auctions respectively, with $k_1 \le k \le k_2$, we have

$$\frac{x'_k(q)}{x'_C(q)} \le \frac{1}{\epsilon}2^n$$

Next we note that $\sup_q x'_C(q) \le n$, and therefore, $\sup_{q:x'_k(q)\ge 1} \frac{x'_C(q)}{x'_k(q)} \le n$. Putting these quantities together with Theorem 4.5, we get that the absolute error in estimating P_k from bids drawn from C is at most

$$\mathbf{E}_{\hat{b}}\Big[|\hat{P}_k - P_k|\Big] \le \frac{20}{\sqrt{N}}\ \sup_q\{x'_k(q)\}\ \ (n + \log 1/\epsilon)$$

∎

Corollary 7.6. *Given N bids of an all-pay auction $C = (1-2\epsilon)A + \epsilon B_1 + \epsilon B_{n-1}$ where A is an arbitrary rank-based auction, and B_1 and B_{n-1} are the highest-bid-wins and highest-$(n-1)$-bids-win auctions respectively, we can estimate the revenue of* any *rank-based auction with absolute error bounded by*

$$\frac{20n(n + \log(1/\epsilon))}{\sqrt{N}}.$$

7.3 Approximation via rank-based auctions

In this section we show that the revenue of optimal rank-based auction approximates the optimal revenue (over all auctions) for position environments. Instead of making this comparison directly we will instead identify a simple non-optimal rank-based auction that approximates the optimal auction. Of course the optimal rank-based auction of Theorem 7.2 has revenue at least that of this simple rank-based auction, thus its revenue also satisfies the same approximation bound.

Our approach is as follows. Just as arbitrary rank-based mechanisms can be written as convex combinations over k-highest-bids-win auctions, the optimal auction can be written as a convex combination over optimal k-unit auctions. We begin by showing that the revenue of optimal k-unit auctions can be approximated by multi-unit highest-bids-win auctions when the agents' values are distributed according to a regular distribution (Lemma 7.7, below). In the irregular case, on the other hand, rank-based auctions cannot compete against arbitrary optimal auctions. For example, if the agents' value distribution contains a very high value with probability $o(1/n)$, then an optimal auction may exploit that high value by setting a reserve price equal to that value; On the other hand, a rank-based mechanism cannot distinguish very well between values correspond to quantiles above $1 - 1/n$. We show that rank-based mechanisms can approximate the revenue of any mechanism that does not iron the quantile interval $[1-1/n, 1]$ (but may arbitrarily optimize over the remaining quantiles). Theorem 7.9 presents the precise statement.

Lemma 7.7. *For regular k-unit n-agent environments, there exists a $k' \leq k$ such that the highest-bid-wins auction that restricts supply to k' units (i.e., a rank reserve) obtains at least half the revenue of the optimal auction.*

Proof. This lemma follows easily from a result of Bulow and Klemperer (1996) that states that for agents with values drawn i.i.d. from a regular distribution the revenue of the k'-unit n-agent highest-bid-wins auction is at least the revenue of the k'-unit $(n-k')$-agent optimal auction. To apply this theorem to our setting, let us use $\mathbf{OPT}(k, n)$ to denote the revenue of an optimal k-unit n-agent auction, and recall that nP_k is the revenue of a k-unit n-agent highest-bids-win auction.

When $k \leq n/2$, we pick $k' = k$. Then,

$$nP_k \geq \mathbf{OPT}(k, n-k) \geq \frac{(n-k)}{n}\mathbf{OPT}(k, n) \geq \frac{1}{2}\mathbf{OPT}(k, n),$$

and we obtain the lemma. Here the first inequality follows from Bulow and Klemperer's theorem and the third from the assumption that $k \leq n/2$. The second inequality follows via by lower bounding $\mathbf{OPT}(k, n-k)$ by the following auction which has revenue exactly $\frac{(n-k)}{n}\mathbf{OPT}(k, n)$: simulate the optimal k-unit n-agent on the $n-k$ real agents and k fake agents with values drawn independently from the distribution. Winners of the simulation that are real agents contribute to revenue and the probability that an agent is real is $(n-k)/n$.

