Sample Complexity for Non-truthful Mechanisms

Jason Hartline* and Samuel Taggart†

December 19, 2018

Abstract

This paper considers sample complexity of non-truthful mechanisms. We show that a polynomial number of samples from the bid distribution are sufficient to identify a mechanism within a restricted family of non-truthful mechanisms that approximates the overall optimal mechanism. The restricted family of mechanisms that we identify has the desirable property that equilibria in various standard payment formats – first-price, second-price, all-pay – are simple to determine and all yield the same outcome. Our mechanism and methods are new to the literature on sample complexity.

*Northwestern University, hartline@eecs.northwestern.edu

†Oberlin College, Sam.Taggart@oberlin.edu

1 Introduction

This paper studies the sample complexity of non-truthful mechanisms. We identify a family of mechanisms that have good sample complexity under standard non-truthful and truthful variants and the objectives of revenue and welfare.[1] Specifically, from a polynomial number – in the desired precision of the approximation and number of agents – of Bayes-Nash equilibrium bids of any mechanism in the family, we identify a mechanism that, with a polynomial number of sampled bids from its equilibrium, approximates the performance of the Bayesian optimal mechanism.

The study of non-truthful mechanisms is necessitated by both practical concerns and theoretical limitations of truthful mechanisms. On the practical side, in some common applications of mechanism design the outcome is a contract, e.g., government procurement auctions, variable commission mechanisms of third-party listing agencies like Booking.com, and ad exchanges. For these applications, the theory of winner-pays-bid mechanisms is most appropriate. For games of effort – like crowdsourcing contests (e.g., Chawla et al., 2015), forecasting (e.g., Osband, 1989), and peer prediction (Dasgupta and Ghosh, 2013, e.g.,) – the theory of all-pay mechanisms is most appropriate. The essay "The lovely but lonely Vickrey auction" by Ausubel and Milgrom (2006) discusses a number of other pragmatic concerns and describes why truthful mechanisms are rarely seen in practice. Akbarpour and Li (2018), for example, gave a formal framework for one of these concerns, proving that winner-pays-bid mechanisms are "credible," in the sense that the agents do not need to trust the designer, while truthful mechanisms are not and, e.g., buyers in a truthful auction may worry about a seller shill.

From the perspective of the theory of computation with strategic agents, a key consideration is how the computation is divvied up between the agents and the designer of the system. From this point of view, truthful mechanisms require that the designer does all the computation, while in general it might be beneficial for agents to do some of the computation. This distinction matters when there are information asymmetries between the agents and the designer. For example, Feng and Hartline (2018) proved that non-truthful mechanisms can be strictly better than truthful mechanisms when robustness to distributional assumptions is desired. Specifically, the so-called "revelation principle" fails when looking for simple and robust mechanisms. This failure necessitates a revisitation of some of the standard questions in mechanism design with a focus on non-truthful mechanisms.

The generalization to non-truthful mechanisms of the existing literature on sample complexity in mechanism design adds two fundamental complications. First, for non-truthful sample complexity it should be assumed that there is sample access to equilibrium bids not the values of the agents. This issue can be partially addressed with methods from econometrics that allow values to be inferred from bids; but this standard approach would impact sample complexity bounds in non-trivial ways (and come with assumptions that bid distributions are well behaved). Second, and more fundamentally, the literature has not yet shown how to design non-truthful mechanisms that have near-optimal equilibria with only estimates of the value distribution. As this indirect approach seems unwieldy (see Appendix A), we instead identify a family of mechanisms for which we can directly analyze the sample complexity under any of the standard variants: truthful payments, winner-pays-bid payments, and all-pay payments. Indeed, the family of mechanisms we identify and the main theorem of our analysis are unlike

[1]For example, the first-price, second-price, and all-pay auctions are standard variants of the highest-bid-wins auction.

the existing approaches for sample complexity in mechanism design.

Our theory of sample complexity for non-truthful mechanisms carefully delineates two aspects of sample complexity: where samples come from and how samples are used. We assume samples are from the bid distribution of non-truthful mechanisms. A non-issue for truthful sample complexity, for non-truthful mechanisms it is important which mechanism the bids are from as the Bayes-Nash equilibrium distributions of bids of distinct non-truthful mechanisms are generally distinct. Samples are generally used in two different ways: samples can be used at design time to select the mechanism to be run, in which case they are observed by the bidders, and samples can be used at run time when the mechanism executes, in which case they are viewed as a source of randomness as agents are bidding. In analyses of mechanisms from samples, incentives and performance guarantees should hold with high probability in design-time samples but can hold in expectation in run-time samples. Surveying the literature on truthful sample complexity: the learned finite support auctions of Elkind (2007) use design-time samples; the single-sample mechanism of Dhangwatnotai et al. (2010) uses run-time samples; Cole and Roughgarden (2014) and its followup papers use design-time samples; Fu et al. (2014) show that for correlated value distributions run-time sample complexity can be much lower than design-time sample complexity. In the related literature on reducing Bayesian-truthful mechanism design to Bayesian algorithm design Hartline and Lucier (2010, 2015), Hartline et al. (2011, 2015b), and Bei and Huang (2011) show that design-time samples are sufficient for ϵ-Bayesian truthfulness; the exact Bayesian-truthful reductions in these papers and Dughmi et al. (2017) use run-time samples. Except for the symmetric environments studied by Dhangwatnotai et al. (2010), the mechanisms in the papers above are not known to have non-truthful variants that satisfy similar bounds.

We consider the problem of designing good mechanisms from samples in general single-dimensional agent environments with independently distributed values in Bayes-Nash equilibrium where a general set system governs the subsets of agents that can be simultaneously served. Truthful mechanisms for this environment are well understood. The Vickrey-Clarke-Groves mechanism maximizes welfare; and a straightforward generalization of Myerson (1981) gives the truthful mechanism that maximizes expected revenue (e.g., Hartline, 2013). The sample complexity of truthful mechanisms of this setting was largely resolved by Devanur et al. (2016) and Gonczarowski and Nisan (2017).

The following statement of the problem of non-truthful sample complexity is the most straightforward generalization of the problem of truthful sample complexity. The problem of *non-truthful sample complexity* is to identify in a parameterized family of mechanisms and polynomials p_{design} and p_{run} such that with n-agent environments and desired loss ϵ:

C1. With $m_{\text{design}} = p_{\text{design}}(n, 1/\epsilon)$ design-time samples of profiles of Bayes-Nash equilibrium bids in any mechanism in the family, parameters of the designed mechanism can be selected.

C2. With $m_{\text{run}} = p_{\text{run}}(n, 1/\epsilon)$ run-time samples of profiles of Bayes-Nash equilibrium bids in the selected mechanism the selected mechanism can be run.

C3. With probability in the m_{design} design-time samples of at least $1 - \epsilon$, the expected performance, in agents values and the m_{run} run-time samples of the selected mechanism, is at most ϵ less than that of the Bayesian optimal mechanism.[2]

[2]Multiplicative versions of C3 are also interesting, in the multiplicative version a $(1 + \epsilon)$ approximation is required.

Note that run-time samples, as discussed above in the discussion of Bayesian reductions, seem to be inherent in obtaining excact incentive properties like Bayes-Nash equilibrium from samples.

The following story fits the above problem of non-truthful sample complexity and describes how design-time and run-time samples can be obtained. This story is implicit in previous papers in the literature. There is a population of agents and the designer aims to run a mechanism on agents drawn from this population. This population model of Bayes-Nash equilibrium is standard, see e.g., Hartline et al. (2015a). A mechanism is sought that performs well on a fresh draw of agents from the population. The mechanism has access to independent samples from the population, e.g., from historical bids in the same or different mechanisms.

Recall that for non-truthful sample complexity (and contrasting with truthful sample complexity), the mechanism from which bid samples are obtained plays a critical role. Our non-truthful mechanism designer fixes a large parameterized family of distributions and has independently drawn profiles of historical bids in one mechanism in the family. The designer uses these historical bid profiles as design-time samples to select new parameters of the mechanism. The agents adapt to the new equilibrium in the new mechanism. The designer collects historical samples in the new mechanism and uses them as run-time samples in its execution.[3]

In practical applications that resemble the story described above, like ad auctions, agents bid in advance of the auction, and it is possible to batch the bid collection for many individual executions of the mechanism together. For these batched executions, the run-time samples can be from the other bid profiles that are collected within the same batch.

Approach and Results. A key primative in our result is the i.i.d. rank-based position auction, a model popularized by ad auctions on search engines, cf. Jansen and Mullen (2008). In such an auction, agents are assigned to positions with higher positions having higher allocation probabilities. In an i.i.d. position auction the agents' values are drawn from the same distribution. Equilibria in these auctions are unique and efficient (Chawla and Hartline, 2013). One way to view our result is as reducing sample complexity questions in general single-dimensional environments with non-identically distributed agents to those in i.i.d. position auctions.

We identify a parameterized family of mechanisms that use run-time samples in their own bid distribution. These mechanisms effectively replace competition between agents from distinct distributions with competition between agents with identical distributions (from the run-time samples). We show that, to any agent, the equilibrium in this mechanism looks like an *i.i.d. position auction* which has a simple, unique, and efficient equilibrium. We show that mechanisms in this family are approximately optimal, i.e., the representation error is small, and we reduce the problem of analyzing the generalization error to the problem of estimating appropriate order statistics of design-time samples from bids in any i.i.d. position auctions with the same agent distributions. For i.i.d. position auctions with standard payment variants like winner-pays-bid and all-pay, Chawla et al. (2017) solved this inference problem. For the truthful variant, we give a straightforward solution and analysis.

Our framework admits several new sample complexity results for general single-dimensional environments and independent but non-identically distributed agents. With n agents, a polynomial in n and $1/\epsilon$ design-time samples are sufficient to identify a mechanism that, with polynomial run-time samples, approximates optimal mechanism to within precision ϵ in the following

[3] It is beyond the scope of this paper to model the adaptive process by which the agents might learn the new equilibrium and the designer might obtain the requisite run-time samples. Instead as is common on the literature on Bayes-Nash mechanism design, we consider the steady state equilibrium.

environments:

- (non-truthful) winner-pays-bid and all-pay mechanisms, additive welfare approximation, and bounded value distributions;
- (non-truthful) winner-pays-bid and all-pay mechanisms, additive revenue approximation, and bounded and regular value distributions; and
- truthful mechanisms, multiplicative revenue approximation, and (unbounded) regular value distributions.[4]

Regular distributions are ones that satisfy a natural convexity property; details are given in Section 2.

2 Preliminaries

This work considers the *single-dimensional independent private value* model of mechanism design. We describe this model in *quantile* space where the geometry of approximation mechanisms is more transparent (cf. Hartline, 2013). There are n agents drawn independently and uniformly at random from n populations. Agents are distinguished by their quantile with respect to their own population. The *quantile* q_i of agent i is the measure of population i with higher values. The *value function* v_i of population i maps agent i's quantile to her value as $v_i(q_i)$ and, with a uniformly drawn quantile, induces a value distribution. The profile of agent value functions and quantiles are denoted by $\boldsymbol{v} = (v_1, \ldots, v_n)$ and $\boldsymbol{q} = (q_1, \ldots, q_n)$, respectively.

An *allocation* is $\boldsymbol{x} = (x_1, \ldots, x_n)$ where $x_i \in \{0, 1\}$ is an indecator for agent i being served. The space of feasible allotions is given by $\mathcal{X} \subset \{0, 1\}^n$. (Notably, we do not require that $\mathcal{X}$ be downward closed.) Agent i can be assigned a non-negative payment denoted p_i and her utility is linear in allocation and payment as $v_i(q_i)\, x_i - p_i$.

A mechanism takes as input a profile of bids $\boldsymbol{b} = (b_1, \ldots, b_n)$ and outputs a feasible allocation $\boldsymbol{x} \in \mathcal{X}$ and agent payments $\boldsymbol{p}$. A mechanism consists of an *allocation algorithm* $\tilde{\boldsymbol{x}}(\boldsymbol{b})$, which maps bid profiles to a feasible allocation, and a payment rule $\tilde{\boldsymbol{p}}(\boldsymbol{b})$, which maps bid profiles to a non-negative payment for each agent. A standard allocation algorithm is *highest-bids-win* which is defined by $\tilde{\boldsymbol{x}}(\boldsymbol{b}) \in \operatorname{argmax}_{\boldsymbol{x} \in \mathcal{X}} \sum_i b_i\, x_i$. We consider payment rules defined directly from the allocation algorithm according to standard payment formats. The *winner-pays-bid* format has payment rule $\tilde{p}_i(\boldsymbol{b}) = b_i\, \tilde{x}_i(\boldsymbol{b})$, and the *all-pay* format has payment rule $\tilde{p}(\boldsymbol{b}) = b_i$. Mechanisms with these payment formats do not have truth-telling as an equilibrium. The truthful payment format is defined according to the payment identity (below, Theorem 1) and can be implemented as an integral or with any of a number of unbiased estimators with expectation equal to the integral (see, e.g., Hartline and Lucier, 2015).

We analyze non-truthful mechanisms in Bayes-Nash equilibrium (BNE): each agent's report to the mechanism is a best response to the distribution of bids induced by other agents' strategies. The strategy of agent i is denoted s_i and maps the agent's quantile to a bid and, with a uniformly drawn quantile, induces a bid distribution. The mechanism $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$, the agents' strategies $\boldsymbol{s}$, and the distribution over quantiles induce interim allocation and payment rules. Agent i's *interim allocation rule* is $x_i(q_i) = \mathbf{E}_{\boldsymbol{q}_{-i}}[\tilde{x}_i(\boldsymbol{s}(\boldsymbol{q}))]$ and *interim payment rule* $p_i(q_i) = \mathbf{E}_{\boldsymbol{q}_{-i}}[\tilde{\boldsymbol{p}}(\boldsymbol{s}(\boldsymbol{q}))]$. Myerson (1981) characterized the interim allocation and payment rules that arise in BNE when agents' values are independently distributed. For truthful mechanisms $\boldsymbol{x} = \tilde{\boldsymbol{x}}$ and $\boldsymbol{p} = \tilde{\boldsymbol{p}}$.

[4] These results are more general than the results of Devanur et al. (2016), who require downward closure.

Theorem 1 (Myerson, 1981)**.** *For independently distributed agents, interim allocation and payment rules are induced by a Bayes-Nash equilibrium with onto strategies if and only if for each agent i,*

1. *(monotonicity) allocation rule $x_i(q)$ is monotone non-increasing in q_i, and*
2. *(payment identity) payment rule $p_i(q_i)$ satisfies $p_i(q_i) = v_i(q_i)\, x_i(v_i) + \int_{q_i}^{1} x_i(r)\, v_i'(r)\, dr + p_i(1)$.*

This paper studies the objectives of welfare and revenue. The *welfare* of a mechanism is $\mathbf{E}[\sum_i v_i(q)\, x_i(q_i)]$. The optimal mechanism for welfare allocates the value-maximizing feasible set, which is monotone and therefore implementable with payments via Theorem 1. The *revenue* of a mechanism is given by $\mathbf{E}[\sum_i p_i(q_i)]$. The revenue of a mechanism is easily analized in quantiles space in terms of revenue curves and marginal revenue as follows.

Lemma 2 (Myerson, 1981; Bulow and Roberts, 1989)**.** *In BNE, the expected payment of an agent i satisfies*

$$\mathbf{E}_{q_i}[p_i(q_i)] = \mathbf{E}_{q_i}[-x_i'(q_i)\, R_i(q_i)] + R_i(1)\, x_i(1) = \mathbf{E}_{q_i}[R_i'(q_i)\, x_i(q_i)] + R_i(0)\, x_i(0)$$

where the revenue curve $R_i(q_i) = v_i(q_i)\, q_i$ *gives the revenue from posting price q_i and the* marginal revenue $R_i'(q_i) = v_i(q_i) + v_i'(q_i) q_i$ *is its derivative. (Note that the derivatitives of the allocation rule $x_i'(\cdot)$ and value function $v_i'(\cdot)$ are non-positive.)*

The first equality follows from revenue equivalence and noting that the allocation rule x_i is equivalent to offering a randomized posted price with price distributed according to the density function $-x_i'(\cdot)$ with a pointmass of $x_i(1)$ at price $v_i(1)$. The second equality follows from integration by parts. The optimal mechanism can be easily identified from the second equality as maximizing the surplus of marginal revenue. The value distributions are called *regular* when the revenue curves are concave, equivalently, the marginal revenues are monotonically non-increasing.

In many environments of interest, the additive terms $R_i(0)\, x_i(0)$ and $R_i(1)\, x_i(1)$ are zero. For example, when the strongest agent $q_i = 0$ in the population has finite value $v_i(0)$, then the revenue when we post price $v_i(0)$ is $R_i(0) = 0$ as only a zero measure of the population will buy at such a price. For example, when the weakest agent in the population $q_i = 1$ has value $v_i(1) = 0$ then $R_i(1) = 0$ as the revenue from posting price 0 is zero.

Position Auctions *I.i.d. rank-by-bid position auctions* play a fundamental role in our analysis. In i.i.d. environments the agents' value functions are identical $v_i = v_j$ for all agents i and j. An n-agent position auction is defined by n position weights $w_1 \geq \ldots \geq w_n \in [0, 1]$ and an outcome is an assignment of agents to positions. If agent i is assigned to position j her allocation is $x_i = 1$ with probability w_j and zero otherwise, i.e., $\mathbf{E}[x_i \mid \text{agent } i \text{ is assigned slot } j] = w_j$. The rank-by-bid allocation algorithm assigns agents to positions assortatively by bid. The following theorem shows that Bayes-Nash equilibria in rank-by-bid position auctions are straightforward.