When $k > n/2$, we pick $k' = n/2$. As before we have:

$$nP_{n/2} \geq \mathbf{OPT}(n/2, n/2) = \frac{1}{2}\mathbf{OPT}(n,n) \geq \frac{1}{2}\mathbf{OPT}(k,n). \qquad \square$$

Lemma 7.8. *In any n-agent setting with an arbitrary (possibly irregular) value distribution with revenue curve $R(\cdot)$, and quantile $q \leq 1 - 1/n$, there exists an integer $k \leq (1-q)n$ such that the revenue of the k-highest-bids-win auction is at least a quarter of $nR(q)$, the revenue from posting a price of $v(q)$.*

Proof. First we get a lower bound on P_k for any k. For any value z, the total expected revenue of the k-highest-bids-win auction is at least zk times the probability that at least $k+1$ agents have value at least z. The median of a binomial random variable corresponding to n Bernoulli trials with with success probability $(k+1)/n$ is $k+1$. Thus, the probability that this binomial is at least $k+1$ is at least $1/2$. Combining these observations by choosing $z = v(1-(k+1)/n)$ we have,

$$n\, P_k \geq v(1-(k+1)/n)\, k/2.$$

Choosing $k = \lfloor (1-q)n \rfloor - 1$, for which $v(1-(k+1)/n) \geq v(q)$, the bound simplifies to,

$$n\, P_k \geq v(q)\, k/2.$$

The ratio of P_k and $R(q) = (1-q)\, v(q)$ is therefore at least

$$\frac{k}{2(1-q)n} > \frac{k}{2(k+2)},$$

which for $q \leq 1 - 3/n$ (or, $k \geq 2$) is at least $1/4$.

For $q \in (1-3/n, 1-1/n]$, we pick $k = 1$. Then, P_1 is at least $1/n$ times $v(q)$ times the probability that at least two agents have a value greater than or equal to $v(q)$. We can verify for $n \geq 2$ that

$$P_1 \geq \frac{v(q)}{n}\left(1 - q^n - n(1-q)q^{n-1}\right) \geq \frac{1}{4}(1-q)\, v(q).$$

$\square$

Theorem 7.9. *For regular value distributions and position environments, the optimal rank-based auction obtains at least half the revenue of the optimal auction. For any value distribution (possibly irregular) and position environments, the optimal rank-based auction obtains at least a quarter of the revenue of the optimal auction that does not iron or set a reserve price on the quantile interval $[1-1/n, 1]$.*

Proof. In the regular setting, the theorem follows from Lemma 7.7 by noting that the optimal auction (that irons by value and uses a value reserve) in a position environment is a convex combination of optimal k-unit auctions: since the revenue of each of the latter can be approximated by that of a k'-unit highest-bids-win auction with $k' \leq k$, the revenue of the convex combination can be approximated by that of the same convex combination over k'-unit highest-bids-win auctions; the resulting convex combination over k'-unit auctions satisfies the same position constraint as the optimal auction.

In the irregular setting, once again, any auction in a position environment is a convex combination of optimal k-unit auctions. The expected revenue of any k-unit auction is bounded from above by the expected revenue of the optimal auction that sells at most k items in expectation. The per-agent revenue of such an auction is bounded by $\bar{R}(1-k/n)$, the revenue of the optimal allocation rule with ex ante probability of sale k/n. Here $\bar{R}(\cdot)$ is the ironed revenue curve (that does not iron on quantiles in $[1-1/n, 1]$). $\bar{R}(1-k/n)$ is the convex combination of at most two points on the revenue curve $R(a)$ and $R(b)$, $a \leq 1-k/n \leq b < 1-1/n$. Now, we can use Lemma 7.8 to obtain an integer $k_a < n(1-a)$ such that P_{k_a} is at least a quarter of $R(a)$, likewise k_b for b. Taking the appropriate convex combination of these multi-unit auctions gives us a 4-approximation to the optimal auction k-unit auction (that does not iron over the quantile interval $[1-1/n, 1]$). Finally, the convex combination of the multi-unit auctions with k_a and k_b corresponds to a position auction with that is feasible for a k unit auction (with respect to serving the top k positions with probability one, service probability is only shifted to lower positions). $\square$

8 Inference methodology and error bounds for first-price auctions

While most of our analysis so far has focused on the case of all-pay auctions, our methodology and results extend as-is to first-price auctions as well. Here we sketch the differences between the two cases.

Recall that in a first-price auction, we can obtain the value distribution from the bid distribution as follows: $v(q) = b(q) + x(q)b'(q)/x'(q)$. Substituting this into the expression for P_y we get:

$$P_y = \mathbf{E}_q\left[(1-q)y'(q)b(q) + \frac{(1-q)y'(q)x(q)b'(q)}{x'(q)}\right] = \mathbf{E}_q\left[(1-q)y'(q)b(q) + Z_y(q)x(q)b'(q)\right]$$

where, as before, $Z_y(q) = \frac{(1-q)y'(q)}{x'(q)}$.