Theorem 3 (Chawla and Hartline, 2013)**.** *In i.i.d. position environments, the rank-by-bid winner-pays-bid and all-pay auctions have a unique and welfare-maximizing Bayes-Nash equilibrium (in which agents are assigned to positions in order of their true values), i.e., $s_i(\cdot) = s_j(\cdot)$ for all agents i and j.*

3 Surrogate-Ranking Mechanisms

In this section, we describe the parameterized family of mechanisms for which demonstrate a polynomial sample complexity. A mechanisms in this family has run-time sample access to the equilibrium bid distribution of each agent. An agent's bid can be compared to these samples to estimate the agent's strength relative to their value distribution. The mechanism then allocates solely on the basis of the agents' ranks. The choice of parameters will determine the exact mapping between ranks and allocations. This approach can be pared with any of the standard payment formats: winner-pays-bid, all-pay, or truthful.

Definition 4. *A* surrogate-ranking mechanism *(SRM) is parameterized by nT* surrogate values *$\mathbf{\Psi}$, with $\Psi_i = \{\psi_i^1 \geq \ldots \geq \psi_i^T\}$ for each agent i. The input to the mechanism is a profile of bids.*

1. *A surrogate value is calculated for each agent i as:*
   (a) *draw $T-1$ run-time samples from the agent's bid distribution,*
   (b) *calculate the rank r_i of the agent's bid relative to these samples,*
   (c) *select the agent's surrogate value $\psi_i = \psi_i^{r_i}$ according to the agent's sample rank.*
2. *For space $\mathcal{X}$ of feasible allocations, the algorithm allocates to maximize the* surrogate surplus $\text{argmax}_{\boldsymbol{x}\in\mathcal{X}} \sum_i \psi_i \, x_i$.
3. *Payments are assigned according to any standard payment format, e.g., winner-pays-bid, all-pay, or truthful.*

In the paper we will focus on surrogate ranking mechanisms as defined above where the allocation is chosen to maximize the surrogate surplus, i.e., $\sum_i \psi_i \, x_i$. Our methods extend in a straightforward manner to settings where computing such an allocation is intractable. For approximation algorithms where surrogate allocations is monotone in surrogate values, all analyses in this paper hold with an additional multiplicative performance loss equal to the approximation factor of the algorithm. Non-monotone algorithms can be made monotone via the methods of Hartline and Lucier (2015) or Hartline et al. (2015b).

Subsequently in Section 5, we will show how to identify good surrogate values from design-time samples. The remainder of this section is devoted to characterizing equilibria in surrogate ranking mechanisms.

3.1 Equilibria of Surrogate-Ranking Mechanisms

We now analyze the equilibrium of winner-pays-bid and all-pay surrogate-ranking mechanisms.[5] To do so, we first give a natural generalization of Bayes-Nash equilibrium to mechnanisms with run-time samples from its own bid distribution.

Definition 5. *A* stationary equilibrium (with samples) *in a mechanism (with samples) is a strategy profile $\boldsymbol{s}$ where the strategy of each agent is in best response to distribution of bids induced by the strategies in the mechanism with sample access to the same bid distributions.*

[5] In Section 5, we also consider truthful SRMs. Equilibrium analysis for truthful SRMs is trivial.

We will characterize the stationary equilibra of surrogate ranking mechanisms as being equivalent to the profile Bayes-Nash equilibria where each agent's strategy is the same as the strategy in unique equilibrium of an i.i.d. position auction with weights induced by the surrogate ranking mechanism. Specifically, rather than competing with other agents in the mechanism, an agent i's bid competes with other bids from her bid distribution which gives an outcome which is equivalent to the equilibrium of an i.i.d. position auction with agents with values drawn only from population i.

We begin by analyzing the distribution of assigned surrogate values in a stationary equilibrium. Recall, each agent's strategy s_i, on a uniform quantile, induces a distribution over bids. Notice that, in the surrogate-ranking mechanism with sample access to this bid distribution, the surrogate value assigned to i will be uniformly distributed from the set Ψ_i of i's surrogate values.

Lemma 6. *In any stationary equilibrium of a surrogate ranking mechanism, and any agent i and surrogate value $\psi_i^j \in \Psi_i$, the ex ante probability agent i is assigned ψ_i^j is $1/T$.*

Lemma 6 implies that in a stationary equilibrium, the probability of allocation associated with a particular surrogate value is fully determined by the other agents' sets of surrogate values, and not by the form of the equilibrium bidding strategies, or even by the agents' value distributions. This characterization of outcomes can be formalized as follows.

Definition 7. *For each agent i, let $\Psi_i = \{\psi_i^1 \geq \ldots \geq \psi_i^T\}$ be agent i's set of surrogate values, and let $\hat{\boldsymbol{x}}$ denote the surrogate surplus maximizing allocation rule $\hat{\boldsymbol{x}}(\boldsymbol{\psi}) = \max_{\boldsymbol{x} \in \mathcal{X}} \sum_i \psi_i \, x_i$. The* characteristic weights *$W_i = \{w_i^1 \geq \ldots \geq w_i^T\}$ for agent i are defined by calculating the allocation probability associated with each surrogate when the surrogates of other populations are drawn uniformly at random, i.e., $w_i^j = \mathbf{E}[\hat{x}_i(\psi_i^j, \boldsymbol{\psi}_{-i})]$ for each surrogate j and uniform random $\boldsymbol{\psi}_{-i}$ from $\boldsymbol{\Psi}_{-i}$.*

We now show that from each agent's perspective, stationary equilibria in surrogate-ranking mechanisms looks like an position auction among agents with the same value function. These agents compete for the characteristic weights of their population's surrogate values. With pay-your-bid or all-pay payment semantics, they therefore inherit the equilibrium of rank-based position auctions, which is shown by Chawla and Hartline (2013) to be efficient (i.e., to rank agents by values) and unique.

Theorem 8. *For any profile of value functions $\boldsymbol{v}$, surrogate values $\boldsymbol{\Psi}$, and characteristic weights $\boldsymbol{W}$; the unique stationary equilibrium of the winner-pays-bid (resp. all-pay) SRM is given by each agent i biding according to the unique and efficient BNE s_i of the i.i.d. winner-pays-bid (resp. all-pay) position auction with position weights W_i and value function v_i.*

Proof. Assume an arbitrary stationary equilibrium and consider an agent i. By Lemma 6, the stationary equilibrium induces a uniform distribution over each other agent's assigned surrogate values. It follows that if agent i is assigned surrogate value ψ_i^j, then they are allocated with probability w_i^j. Moreover, since the surplus-maximizing allocation algorithm is monotone, characteristic weights for agent i are monotone as well. Hence, placing the jth highest bid among the run-time samples will cause i to be assigned the jth highest characteristic weight. Thus, agent i faces the same bidding problem as if they played in the equilibrium of the i.i.d. position auction with position weights $w_i^1, \ldots, w_i^T$ and value function v_i. Thus, agent i bids according to the BNE of this i.i.d. position auction. This BNE is efficient, i.e., bids are in the same order as values, and unique.

Uniqueness of the stationary equilibrium follows by uniqueness of characteristic weights under any stationary equilibrium (Definition 7), which are determined only by the set of surrogate values $\mathbf{\Psi}$, and the uniqueness of Bayes-Nash equilibrium in i.i.d. position auctions. □

3.2 Equivalence of Surrogate Ranking Mechanisms

Surrogate ranking mechanisms are equivalent for revenue and welfare, irrespective of their payment format. It is helpful to relate this equivalance to the famous revenue equivalence result of Myerson (1981). In the latter, two mechanisms with the same equilibrium outcome (and with the same expected payment of the agent with the lowest value in the support of the distribution, usually zero) have the same expected revenue. For instance, with i.i.d. distirbutions the single-item first-price and second-price auctions have equilibrium outcome, i.e., the highest valued agent wins, and thus, by revenue equivalence, the same expected revenue. With non-identical distributions, these auctions do not have the same equilibrium outcome and, thus, do not generally have the same expected revenue. Our equivalence result, in contrast, holds for surrogate-ranking mechanisms in asymmetric environments (distributions and feasibility constraints).

Theorem 9. *For any fixed surrogate values and value functions, the expected welfare (resp. revenue) of the winner-pays-bid, all-pay, and truthful surrogate-ranking mechanisms in stationary equilibrium with samples are equal.*

Proof. This theorem follows because to each agent the mechanism looks like an i.i.d. position auction with characteristic weights that are independent of the payment format. Such i.i.d. position auctions admit only the efficient equilibrium and thus are welfare equivalent. Welfare equivalence implies, by the usual argument, revenue equivalence. □

Theorem 9 gives a revelation principle for surrogate ranking mechanisms. Bounds on the revenue and welfare of the truthful surrogate ranking mechanism implies the same bounds on that of the non-truthful ones because their equilibrium outcomes are the same in expectation.

4 Approximation Analysis

In this section, we state and briefly discuss our approximation analysis for surrogate ranking mechanisms. Condition C3 requires that there exist some choice of surrogate values which induces near-optimal expected performance. In Appendix C, we show how to construct a choice of surrogate values which proves the following guarantee:

Theorem 10. *There exists a surrogate-ranking mechanism that attains a $(1 - \tilde{O}(\sqrt[3]{n/T}))$-fraction of the optimal welfare. With regular distributions, there exists such a mechanism which attains a $(1 - \tilde{O}(\sqrt[5]{n^3/T}))$-fraction of the optimal revenue.*

To prove Theorem 10, we first argue that there exists an approximately-optimal mechanism which coarsens quantile space into bins, and treats agents based solely on their bin. We then show that this bin-based mechanism can be transformed into a surrogate-ranking mechanism with only polynomial loss in performance. See Appendix C for further details.

A main challenge in proving approximation bounds in sample complexity is arguing that agents at the extremes, i.e., quantiles close to 0 and 1, are treated appropriately. For revenue, allocation to quantiles close to 0 could have high positive contribution to the revenue while allocation to quantiles close to 1 could have high negative contribution to the revenue. Our approach is to treat all agents that have quantile close to 0 as if they have quantile 0 and all who have quantile close to 1 as if they have quantile 1. The key step in the analysis is then to show that the loss from this change, which disrupts the allocation to other agents, is bounded. The subsequent analysis is then relatively straightforward.

A key lemma in the analysis gives bounds virtual surplus from allocation rules that are approximately similar and may be of independent interest. Recall the characterization of expected revenue in terms of revenue curves and marginal revenue from Lemma 2:

$$\mathbf{E}_q[p(q)] = \mathbf{E}_q[-x'(q)\, R(q)] + R(1)\, x(1) = \mathbf{E}_q[R'(q)\, x(q)] + R(0)\, x(0). \tag{1}$$

The first equality enables a geometric understanding of revenue. Given an fixed allocation rule x in quantile space, for two value functions v_1 and v_2 where $R_i(q) = q\, v_i(q)$ satisfies $R_1(q) \geq R_2(q)$, then the revenue from v_1 on x is at least the revenue of v_2 on x. This follows from the first equality of equation (1), where the expressions for revenue of both value functions are weighted integrals over $q \in [0, 1]$ with non-negative weights $-x_i'(q)$. Approximation bounds hold as well, specifically, if R_2 approximates R_1 at all $q \in [0, 1]$ then the same approximation holds for the revenue of any fixed allocation rule x. Note however that a similar result, with a fixed distribution and similar allocation rules is not implied by the second equation, as the weights $R'(q)$ are not generally all the same sign. Instead, the following lemma shows that two allocation rules with *inverses* that are approximately close have approximately the same revenue. We state the result for general virtual value functions $\phi(\cdot)$ and cumulative virtual curves $\Phi(q) = \int_0^q \phi(r)\, dr$. The assumption in the lemma on the cumulative virtual curve is that lines from the origin pass from below to above the curve. This assumption, for example, is satisfied by any revenue curve, specifically, it does not require regularity.

Lemma 11. *For virtual value function $\phi(\cdot)$ and cumulative virtual value $\Phi(q) = \int_0^q \phi(r)\, dr$ satisfying $\Phi(\alpha\, q) \geq \alpha\, \Phi(q)$ for all quantiles q and $\alpha \in [0, 1]$, and any two allocation rules x_1 and x_2 that satisfy $x_1^{-1}(z) \geq x_2^{-1}(z) \geq \frac{1}{\alpha} x_1^{-1}(z)$, the expected virtual surpluses satisfy*

$$\mathbf{E}_q[\phi(q)\, x_2(q)] + \Phi(0)\, x_2(0) \geq \tfrac{1}{\alpha}[\mathbf{E}_q[\phi(q)\, x_1(q)] + \Phi(0)\, x_1(0)].$$

Proof. The virtual surplus of any allocation rule x can be rewritten as $\int_0^1 \phi(q)\, x(q)\, dq + \Phi(0)\, x(0) = \int_0^1 \Phi(x^{-1}(z))\, dz$. This follows by the first equality of equation (1) and a change of variables to integrate the vertical axis rather than the horizontal axis as follows:

$$\begin{aligned} \int_0^1 -x'(q)\, \Phi(q)\, dq + \Phi(1)\, x(1) &= \int_{x(1)}^{x(0)} \Phi(x^{-1}(z))\, dz + \int_0^{x(1)} \Phi(1)\, dz \\ &= \int_0^{x(0)} \Phi(x^{-1}(z))\, dz. \end{aligned} \tag{2}$$

Notice that the second line follows from the first line because $x^{-1}(z) = 1$ for $z \in [0, x(1)]$.

Now consider two arbitrary quantiles q_1 and q_2 satisfying $\frac{1}{\alpha} q_1 \leq q_2 \leq q_1$. By assumption, we have $\Phi(q_2) \geq q_2 \Phi(q_1)/q_1 \geq \frac{1}{\alpha}\Phi(q_1)$. The assumption on the approximation of the two allocation rules, namely $x^{-1}(z) \geq \hat{x}^{-1}(z) \geq \frac{1}{\alpha} x^{-1}(z)$ for all $z \in [0, 1]$, and the expected virtual surplus written as rewritten in equation (2) both both x_1 and x_2, then, suffice to prove the lemma. ☐

5 Reduction from Sample Complexity to Rank-Based Inference

We have shown that surrogate-ranking mechanisms possess a unique stationary equilibrium (Theorem 8), and that this equilibrium may be analyzed as if it was truthful (Theorem 9). In this section, we show how to use bid data to design a surrogate-ranking mechanism with near-optimal welfare or revenue in stationary equilibrium. Specifically, we reduce this design problem to an inference problem which is better-understood: estimating expected order statistics from bid data in i.i.d. position auctions. This inference problem was solved by Chawla et al. (2017) for first-price and all-pay mechanisms and is straighforward for truthful mechanisms.

Before giving details, we describe the high-level approach. Recall from Section 4 that, for revenue with regular distributions or welfare with general distributions, polynomially many surrogate values T per agent suffice to obtain a $(1-\epsilon)$-fraction of the optimal revenue or welfare via a surrogate-ranking mechanism. Consider a surrogate-ranking mechanism with T surrogate values per agent. We first show in Section 5.1 that the revenue- or welfare-optimal choice of these nT surrogate values requires only knowledge of the order statistics of the distribution. In Section 5.2 we oberve that this design approach is robust to error from inference: if one uses imperfect estimates of expected order statistics to design a SRM, then estimation error will propagate cleanly to revenue or welfare loss. Composing these three observations above yields the desired reduction.

Proposition 12. *For any n-agent single-dimensional independent-private-value environment, any unknown monotonic virtual value functions for the agents, and any estimator of expected order statistics of T samples from a virtual value distribution that attains error at most ϵ with probability at least $1-\epsilon$ from a polynomial number of samples in T and $1/\epsilon$ (from the bid distirbution of an i.i.d. position auction); the surrogate-ranking mechanisms with T selected to be polynomial in n and $1/\epsilon$ and surrogate values given by these estimated order statistics gives a mechanism that attains expected virtual surplus at most $n\epsilon$ less than optimal.*

Section 5.3 concludes by instantiating the reduction with an estimator for the requisite order statistics. Specifically, Chawla et al. (2017) show how to estimate expected order statistics using bid data from all-pay and winner-pays-bid position auctions with bounded distributions, and we show in Appendix D how to estimate expected order statistics with truthful data from unbounded regular distributions. These results imply polynomial sample complexity for winner-pays-bid, all-pay, and truthful mechanism design.

5.1 Optimal Surrogate-Ranking Mechanisms

Surrogate-ranking mechanisms are parametrized by nT surrogate values. Each choice of surrogate values induces a different allocation rule in stationary equilibrium, but by Theorem 9, this equilibrium allocation rule is the same under any of the standard payment formats. Thus optimal surrogate values can be determined assuming bids are truthful. In this section, we characterize the welfare- and revenue-optimal choices of surrogate values. To choose our surrogate values optimally, we consider a relaxed problem of maximizing a generic virtual surplus quantity subject to the constraint that the allocation rule depend only on each agent's rank among $T-1$ other truthful run-time samples. The solution to this problem is straightforward given the simple observation that if the only information we have to make decisions on is the rank of an agent against samples from her value distribution, then decisions should be made based on expected order statistics. The right choice of surrogate values is the expected order statistics of the quantity of interest for the objective, i.e., for welfare maximization its order statistics for the value distribution for revenue maximization its order statistics for the distribution of marginal revenues (via the characterization of expected revenue in Lemma 2). A formal proof of this theorem is in Appendix B.

Theorem 13. *The welfare-optimal surrogate-ranking mechanism uses surrogate values* $\psi_i^j = \mathbf{E}_{q_i}[v_i(q_i) \mid r_i = j]$. *For regular distributions, the revenue-optimal surrogate-ranking mechanism uses surrogate values* $\psi_i^j = \mathbf{E}_{q_i}[R_i'(q_i) \mid r_i = j]$ *where* $R_i'(q_i) = v_i(q_i) + q_i\, v_i'(q_i)$ *is the marginal revenue of agent* i *at quantile* q_i.