Integrating the second expression by parts, we get

$$\begin{aligned} \int_0^1 Z_y(q)x(q)b'(q)\,dq &= Z_y(q)x(q)b(q)|_0^1 - \int_0^1 (Z_y'(q)x(q) + Z_y(q)x'(q))b(q)\,dq \\ &= -\int_0^1 Z_y'(q)x(q)b(q)\,dq - \int_0^1 (1-q)y'(q)b(q)\,dq \end{aligned}$$

When we put this back in the expression for P_y two of the terms cancel, and we get the following lemma.

Lemma 8.1. *The per-agent revenue of an auction with allocation rule y can be written as a linear combination of the bids in a first-pay auction:*

$$P_y = \mathbf{E}_q\left[-x(q)Z_y'(q)b(q)\right]$$

where $Z_y(q) = (1-q)\frac{y'(q)}{x'(q)}$ and $x(q)$ are known precisely.

As in the case of the all-pay auction format, we can write the error in P_y as:

$$\begin{aligned} |\hat{P}_y - P_y| &= \mathbf{E}_q\left[\left|-x(q)Z_y'(q)(\widehat{b}(q) - b(q))\right|\right] \\ &\leq E\left[\frac{(\log(1+Z_y(q)))^\alpha}{Z_y(q)}|Z_y'(q)|\right]\sup_q\left|x(q)\frac{Z_y(q)}{(\log(1+Z_y(q)))^\alpha}(\hat{b}(q) - b(q))\right| \end{aligned}$$

With an appropriate choice of α we obtain the following theorem.

Theorem 8.2. *The expected absolute error in estimating the revenue of a position auction with allocation rule y using N samples from the bid distribution for a first-pay position auction with allocation rule x is bounded as below; Here n is the number of positions in the two position auctions.*

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_y - P_y|\right] \leq \frac{20}{\sqrt{N}}\sqrt{n\log n}\ \sup_q\{y'(q)\}\ \log\max\left\{\sup_{q:y'(q)\geq 1}\frac{x'(q)}{y'(q)}, \sup_q\frac{y'(q)}{x'(q)}\right\}$$
$$+ \frac{O(1)}{N}\sup_q\{x'(q)\}\ \sup_q\left\{\frac{y'(q)}{x'(q)}\right\}.$$

When y is the highest-k-bids-win allocation rule, the error improves to:

$$\mathbf{E}_{\hat{b}}\left[|\hat{P}_k - P_k|\right] \leq \frac{20}{\sqrt{N}}\ \sup_q\{x_k'(q)\}\ \log\max\left\{\sup_{q:x_k'(q)\geq 1}\frac{x'(q)}{x_k'(q)}, \sup_q\frac{x_k'(q)}{x'(q)}\right\}.$$

Because the error bounds in Theorem 8.2 are identical to those in Theorems 4.5 and 4.6, Lemma 7.5, Corollaries 6.1, 6.2, 7.4, 7.6, and Theorem 6.3 continue to hold when bids are drawn from a first-price auction.

9 Discussion and Conclusions

We conclude with some observations and discussion.

- Good inference requires careful design of the mechanism. Perfect inference and perfect optimality cannot be achieved together.
- We cannot achieve good accuracy in infering the revenue of an arbitrary mechanism, or in infering the entire revenue curve. In contrast, the multi-unit revenues P_k are special functions that depend linearly on the bid distribution (and not, for example, on bid density). This property enables them to be learned accurately.
- Rank based mechanisms achieve a good tradeoff between revenue optimality and quality of inference in position environments: (1) They are close to optimal regardless of the value distribution; (2) Optimizing over this class for revenue requires estimating only n parameters P_k that, by our observation above, are "easy" to estimate accurately; (3) Rank based mechanisms satisfy the necessary conditions on the slope of the allocation function that enable good inference.

References

Ai, C. and Chen, X. 2003. Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica 71,* 6, 1795–1843.

Alaei, S., Fu, H., Haghpanah, N., Hartline, J., and Malekian, A. 2012. Bayesian optimal auctions via multi- to single-agent reduction. In *ACM Conference on Electronic Commerce.*

Athey, S. and Ellison, G. 2011. Position auctions with consumer search. *The Quarterly Journal of Economics 126*, 1213–1270.

Athey, S. and Haile, P. A. 2007. Nonparametric approaches to auctions. *Handbook of econometrics 6*, 3847–3965.

Babaioff, M., Dughmi, S., Kleinberg, R., and Slivkins, A. 2012. Dynamic pricing with limited supply. In *Proceedings of the 13th ACM Conference on Electronic Commerce*. EC '12. ACM, New York, NY, USA, 74–91.

Babaioff, M., Sharma, Y., and Slivkins, A. 2009. Characterizing truthful multi-armed bandit mechanisms: Extended abstract. In *Proceedings of the 10th ACM Conference on Electronic Commerce*. EC '09. ACM, New York, NY, USA, 79–88.

Blundell, R. W. and Powell, J. L. 2004. Endogeneity in semiparametric binary response models. *The Review of Economic Studies 71,* 3, 655–679.

Brown, J. and Morgan, J. 2009. How much is a dollar worth? tipping versus equilibrium coexistence on competing online auction sites. *Journal of Political Economy 117,* 4, 668–700.

Bulow, J. and Klemperer, P. 1996. Auctions vs negotiations. *American Economic Review 86,* 1, 180–194.

Chawla, S. and Hartline, J. D. 2013. Auctions with unique equilibria. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*. EC '13. ACM, New York, NY, USA, 181–196.

Chen, X. and Pouzo, D. 2012. Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals. *Econometrica 80,* 1, 277–321.

Cheng, C. and Parzen, E. 1997. Unified estimators of smooth quantile and quantile density functions. *Journal of statistical planning and inference 59,* 2, 291–307.

Chernozhukov, V., Fernández-Val, I., and Melly, B. 2013. Inference on counterfactual distributions. *Econometrica 81,* 6, 2205–2268.

Csörgö, M. 1983. *Quantile processes with statistical applications*. SIAM.

Csorgo, M. and Revesz, P. 1978. Strong approximations of the quantile process. *The Annals of Statistics*, 882–894.

Devanur, N. R., Ha, B. Q., and Hartline, J. D. 2013. Prior-free auctions for budgeted agents. In *ACM Conference on Electronic Commerce*. 287–304.

Devanur, N. R. and Kakade, S. M. 2009. The price of truthfulness for pay-per-click auctions. In *Proceedings of the 10th ACM Conference on Electronic Commerce*. EC '09. ACM, New York, NY, USA, 99–106.

Doraszelski, U. and Pakes, A. 2007. A framework for applied dynamic analysis in io. *Handbook of industrial organization 3*, 1887–1966.

Edelman, B., Ostrovsky, M., and Schwarz, M. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *The American Economic Review 97,* 1, 242–259.

Fain, D. and Pedersen, J. 2006. Sponsored search: A brief history. *Bulletin of the American Society for Information Science and Technology 32,* 2, 12–13.

Gatti, N., Lazaric, A., and Trovò, F. 2012. A truthful learning mechanism for contextual multi-slot sponsored search auctions with externalities. In *Proceedings of the 13th ACM Conference on Electronic Commerce*. EC '12. ACM, New York, NY, USA, 605–622.

Guerre, E., Perrigne, I., and Vuong, Q. 2000. Optimal nonparametric estimation of first-price auctions. *Econometrica 68,* 3, 525–574.

Hardy, G., Littlewood, J., and Pólya, G. 1929. Some simple inequalities satisfied by convex functions. *Messenger of Math 58*, 145–152.

Hartline, J. and Yan, Q. 2011. Envy, truth, and profit. In *Proceedings of the 12th ACM conference on Electronic commerce*. ACM, 243–252.

Myerson, R. 1981. Optimal auction design. *Mathematics of Operations Research 6*, 58–73.

Newey, W. K. 1994. The asymptotic variance of semiparametric estimators. *Econometrica: Journal of the Econometric Society*, 1349–1382.

Ostrovsky, M. and Schwarz, M. 2011. Reserve prices in internet advertising auctions: A field experiment. In *Proceedings of the 12th ACM Conference on Electronic Commerce*. EC '11. ACM, New York, NY, USA, 59–60.

Paarsch, H. J. and Hong, H. 2006. *An introduction to the structural econometrics of auction data*. Vol. 1. The MIT Press.

Reiley, D. H. 2006. Field experiments on the effects of reserve prices in auctions: more magic on the internet. *The RAND Journal of Economics 37,* 1, 195–211.

Varian, H. 2007. Position auctions. *International Journal of Industrial Organization 25,* 6, 1163–1178.

A Finding the optimal iron by rank auction

Recall that iron by rank auctions are weighted sums of multi-unit auctions. Therefore, their revenue can be expressed as a weighted sum over the revenues P_k of k-unit auctions. We consider a position environment given by non-increasing weights $\boldsymbol{w} = (w_1, \ldots, w_n)$, with $w_0 = 0$, $w_1 = 1$, and $w_{n+1} = 0$. Define the cumulative position weights $\boldsymbol{W} = (W_1, \ldots, W_n)$ as $W_k = \sum_{j \leq i} w_j$.