Because the welfare- and revenue-optimal surrogate ranking mechanisms are at least as good as any other surrogate ranking mechanism, it follows that they inherit the welfare and revenue guarantees of any other such mechanism. In particular, we obtain the following corollary to Theorem 10:

Corollary 14. *The welfare- (resp. revenue-) optimal surrogate ranking mechanism obtains a* $(1 - \tilde{O}(\sqrt[3]{n/T}))$*-fraction of the optimal welfare (resp. a* $(1 - \tilde{O}(\sqrt[5]{n^3/T}))$*-fraction of the optimal revenue with regular distributions) in stationary equilibrium.*

5.2 Propagation of Error

The optimal surrogate ranking mechanism for welfare (resp. revenue) uses surrogate values equal to the expected order statistics of each agent's value (resp. virtual value) distribution. We now show that a designer can in fact use noisy estimates of these quantities, and the performance will degrade smoothly with the estimation error. As before, we present our result for an arbitrary monotonic virtual value function ϕ_i for each agent.

For monotonic ϕ_i, though the expeted order statistics are monotone, estimates of these order statistics may not be. However, if the estimates of agent i's order statistics are within an agent specific error ϵ_i of being correct, than so will any natural method for making these estimates monotone, e.g., using the jth estimate of $\max_{j' \leq j} \hat{\psi}_i^{j}$ instead of $\hat{\psi}_i^j$. Of course, the recommended method is the standard approach of ironing, which for quantities like order statistics is formally described in Devanur et al. (2015) and Chawla et al. (2017). In fact ironing is equivalent in this setting to isotonic regression. An advantage of ironing is that it is the correct approach when the original ϕ_i is non-monotonic and, omitting the details and consequences, our results for revenue can be extended to *tail regular* distributions using this approach, cf. Devanur et al. (2015). For the remainder of the discussion, without loss of generality, we assume that the error estimates are monotonic.

The following theorem shows that errors in estimated order statistics propagate in a well-behaved fashion in surrogate-ranking mechanisms.

Theorem 15. *For all* i *and* j*, let* ψ_i^j *be the expected* j*th order statistic of agent* i*'s virtual value distribution, and let* $\hat{\psi}_i^j$ *be an estimate of* ψ_i^j *satisfying* $|\hat{\psi}_i^j - \psi_i^j| < \epsilon_i$*, where* ϵ_i *is an agent-specific error bound. The difference between the expected virtual surplus of the surrogate ranking mechanisms with the true expected order statistics and estimated order statistics is at most* $2 \sum_i \epsilon_i$.

Proof. Let $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ denote the allocation rule of the surrogate-ranking mechanism as a function of agents' ranks $\boldsymbol{r}$ among their run-time samples with optimal surrogate values $\boldsymbol{\Psi}$ and estimated and ironed surrogate values $\hat{\mathbf{\Psi}}$, respectively. The theorem follows from:

$$\begin{aligned}
\mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i \mathbf{E}_{q_i}\left[\psi_i(q_i)\,|\,r_i\right]\hat{x}_i(\boldsymbol{r})\Big] &= \mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i \psi_i^{r_i}\hat{x}_i(\boldsymbol{r})\Big] \\
&\geq \mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i (\hat{\psi}_i^{r_i} - \epsilon_i)\hat{x}_i(\boldsymbol{r})\Big] \\
&\geq \mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i \hat{\psi}_i^{r_i}\,\hat{x}_i(\boldsymbol{r})\Big] - \sum\nolimits_i \epsilon_i \\
&\geq \mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i \hat{\psi}_i^{r_i}\,x_i(\boldsymbol{r})\Big] - \sum\nolimits_i \epsilon_i \\
&\geq \mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i (\psi_i^{r_i} - \epsilon_i)\,x_i(\boldsymbol{r})\Big] - \sum\nolimits_i \epsilon_i \\
&\geq \mathbf{E}_{\boldsymbol{r}}\Big[\sum\nolimits_i \psi_i^{r_i}\,x_i(\boldsymbol{r})\Big] - 2\sum\nolimits_i \epsilon_i.
\end{aligned}$$

The second and fifth lines follow from the assumption that $|\hat{\psi}_i^j - \psi_i^j| < \epsilon_i$, and the fourth line follows from the fact that $\hat{\boldsymbol{x}}$ is the allocation rule that maximizes $\mathbf{E}_{\boldsymbol{r}}\left[\sum_i \hat{\psi}_i^{r_i}\,\hat{x}_i(\boldsymbol{r})\right]$. The last line is the expected virtual surplus of the optimal surrogate-ranking mechanism $\boldsymbol{x}$, which implies the result. □

5.3 Sample Complexity of Non-Truthful Mechanisms

We can now formalize the reduction from non-truthful sample complexity to inference in rank-based position auctions. In Section 3, we observed that the data generated by an agent in a SRM is distributed according to the unique BNE of an i.i.d. position auction. In Section 4, we demonstrated the existence of SRMs with near-optimal welfare and revenue, and in Section 5.1 and Section 5.2, we showed that it was possible to construct such a mechanism from noisy estimates of expected order statistics. We now instantiate the reduction by noting that one can use bid data from SRMs to infer these parameters efficiently.

Chawla et al. (2017) study the problem of inferring order statistics from bid data in all-pay and winner-pays-bid i.i.d. position auctions. They show that for any non-trivial position weights, it is possible to efficiently infer order statistics for both the values and marginal revenues. We summarize their results below:

Theorem 16 (Chawla et al., 2017)**.** *Consider a T-agent all-pay or winner-pays-bid i.i.d. position auction with arbitrary position weights and values in $[0,1]$. There exists an estimator $\hat{V}_k$ for the expected kth order statistic of the value distribution V_k such that with $N \geq O(T^4(\log^2(1/\delta + T) + \log^2(1/\epsilon + T))\epsilon^{-2}\delta^{-2})$ sampled bids from the unique BNE, $|\hat{V}_k - V_k| \leq \epsilon$ with probability at least $1-\delta$.*

Theorem 17 (Chawla et al., 2017)**.** *Consider a T-agent all-pay or winner-pays-bid i.i.d. position auction with arbitrary position weights and values in $[0,1]$. There exists an estimator $\hat{\Psi}_k$ for the expected kth order statistic of the value distribution Ψ_k such that with $N \geq O(T^4(\log^2(1/\delta) + \log^2(1/\epsilon))\epsilon^{-2}\delta^{-2})$ sampled bids from the unique BNE, $|\hat{\Psi}_k - \Psi_k| \leq \epsilon$ with probability at least $1-\delta$.*

We note that the above results combine with Proposition 12 to imply a solution to the non-truthful sample complexity problem for agents with values in $[0,1]$ and additive loss. We summarize below:

Theorem 18. *For agents with bounded values in* $[0, 1]$*, there are families of winner-pays-bid and all-pay mechanisms that satisfy C1, C2, and C3 for additive loss and the welfare objective. If the agents' distributions are regular and bounded, then the same result also holds for the revenue objective.*

Our results have new implications for the truthful sample complexity literature as well. Specifically, we give polynomial sample complexity for revenue maximization with unbounded regular distributions and general feasibility settings. This extends the result of Devanur et al. (2016) by dropping the downward-closure requirement on the feasibility constraint. Details of this instantion of the reduction are in Appendix D.

Theorem 19. *For agents with regularly distributed values (but potentially unbounded), there is a family of truthful mechanisms that satisfies C1, C2, and C3 for multiplicative approximation and the revenue objective.*

We conclude by noting that the polynomials in the sample complexity guarantees of Theorem 18 and Theorem 19 are unwieldy. We leave improving them to future work.

References

Agrawal, S., Ding, Y., Saberi, A., and Ye, Y. (2010). Correlation robust stochastic optimization. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 1087–1096.

Akbarpour, M. and Li, S. (2018). Credible mechanisms. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 371–371. ACM.

Ausubel, L. M. and Milgrom, P. (2006). The lovely but lonely Vickrey auction. *Combinatorial auctions*, pages 17–40.

Bei, X. and Huang, Z. (2011). Bayesian incentive compatibility via fractional assignments. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 720–733.

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

Chawla, S. and Hartline, J. D. (2013). Auctions with unique equilibria. In *ACM Conference on Electronic Commerce*, pages 181–196.

Chawla, S., Hartline, J. D., and Nekipelov, D. (2017). Mechanism redesign. *arXiv preprint arXiv:1708.04699*.

Chawla, S., Hartline, J. D., and Sivan, B. (2015). Optimal crowdsourcing contests. *Games and Economic Behavior*.

Cole, R. and Roughgarden, T. (2014). The sample complexity of revenue maximization. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 243–252. ACM.

Dasgupta, A. and Ghosh, A. (2013). Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd international conference on World Wide Web*, pages 319–330. ACM.

Devanur, N. R., Hartline, J. D., and Yan, Q. (2015). Envy freedom and prior-free mechanism design. *Journal of Economic Theory*, 156:103–143.

Devanur, N. R., Huang, Z., and Psomas, C. (2016). The sample complexity of auctions with side information. In *STOC 2016*, pages 426–439.

Dhangwatnotai, P., Roughgarden, T., and Yan, Q. (2010). Revenue maximization with a single sample. In *ACM Conference on Electronic Commerce*, pages 129–138.

Dughmi, S., Hartline, J. D., Kleinberg, R., and Niazadeh, R. (2017). Bernoulli factories and black-box reductions in mechanism design. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 158–169. ACM.

Elkind, E. (2007). Designing and learning optimal finite support auctions. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 736–745. Society for Industrial and Applied Mathematics.

Feng, Y. and Hartline, J. D. (2018). An end-to-end argument in mechanism design (prior-independent auctions for budgeted agents). In *IEEE 59th Annual Symposium on Foundations of Computer Science*.

Fu, H., Haghpanah, N., Hartline, J., and Kleinberg, R. (2014). Optimal auctions for correlated buyers with sampling. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 23–36. ACM.

Gonczarowski, Y. A. and Nisan, N. (2017). Efficient empirical revenue maximization in single-parameter auction environments. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 856–868. ACM.

Hartline, J. and Lucier, B. (2010). Bayesian algorithmic mechanism design. In *ACM Symposium on Theory of Computing*, pages 301–310.

Hartline, J., Syrgkanis, V., and Tardos, E. (2015a). No-regret learning in bayesian games. In *Advances in Neural Information Processing Systems*, pages 3061–3069.

Hartline, J. D. (2013). Bayesian mechanism design. *Foundations and Trends® in Theoretical Computer Science*, 8(3):143–263.

Hartline, J. D., Kleinberg, R., and Malekian, A. (2011). Bayesian incentive compatibility via matchings. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 734–747.

Hartline, J. D., Kleinberg, R., and Malekian, A. (2015b). Bayesian incentive compatibility via matchings. *Games and Economic Behavior*, 92:401–429.

Hartline, J. D. and Lucier, B. (2015). Non-optimal mechanism design. *American Economic Review*, 105(10):3102–24.

Jansen, B. J. and Mullen, T. (2008). Sponsored search: an overview of the concept, history, and technology. *International Journal of Electronic Business*, 6(2):114–131.

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

Osband, K. (1989). Optimal forecasting incentives. *Journal of Political Economy*, 97(5):1091–1112.

Roughgarden, T. and Schrijvers, O. (2016). Ironing in the dark. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 1–18.

Yan, Q. (2011). Mechanism design via correlation gap. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 710–719.

A Undoing the Revelation Principle

Good first-price and all-pay mechanisms for a given environment can be found by undoing the revelation principle (ignoring computational complexity). This construction applies to any revelation mechanism $\mathcal{M}$. For concreteness, imagine a applying this approach to a single-minded combinatorial auction problem where $\mathcal{M}$ is the Vickrey-Clarke-Groves (VCG) mechanism. We give the all-pay version of the construction which is slightly simpler, but exhibits the same issues.

Definition 20. The all-pay unrevelation mechanism *for a revelation mechanism* $\mathcal{M}$ *is:*

1. *For each agent* i *and value* v_i*, calculate* $s_i(v_i)$ *as the expected payment in* $\mathcal{M}$ *when the agent's value is* v_i *and other agents' values are drawn from the distribution.*
2. *For each agent* i*, given bid* b_i *in the un-revelation mechanism, calculate the agent's value as* $v_i = s_i^{-1}(b_i)$*.*
3. *Serve the agents who are served by* $\mathcal{M}$ *on values* $\mathbf{v} = (v_1, \ldots, v_n)$*; all agents pay their bids.*

The characterization of Bayes-Nash equilibrium (Theorem 1) implies that s_i is the strategy that agents will employ in equilibrium of the constructed all-pay mechanism. Thus, the all-pay mechanism has the same equilibrium outcome.

From this definition we can see why symmetric and ordinal environments (i.e., IID position environments) are special. For these environments all agents will have the same strategy function, this strategy function will order higher valued bidders higher (by monotonicity), and the ordinal environment then implies that all that is needed to select an outcome is the order of values not their cardinal values. Thus, the mechanism simplifies to simply ordering the bids and the strategy function does not need to be calculated.

Even absent computational issues in estimating the strategy functions so as to implement this mechanism, it is clear that very detailed distributional information is needed to run the unrevelation mechanism. Moreover, the resulting outcomes may be very sensitive to small errors with the inversion of the strategy function. This unrevelation mechanism is not to be considered practical.

B Proof of Theorem 13

Theorem 13. *The welfare-optimal surrogate-ranking mechanism uses surrogate values* $\psi_i^j = \mathbf{E}_{q_i}[v_i(q_i) \mid r_i = j]$*. For regular distributions, the revenue-optimal surrogate-ranking mechanism uses surrogate values* $\psi_i^j = \mathbf{E}_{q_i}[R_i'(q_i) \mid r_i = j]$ *where* $R_i'(q_i) = v_i(q_i) + q_i\, v_i'(q_i)$ *is the marginal revenue of agent* i *at quantile* q_i*.*

Proof. We define the *rank-based allocation problem* as follows: the designer must choose an allocation rule $\bar{x}$ which takes as input the rank $\boldsymbol{r} = (r_1, \ldots, r_n)$ of each agent among $T-1$ runtime samples for their distributions and outputs a (possibly randomized) feasible allocation $\bar{\boldsymbol{x}}(\boldsymbol{r})$. As a constraint, $\bar{\boldsymbol{x}}$ must be monotone in the ranks of each agent. The objective is to maximize $\mathbf{E}[\sum_i \phi_i(q_i)\bar{x}_i(\boldsymbol{r})]$ for some given virtual value function $\phi_i(\cdot)$, where the expectation is over agents' quantiles being uniformly distributed and over the runtime samples used to compute $\boldsymbol{r}$. For example, $\phi_i(q_i) = v_i(q_i)$ corresponds to welfare maximization and $\phi_i(q_i) = R_i'(q_i)$ corresponds to revenue maximization.

The rank-based allocation problem can be solved by inspection. Fixing an allocation rule, the objective can be rewritten as $\sum_i \mathbf{E}[\phi_i(q_i) \mid r_i]\ \bar{x}_i(\boldsymbol{r})$ by linearity of expectation. From this expression, it becomes clear that the optimal solution chooses the allocation which maximizes the quantity $\sum_i \mathbf{E}[\phi_i(q_i) \mid r_i]\ \bar{x}_i(\boldsymbol{r})$. Note that if $\phi_i(\cdot)$ is monotone, then this allocation rule will be monotone as well, and therefore feasible.[6] Setting these expected order statistics as surrogate values, the surrogate-ranking mechanism (Definition 4) optimizes this quantity. □

C Proof of theorem 10

In this appendix, we derive a surrogate-ranking mechanism with near-optimal welfare or revenue, which proves the following theorem, restated for convenience:

Theorem 10. *There exists a surrogate-ranking with winner-pays-bid, all-pay, or truthful payment semantics which attains a* $(1 - \tilde{O}(\sqrt[3]{n/T}))$*-fraction of the optimal welfare. With regular distributions, there exists such a mechanism which attains a* $(1 - \tilde{O}(\sqrt[5]{n^3/T}))$*-fraction of the optimal revenue in stationary equilibrium.*

Before proving the main theorem, we observe that a less general result follows from a theorem of Hartline et al. (2011). In this paper, the authors consider an allocation procedure which can be interpreted as a surrogate ranking mechanism. They show that for agents whose values are distributed on the interval $[0, 1]$, their mechanism has an additive welfare loss of at most $n/(4\sqrt{T})$. We improve on this welfare guarantee in three ways. First, we derive a guarantee for arbitrary distributions, even those with unbounded support. Second, our bounds will be multiplicative. Finally, our bounds apply to revenue in addition to welfare.[7]

C.1 Ranking Versus Pricing

For a mechanism to maximize welfare or revenue effectively, it must be able to discriminate between agents with high and low values. To prove Theorem 10 we must show that ranking mechanisms can do this effectively. We build towards this goal by first showing that ranking mechanisms can approximate the simplest form of discrimination: allocating all quantiles below a threshold q. This is equivalent to posting a price at $v(q)$. We in particular consider thresholds for which q is an integral multiple of $1/T$. Formally:

Definition 21. *The* k/T*-price posting algorithm allocates agents if and only if their quantile is below* k/T*, for some integer* k*. This can be achieved by posting the price with quantile* k/T*.*

[6] If it is not monotone, then then the resulting surrogate values may not be monotone, if the surrogate values by this approach are not monotone, they can be ironed using the standard procedure.

[7] With the additional assumption that virtual values are bounded below by $-\underline{\phi}$, the guarantee of Hartline et al. (2011) applies to revenue as well, with an additional factor of $1 + \underline{\phi}$ applied to the loss.