Define the *multi-unit revenue curve* as the piece-wise constant function connecting the points $(0, P_0, \ldots, (n, P_n)$. This function may or may not be concave. Define the *ironed multi-unit revenue curve* as $\bar{\boldsymbol{P}} = (\bar{P}_1, \ldots, \bar{P}_n)$ the smallest concave function that upper bounds the multi-unit revenue curve. Define the multi-unit marginal revenues as $\boldsymbol{P'} = P'_1, \ldots, P'_n$ and $\bar{\boldsymbol{P}}' = \bar{P}'_1, \ldots, \bar{P}'_n$ as the left slope of the multi-unit and ironed multi-unit revenue curves, respectively. I.e., $P'_k = P_k - P_{k-1}$ and $\bar{P}'_k = \bar{P}_k - \bar{P}_{k-1}$.

We now see how the revenue of any position auction can be expressed in terms of the multi-unit revenue curves and marginal revenues.

$$\begin{aligned}\mathbf{E}[\text{revenue}] &= \sum_{k=0}^{n} P_k\, w'_k = \sum_{k=0}^{n} P'_k\, w_k \\ &\leq \sum_{k=0}^{n} \bar{P}_k\, w'_k = \sum_{k=0}^{n} \bar{P}'_k\, w_k.\end{aligned}$$

The first equality follows from viewing the position auction with weights $\boldsymbol{w}$ as a convex combination of multi-unit auctions (where its revenue is the convex combination of the multi-unit auction revenues). The second and final inequality follow from rearranging the sum (an equivalent manipulation to integration by parts). The inequality follows from the fact that $\bar{\boldsymbol{P}}$ is defined as the smallest concave function that upper bounds $\boldsymbol{P}$ and, therefore, satisfies $\bar{P}_k \geq P_k$ for all k. Of course the inequality is an equality if and only if $w'_k = 0$ for every k such that $\bar{P}'_k > P'_k$.

We now characterize the optimal ironing-by-rank position auction. Given a position auction weights $\boldsymbol{w}$ we would like the ironing-by-rank which produces $\bar{\boldsymbol{w}}$ (with cumulative weights satisfying $\boldsymbol{W} \geq \bar{\boldsymbol{W}}$) with optimal revenue. By the above discussion, revenue is accounted for by marginal revenues, and upper bounded by ironed marginal revenues. If we optimize for ironed marginal revenues and the condition for equality holds then this is the optimal revenue. Notice that ironed revenues are concave in k, so ironed marginal revenues are monotone (weakly) decreasing in k. The position weights are also monotone (weakly) decreasing. The assignment between ranks and positions that optimizes ironed marginal revenue is greedy with positions corresponding to ranks with negative ironed marginal revenue discarded. Tentatively assign the kth rank agent to slot k (discarding agents that correspond to discarded positions). This assignment indeed maximizes ironed marginal revenue for the given position weights but may not satisfy the condition for equality of revenue with ironed marginal revenue. To meet this condition with equality we can randomly permute (a.k.a., iron by rank) the positions that corresponds to intervals where the revenue curve is ironed. This does not change the surplus of ironed marginal revenue as the ironed marginal revenues on this interval are the same, and the resulting position weights $\bar{\boldsymbol{w}}$ satisfy the condition for equality of revenue and ironed marginal revenue.

B Proofs for Section 3

PROOF OF LEMMA 3.2. Consider estimation of the bid function using the sorted bids $b^{(1)} \geq b^{(2)} \geq \ldots \geq b^{(N)}$. Then the bid function is estimated as

$$\widehat{b}(q) = b^{([qN])},$$

where $[\cdot]$ is the floor integer. Let $G(\cdot)$ be the population cdf of the distribution of bids and $\widehat{G}(\cdot) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{\cdot \leq b_i\}$ be its empirical analog. Note that the derivative of the quantile function of the distribution of bids $b'(\cdot)$ is the inverse of the bid density. We now impose the restriction on the population distribution of bids ensuring its desirable smoothness.

Assumption 1. *For any $b_1, b_2 \in [0,\, 1]$ such that $|b_2 - b_1| < \delta$*

$$|G(b_2) - G(b_1) - b'(b^{-1}(b_1))(b_2 - b_1)| = o(\delta).$$

Then, we have by definition $\widehat{G}(\widehat{b}(q)) = q = G(b(q))$,where $G(\cdot)$ is the cdf of bids and $\widehat{G}(\cdot)$ is the empirical cdf.