We first show that price-posting can be approximated by ranking. With unlimited supply, posting a price at the quantile k/T will result in allocation to k of the T agents in expectation. The rank-based equivalent enforces this quota pointwise, allocating the k highest-valued agents from T identically distributed agents.

Definition 22. *The top k-of-T algorithm for T agents ranks agents by value and allocates the k agents with the highest values.*

As the law of large numbers might suggest, these two algorithms perform comparably for large n when k is bounded away from the extremes (1 and $T-1$). Formally, we show:

Lemma 23. *The top k-of-T algorithm attains at least a $\rho(k)$-fraction of the welfare of the k/T-price posting algorithm with n agents. If values are regularly distributed, then it attains a $\eta(\min(k, T-k), T)$-fraction of the revenue of the k/T-price posting algorithm, where*

$$\rho(k) = 1 - O\left(\sqrt{\frac{\ln k}{k}}\right) \quad \textit{and} \quad \eta(k,T) = 1 - O\left(\frac{T-k}{k^{3/2}}\sqrt{\log k}\right),$$

Proof. We argue separately for the objectives of welfare and virtual surplus, but in both cases, the proof strategy will be the same. We will first explicitly characterize the worst-case distribution for each objective, and then we will analyze the performance ratio of the top-k-of-T algorithm and k/T-price posting algorithms using standard concentration bounds. In what follows, we will suppress subscripts denoting a particular agent when the agent's identity is irrelevant.

Key to the analysis will be two formulae for the expected surplus of an algorithm, in terms of its interim allocation rule $x(\cdot)$ and the distribution's value function $v(\cdot)$. We have that an algorithm's surplus is:

$$\mathbb{E}_{q\sim U[0,1]}[x(q)v(q)] = \mathbb{E}_{q\sim U[0,1]}[-x'(q)V(q)], \tag{3}$$

where $V(q) = \int_0^q v(z)\,dz$, and the equality follows from integration by parts. An analogous formula holds for virtual surplus, with $v(q)$ replaced by the Myerson virtual value at q. The only real difference between the two objectives is the fact that values are always positive, whereas virtual values may be negative. This will change the nature of the approximation, as allocating the wrong agent becomes actively harmful to the performance of the algorithm.

Welfare. We begin by normalizing the per-agent surplus of the price-posting mechanism to 1. Note that for the k/T-price posting algorithm, the allocation rule is 1 until quantile k/T, and then drops to 0. It follows from equation (3) that our normalization is equivalent to the assumption that $V(k/T) = 1$.

Next, we note that because $v(\cdot)$ is positive and decreasing, $V(\cdot)$ is increasing and concave, with $V(0) = 0$. Let $x(\cdot)$ be the allocation rule of the top-k-of-T algorithm. Given our normalization, the problem of finding the worst-case distribution then becomes:

$$\begin{aligned} \min_{V(\cdot)} \quad & \mathbb{E}_{q\sim U[0,1]}[-x'(q)V(q)] \\ \text{subject to} \quad & V(0) = 0 \\ & V(k/T) = 1 \\ & V(\cdot) \text{ concave} \\ & V(\cdot) \text{ increasing} \end{aligned}$$

This program can be solved by inspection by noticing that there is pointwise minimal function satisfying the constraints of the program: namely, the optimal $V(q)$ is linear with slope $v(q) = T/k$ for $q \leq k/T$, and constant at 1 for $q \geq k/T$. This corresponds to the distribution with k/T mass on the value T/k, and the rest on 0.

Having derived the distribution with the worst gap between ranking and pricing, we now analyze the performance of ranking on this distribution. Let X be the random variable denoting the number of agents with positive value, and for all $i \in \{1, \ldots, T\}$, let X_i be an indicator random variable for the event that agent i's value is positive. Note that $X = \sum_{i=1}^{T} X_i$ and $\mathbb{E}[X] = k$. Further note that the expected welfare from the top-k-of-T algorithm is $\mathbb{E}[\min(X, k)T/k]$. The standard multiplicative Chernoff bound states that for any $\delta \in [0, 1]$,

$$\Pr[X \leq (1 - \delta)k] \leq e^{-\frac{\delta^2 k}{2}}$$

We may therefore lower bound the welfare of the top k-of-T algorithm as

$$\mathbb{E}[\min(X, k)T/k] \geq (1 - \delta)(1 - e^{-\frac{\delta^2 k}{2}})T.$$

Choosing $\delta = \sqrt{\frac{\ln k}{k}}$ yields the result. Note that the analysis above also bounds the correlation gap for k-uniform matroids (Agrawal et al., 2010; Yan, 2011).

Virtual Surplus. We now adapt the above proof to virtual surplus. The main difference is that the Myerson virtual value, denoted $\phi(q)$, can be negative. We will additionally use the fact that the Myerson virtual value is the derivative of the price-posting revenue curve. That is, $\phi(q) = R'(q) = \frac{d}{dq} v(q)(1 - q)$. It follows that cumulative virtual value has the convenient form $R(q) = v(q)q$.

As before, we normalize the virtual surplus from price-posting to 1. This corresponds with setting $R(q) = 1$. Subject to normalization, we use properties of revenue curves to derive the worst-case distribution for virtual surplus. We assume values are regularly distributed, which implies that $R(q)$ is concave. Moreover, since $R(q) = v(q)q$, we have that $R(0) = R(1) = 0$. These properties yield the following program for the worst-case distribution:

$$\begin{aligned} \min_{R(\cdot)} \quad & \mathbb{E}_{q \sim U[0,1]}[-x'(q)R(q)] \\ \text{subject to} \quad & R(0) = R(1) = 0 \\ & R(k/T) = 1 \\ & R(\cdot) \text{ concave} \end{aligned}$$

Again, this may be solved by inspection. The worst-case $R(\cdot)$ is triangular, with its apex at $(k/n, 1)$. That is, on $[0, k/T]$, $R(q)$ has slope T/k, and on $[k/T, 1]$, it has slope $-T/(T - k)$. In other words, the worst-case distribution for virtual surplus has virtual value T/k with probability k/T, and virtual value $-T/(T - k)$ otherwise.

We divide the analysis of the worst-case distribution for virtual surplus into two cases: $k \leq T/2$ and $k > T/2$. For $k \leq T/2$, we will prove that the expected virtual surplus of the top-k-of-T algorithm is at least a $\eta_1(T, k)$-fraction of that of the k/T-price posting algorithm for some

$$\eta_1(k, T) = 1 - O\left(\sqrt{\log k}\frac{n - k}{k^{3/2}}\right).$$

For $k \geq T/2$, we will instead prove an approximation ratio of

$$\eta_2(k,T) = 1 - O\left(\sqrt{\log(T-k)}\frac{k}{(T-k)^{3/2}}\right).$$

Since for $k \leq T/2$, $\eta_2(k,T)$ will represent a better approximation than $\eta_1(k,T)$ and vice versa for $k \geq T/2$, the result will follow.

First assume $k \leq T/2$. Let $X_1, \ldots, X_T$ be indicator variables for the event that agent i has a realized virtual value of T/k. Let $X = \sum_{i=1}^{T} X_i$ denote the number of agents with realized virtual value T/k. As we did for welfare, for any $\delta \in [0,1]$, we may use a multiplicative Chernoff bound to write

$$\Pr[X \leq (1-\delta)k] \leq e^{-\frac{\delta^2 k}{2}}.$$

We may use this fact to obtain the following lower bound on the expected virtual surplus of the top-k-of-T algorithm:

$$\left(n - \delta\frac{T}{T-k}\right)\left(1 - e^{-\frac{\delta^2 k}{2}}\right) - e^{-\frac{\delta^2 k}{2}}\frac{kT}{T-k}. \tag{4}$$

The first term comes from the event that $X \geq (1-\delta)k$, in which case the virtual surplus is at least $(1-\delta)T - \delta\frac{kT}{T-k} = (1 - \delta T/(T-k))$. The second term comes from the event that $X \leq (T-\delta)k$, in which case the virtual surplus is at least $kT/(T-k)$. Using the fact that $k \leq T/2$, we may lower bound (4) as:

$$T\left[(1-2\delta)(1-e^{-\frac{\delta^2 k}{2}}) - e^{-\frac{\delta^2 k}{2}}\right].$$

Taking $\delta = \sqrt{\log k/k}$ yields an lower bound of $n(1 - O(\sqrt{\log k/k}))$. We may weaken this guarantee to produce the bound in the lemma statement by noting that $(T-k)/k \geq 1$.