Now we decompose $\widehat{G}(\widehat{b}(q)) - G(b(q))$ as

$$0 = \widehat{G}(\widehat{b}(q)) - G(b(q)) = \widehat{G}(\widehat{b}(q)) - G(\widehat{b}(q)) + G(\widehat{b}(q)) - G(b(q)). \tag{16}$$

We can bound the first term on the right as follows. By the Donsker theorem $\sqrt{N}(\widehat{G}(t) - G(t))$ converges to a tight mean zero stochastic process $\mathbb{G}(t)$ over t with covariance function such that $H(t,t) = G(t)(1 - G(t))$. Note that

$$\sup_t |H(t,t)| \leq \frac{1}{4}.$$

This means that

$$E\left[\left|\sqrt{N}\sup_t(\widehat{G}(t) - G(t))\right|\right] \leq \frac{1}{2}.$$

Next, we can relate the second term on the right hand side of (16) to the error in the bid estimator. Consider the following expansion:

$$G(\widehat{b}(q)) - G(b(q)) = g(b(q))(\widehat{b}(q) - b(q)) + o(|\widehat{b}(q) - b(q)|^2).$$

Combining this result together with the decomposition above and recalling that $b'(q) = 1/g(b(q))$, we write

$$\sqrt{N}(\widehat{b}(q) - b(q)) = -b'(q)\sqrt{N}(\widehat{G}(\widehat{b}(q)) - G(\widehat{b}(q))) + o_p(1).$$

Then we write

$$E\left[\left|\sqrt{N}\sup_q(\widehat{b}(q) - b(q))\right|\right] \leq \sup_q b'(q) E\left[\left|\sqrt{N}\sup_t(\widehat{G}(t) - G(t))\right|\right].$$

This means that the mean absolute error in bids is bounded by

$$\frac{\sup_q b'(q)}{2}\frac{1}{\sqrt{N}}.$$

Then, recalling that $v(q) \leq 1$ and $v(q) - b(q) \leq 1$, we can replace the upper bound on $b'(q)$ by $x'(q)$ for an all-pay auction and by $x'(q)/x(q)$ for the first-price auction. $\square$

C Proofs for Section 4

Proof of Lemma 4.2. Consider the function $A(q) = 1/Z_k(q) = x'(q)/(1-q)x'_k(q)$. $x'(q)$ is a weighted sum over $x'_j(q)$ for $j \in \{1, \cdots, n-1\}$. So, $A(q)$ is a weighted sum over terms $x'_j(q)/(1-q)x'_k(q)$. Let us look at these terms closely.

$$\frac{x'_j(q)}{(1-q)x'_k(q)} = \alpha_{k,j} q^{k-j}(1-q)^{j-k-1}$$

where $\alpha_{k,j}$ is independent of q. The functions $q^{k-j}(1-q)^{j-k-1}$ are convex. This implies that $A(q)$ which is a weighted sum of convex functions is also convex. Consequently, it has a unique minimum. Therefore, $Z_k(q) = 1/A(q)$ has a unique maximum. $\square$

Proof of Theorem 4.5. Recall that for $\alpha > 0$ we can write[4]

$$|\hat{P}_k - P_k| \leq E\left[\frac{(\log(1+Z_k(q)))^\alpha}{Z_k(q)}|Z'_k(q)|\right] \sup_q \left|\frac{Z_k(q)}{(\log(1+Z_k(q)))^\alpha}(\hat{b}(q)-b(q))\right|$$

We start by considering the first term. Lemma 4.2 shows that $Z'_k(\cdot)$ changes sign only once. Consider the region where the sign of $Z'_k(\cdot)$ is constant and make the change of variable $t = Z_k(q)$. Recall that $Z^*_k = \sup_q Z_k(q)$ and we note that $\inf_q Z_k(q) \geq 0$. Then we can evaluate the first term as

$$E\left[\frac{(\log(1+Z_k(q)))^\alpha}{Z_k(q)}|Z'_k(q)|\right] \leq 2\int_0^{Z^*_k} \frac{(\log(1+t))^\alpha}{t}\, dt$$

Note that for any $t > 0$, $\log(1+t) \leq t$. Thus,

$$\int_0^\delta \frac{(\log(1+t))^\alpha}{t}\, dt < \frac{\delta^\alpha}{\alpha}$$

Now split the integral into two pieces as

$$\int_0^{Z^*_k} \frac{(\log(1+t))^\alpha}{t}\, dt = \int_0^1 \frac{(\log(1+t))^\alpha}{t}\, dt + \int_1^{Z^*_k} \frac{(\log(1+t))^\alpha}{t}\, dt$$

We just proved that the first piece is at most $1/\alpha$. Now we upper bound the second piece and consider the integrand at $t \geq 1$. First, note that

$$(\log(1+t))^\alpha = \left(\log t + \log(1+\frac{1}{t})\right)^\alpha \leq \left(\log t + \frac{1}{t}\right)^\alpha \leq (\log t + 1)^\alpha.$$

Thus, the integral behaves as

$$\int_1^{Z^*_k} \frac{(\log(1+t))^\alpha}{t}\, dt \leq \int_1^{Z^*_k} \frac{(\log(t)+1)^\alpha}{t}\, dt = \frac{1}{1+\alpha}(\log Z^*_k + 1)^{1+\alpha}.$$

Thus, we just showed that

$$E\left[\frac{(\log(1+Z_k(q)))^\alpha}{Z_k(q)}|Z'_k(q)|\right] \leq \frac{2}{\alpha} + \frac{2}{1+\alpha}(\log Z^*_k + 1)^{1+\alpha},$$

which is at most $2(1+e)/\alpha$ for $\alpha < 1/\log Z^*_k$.

Now consider the term

$$\sup_q \left|\frac{Z_k(q)}{(\log(1+Z_k(q)))^\alpha}(\hat{b}(q)-b(q))\right|$$

Note that $\log(1+t) \geq \min\{1,t\}/2$. So the first term can be bounded from above as

$$\frac{Z_k(q)}{(\log(1+Z_k(q)))^\alpha} \leq 2^\alpha \max\left\{Z_k(q),\, (Z_k(q))^{1-\alpha}\right\}$$

The second term behaves as

$$(\hat{b}(q) - b(q)) = -x'(q)v(q)(\hat{G}(b(q)) - G(b(q)))$$

[4] In this entire proof the logarithms are natural logarithms.

where G and $\hat{G}$ are the real and empirical bid distributions, respectively. We recall that $\mathbf{E}\left[\sup_q \left|\hat{G}(b(q)) - G(b(q))\right|\right]$ $\frac{1}{2\sqrt{N}}$. Thus

$$\mathbf{E}\left[\sup_q \left|\frac{Z_k(q)}{(\log(1+Z_k(q)))^\alpha}(\hat{b}(q)-b(q))\right|\right] \le 2^\alpha \sup_q \left(\max\left\{x_k'(q),\, (x_k'(q))^{1-\alpha}(x'(q))^\alpha\right\}\right)\frac{1}{2\sqrt{N}}$$

$$\le 2^\alpha \sup_q \left(x_k'(q)\right)\left(\underbrace{\max\left(1, \sup_{q: x_k'(q)\ge 1}\frac{x'(q)}{x_k'(q)}\right)}_{=:A}\right)^\alpha \frac{1}{2\sqrt{N}}$$

Now we combine the two evaluations together and pick $\alpha = \min\{1/\log A, 1/\log Z_k^*\}$, with A defined as above, to obtain

$$\mathbf{E}\left[|\hat{P}_k - P_k|\right] \le \frac{2(1+e)}{\alpha}2^\alpha A^\alpha \frac{1}{2\sqrt{N}}\sup_q\left(x_k'(q)\right)$$

$$\le \frac{20}{\sqrt{N}}\ \sup_q\{x_k'(q)\}\ \max\left\{\log A, \log\sup_q\left\{\frac{x_k'(q)}{x'(q)}\right\}\right\}$$

□

Proof of Lemma 4.7. We can write the function $\tilde{b}(i/N)$ as

$$\tilde{b}\left(\frac{i}{N}\right) = \frac{N!}{(i-1)!(N-i)!}\int G(t)^{i-1}(1-G(t))^{N-i}g(t)t\,dt$$

$$= \frac{N!}{(i-1)!(N-i)!}\int t^{i-1}(1-t)^{N-i}b(t)\,dt$$

Note that

$$\frac{N!}{(i-1)!(N-i)!}t^{i-1}(1-t)^{N-i}$$

is the density of the beta-distribution with parameters $\alpha = i$ and $\beta = N-i+1$. Denote this density $f(t;\alpha,\beta)$. Then we can write

$$\tilde{b}\left(\frac{i}{N}\right) = \int_0^1 b(t)f(t;\alpha,\beta)\,dt.$$

Now let $q \in [i/N,\, (i+1)/N]$, and consider an expansion of $b(t)$ at q such that

$$b(t) = b(q) + b'(q)(t-q) + O((t-q)^2)$$

Now we substitute this expansion into the formula for $\tilde{b}(\cdot)$ above to get

$$\tilde{b}\left(\frac{i}{N}\right) = b(q) + b'(q)\int_0^1 (t-q)f(t;\alpha,\beta)\,dt + O(\int_0^1 (t-q)^2 f(t;\alpha,\beta)\,dt)$$

The mean of the beta distribution is $\alpha/(\alpha+\beta)$ and the variance is $\alpha\beta/((\alpha+\beta)^2(\alpha+\beta+1))$. This means that

$$\tilde{b}\left(\frac{i}{N}\right) - b(q) = b'(q)\left(\frac{i}{N+1} - q\right) + O\left(\frac{1}{N^2}\right).$$

Thus

$$\sup_{q\in[i/N,(i+1)/N]}\left|\tilde{b}\left(\frac{i}{N}\right)-b(q)\right|\le \sup_q b'(q)\frac{2}{N}+O\left(\frac{1}{N^2}\right).$$

Therefore, the expectation $\left|\hat{P}_y-\mathbf{E}\left[\hat{P}_y\right]\right|$ is at most $O(1)/N\ \sup_q\{x'(q)\}\ \sup_q Z_y(q)$. $\square$

Proof of Lemma 4.8. Fix I and ℓ, and note that the function $\tilde{b}$ is fixed (that is, it does not depend on the empirical bid sample). Then, the sum $\mathbf{T}_{i_\ell,i_{\ell+1}}$ depends only on the empirical bid values $\widehat{b}(q)$ for quantiles in the interval $[i_\ell/N, i_{\ell+1}/N)$. By the definition of I, we know that the smallest i_ℓ bids in the sample are all smaller than $\tilde{b}((i_\ell-1)/N)\le\tilde{b}(i_\ell/N)$, and the largest $N-i_{\ell+1}$ bids in the sample are all larger than $\tilde{b}(i_{\ell+1}/N)\ge\tilde{b}((i_{\ell+1}-1)/N)$. On the other hand, the empirical bids $\widehat{b}(q)$ for $q\in[i_\ell/N, i_{\ell+1}/N)$ lie within $[\tilde{b}(i_\ell/N),\tilde{b}((i_{\ell+1}-1)/N)]$. Therefore, conditioned on i_ℓ and $i_{\ell+1}$, the latter set of empirical bids is independent of the former set of empirical bids. $\square$

Proof of Lemma 4.9 and Theorem 4.6. We will use Chernoff-Hoeffding bounds to bound the expectation of $\mathbf{T}_{0,N}$ over the bid sample, conditioned on I and $\Delta=\sup_{q\in[0,1]}|\hat{G}(b(q))-G(b(q))|$. We first note that $\mathbf{T}_{0,N}$ has mean zero because for any integer $i\in[0,N]$, $\mathbf{E}_{\text{samples}}\left[\widehat{b}(i/N)\right]=\tilde{b}(i/N)$.

Next we note that the $\mathbf{T}_{i,j}$'s are bounded random variables. Specifically, let Q be an interval of quantiles over which the difference $\widehat{b}(q)-\tilde{b}(q)$ does not change sign. Then, following the proof of Theorem 4.5, we can bound

$$\begin{aligned}|\mathbf{T}_Q| &= \left|\int_Q Z_y'(q)(\widehat{b}(q)-\tilde{b}(q))\,\mathrm{d}q\right|\\ &\le 20\Delta\ \underbrace{\sup_q\{y'(q)\}\ \log\max\left\{\sup_{q:y'(q)\ge 1}\frac{x'(q)}{y'(q)},\sup_q\frac{y'(q)}{x'(q)}\right\}}_{=:C}.\end{aligned}$$

Likewise, over an interval Q where Z_y' does not change sign, we again get $|\mathbf{T}_Q|\le 20\Delta C$ with C defined as above. Moreover, for an interval Q over which Z_y' changes sign at most t times, we have

$$\int_Q |Z_y'(q)(\widehat{b}(q)-\tilde{b}(q))|\,\mathrm{d}q\le t\cdot 20\Delta C.$$

Finally, noting that Z_y is a weighted sum over the n functions Z_k defined for the k-unit auctions, and that by Lemma 4.2 each Z_k has a unique maximum, we note that Z_y' changes sign at most $2n$ times.

We now apply Chernoff-Hoeffding bounds to bound the probability that the sum $\sum_{\ell=0}^{\ell=k-1}\mathbf{T}_{i_\ell,i_{\ell+1}}$ exceeds some constant a. With τ_ℓ denoting the upper bound on $|\mathbf{T}_{i_\ell,i_{\ell+1}}|$, this probability is at most

$$\exp\left(-\frac{a^2}{\sum_\ell \tau_\ell^2}\right).$$

By our observations above, for all ℓ, $\tau_\ell\le 40\Delta C$, and $\sum_\ell\tau_\ell\le\int_0^1|Z_y'(q)(\widehat{b}(q)-\tilde{b}(q))|\,\mathrm{d}q\le 40n\Delta C$. Therefore, $\sum_\ell\tau_\ell^2\le n(40\Delta C)^2$. We can now choose $a=\sqrt{n\log n}\,40\Delta C$ to make the above probability at most $1/n$.

Putting everything together, we get that conditioned on I and Δ, the expected value of $|\mathbf{T}_{0,N}|$ over the bid sample is at most $a+1/n\cdot 40n\Delta C=O(1)\sqrt{n\log n}\Delta C$. Since this bound is independent of I, the same bound holds when we remove the conditioning on I. The theorem now follows by plugging in the expected value of Δ from Lemma 3.2. $\square$