We now argue the $k \geq T/2$ case. Let $Y_1, \ldots, Y_T$ be indicator variables for the event that agent i has realized virtual value $-T/(T-k)$. Let $Y = \sum_{i=1}^{n} Y_i$ denote the number of agents with virtual value $-T/(T-k)$. Again using a standard multiplicative Chernoff bound, we may choose any $\delta \in (0,1)$ and write:

$$\Pr[Y \geq (1+\delta)(n-k)] \leq e^{-\frac{\delta^2(T-k)}{3}}.$$

The expected virtual surplus of the top-k-of-T algorithm is at least

$$\left((T-(1+\delta)(T-k))\frac{T}{k} - \frac{T}{T-k}\delta(T-k)\right)\left(1 - e^{\frac{-\delta^2(T-k)}{3}}\right) - e^{\frac{-\delta^2(T-k)}{3}}\frac{Tk}{T-k} \tag{5}$$

$$= T\left[(1-\delta T/k)\left(1 - e^{\frac{-\delta^2(T-k)}{3}}\right) - e^{\frac{-\delta^2(T-k)}{3}}\frac{k}{T-k}\right] \tag{6}$$

The first term in (5) represents the virtual surplus from the event that $Y \leq (1+\delta)(T-k)$. At least $n - (1+\delta)(T-k)$ agents win with virtual surplus n/k, and at most $\delta(T-k)$ agents win with virtual surplus $-n/(n-k)$. The second term lower bounds the virtual surplus from the event that $Y \geq (1+\delta)(T-k)$. At most k winners can have virtual surplus $-T/(T-k)$. Using the fact that $k \geq T/2$, we may further lower bound (6) by

$$= T\left[(1-2\delta)\left(1 - e^{\frac{-\delta^2(T-k)}{3}}\right) - e^{\frac{-\delta^2(T-k)}{3}}\frac{k}{T-k}\right]$$

Taking $\delta = \sqrt{\frac{3\log(T-k)}{2(T-k)}}$ yields the desired guarantee. □

We can further generalize Lemma 23 by comparing distributions over pricing algorithms with the analogous distributions over top-k algorithms. As long as prices avoid the extremes of the distribution, ranking performs well with respect to pricing.

Lemma 24. *Consider a distribution over k/T-price posting algorithms for T agents, where the highest price is at quantile $\underline{k}/T$ and the lowest price is at quantile $\overline{k}/T$. The same distribution over corresponding top-k-of-T algorithms attains a $\rho(\min(\underline{k}, T-\overline{k}), T)$-fraction of the welfare of the distribution over price-posting algorithms. If values are regularly distributed, then the distribution over top-k-of-T algorithms attains an $\eta(\min(\underline{k}, T-\overline{k}), T)$-fraction of the price-posting revenue as well.*

Proof. Lemma 23 implies that for each price in the distribution of the price-posting algorithm, there is a top-k-of-T algorithm which approximates it and which appears with the same probability. The approximation ratio of a distribution over pairwise approximations is at least the approximation from the worst pair. Note that the approximations from Lemma 23 are symmetric about 1/2, and are worst for very low and very high k. It follows that the approximation ratio is driven by the $\underline{k}/T$- and $\overline{k}/T$-price posting algorithm. □

C.2 Binning Versus Ranking

Lemma 24 shows that in simple settings, ranking can discriminate with almost as much accuracy as pricing. We now extend this idea the setting of general mechanism design with runtime samples. We compare the allocation rule of a surrogate ranking mechanism (i.e. a *surrogate ranking algorithm*) with a *surrogate binning algorithm*, which assigns agents a surrogate value by dividing quantile space into coarse "bins" of uniform probability and associating each bin with a surrogate values. We formally define this algorithm as follows.

Definition 25. *Given a set of T surrogate values Ψ with $\psi^1 \geq \ldots \geq \psi^T$, the* binning *selection rule for Ψ maps each input quantile q to ψ^j for the j for which $q \in [(j-1)/T, j/T]$.*

Definition 26. *For surrogate values $\{\Psi_i\}_{i=1}^n$, and binning surrogate selection rules $\{\sigma_i\}_{i=1}^n$, the* surrogate binning algorithm *is given by computing $\tilde{\mathbf{x}}(\mathbf{q}) = \text{argmax}_{\mathbf{x}} \sum_i x_i \sigma_i(q_i)$.*

Any allocation rule can be viewed by agents as a distribution over posted prices. We note several properties of surrogate-binning algorithms which make this distribution easy to analyze. First, note that because quantiles are uniform for each agent, the distribution of surrogate values input to the allocation algorithm are also uniform. Second, because surrogate-binning mechanisms treat agents coarsely based on the bin into which their quantile falls, and because bins are distributed evenly in quantile space, it follows that for surrogate-binning mechanisms, the distribution over posted prices can actually be viewed as a distribution over k/T-quantile price posting algorithms.

Meanwhile, a standard fact from the study of position auctions is that any rank-based position auction can be represented as a convex combination of k-unit auctions. Since the allocation rule agents face in a surrogate-ranking mechanism is identical to that of a rank-based position auction, a similar analysis applies. We summarize the above discussion with the following lemma:

Lemma 27. *Any surrogate-binning (resp. surrogate-ranking) algorithm with surrogate values $\{\psi_i^j\}_{i=1,\ldots,n}^{j=1,\ldots,T}$ appears to agents in each distribution i as a distribution over price-posting (resp.*

top-k) algorithms. The probability of offering the price with quantile $\frac{j}{T}$ (resp. of allocating j units) is given by $w_i^j - w_i^{j+1}$, where $w_i^0 = 1$, $w_i^{T+1} = 0$, and w_i^j is the characteristic weight for ψ_i^j for $j = 1, \ldots, T$.

Notice that if $\psi_i^1 = \ldots = \psi_i^{\underline{k}}$, the binning algorithm's allocation rule on the first $\underline{k}$ intervals of distribution i's quantile space will be constant. In terms of distribution i's randomization over posted pricings, the highest nontrivial price offered has quantile $\underline{k}/n$, and the lowest has quantile $\overline{k}/n$. These extremal quantiles drive the approximation guarantees relating pricing to ranking and, thus, good approximation bounds can be obtained via Lemma 24 if there is not much loss in restricting to binning algorithms that price at moderate quantiles. Consequently, we obtain the following theorem:

Theorem 28. *For surrogate values $\psi_i^1 \geq \psi_i^2 \geq \ldots \geq \psi_i^T$ with $\psi_i^1 = \psi_i^2 = \ldots = \psi_i^{\underline{k}}$ and $\psi_i^{\overline{k}} = \psi_i^{\overline{k}+1} = \ldots = \psi_i^T$ for each population i, the surrogate ranking algorithm attains a $\rho(\underline{k})$-fraction of the welfare of the binning algorithm. If distributions are regular, then the surrogate ranking algorithm attains a $\eta(\min(\underline{k}, T - \overline{k}), T)$-fraction of the binning algorithm's virtual surplus.*

Proof. First note that to agent i, the distribution of surrogate values among agents from other agents is identical between binning and ranking - each surrogate value appears with probability $1/T$, independently across distributions. It follows that each surrogate value has the same characteristic weight in both mechanisms. It follows that the distribution of prices from the surrogate-binning mechanism matches the distribution over top-k mechanisms from the surrogate-ranking mechanism.

As we have seen, the respective distributions over prices and top-k mechanisms for binning and ranking are identical. Note that the above distributions might involve prices and k for which 0 or T agents get allocated. For purposes of approximation, these cases can be ignored, as the corresponding pricing and top-k mechanisms perform identically. It follows that we may apply Lemma 24. If $\psi_i^1 = \ldots = \psi_i^{\underline{k}}$ and $\psi_i^{\overline{k}} = \ldots = \psi_i^T$, then we have that the most extreme nontrivial values of k are $\underline{k}$ and $\overline{k}$. This yields the stated bound. □

C.3 Approximately Optimal Binning

We now show how to construct an approximately welfare- or revenue-optimal binning algorithm with large first and last bin. Combined with Lemma 27, this construction implies Theorem 10. We will choose surrogate values via resampling. That is, we will map the jth surrogate of agent i, which corresponds to an agent with quantile in $[\frac{j-1}{T}, \frac{j}{T}]$ to a redrawn value from this conditional distribution. Such resampling does not change the induced allocation rule for any other agents, and replaces the allocation rule for agent i on their jth quantile interval with its average.

The basic approach does not directly lead to the desired approximation bound. For an example where this fails, note that for the top quantile interval, the allocation probability at the very top of the interval may be much higher than its average across the interval, while the highest values on the interval may be much higher than the interval's average. For example, if the value and allocation rule are both one for an ϵ measure and zero otherwise, then the original welfare is ϵ and the welfare from resampling is ϵ^2. A second issue is that we wish to be able to apply Lemma 27 to get a good approximation bound. For example, we would get a $1 - O(\sqrt{\log k/k})$ bound for welfare if the top k intervals have the same allocation probability.

To resolve both these issues, we will first modify the allocation rule to treat agents with values in the top k intervals as if they had the highest value in the support of their distributions, and treat agents with values in the bottom k intervals as if they had the lowest value in the support of their distributions, for some given positive integer k. The quantiles of the remaining agents will be rescaled. Conditioned on the values not being in the top or bottom k intervals, the value distribution after rescaling will match the original unconditioned value distribution. We refer to this transformation as *extremal buffering*. Note that applying this procedure to one agent does change the outcomes for other agents. We show that this change does not have a significant impact on the outcomes other agents receive, and approximately preserves welfare and revenue from each population. Our analysis will hold for any choice of k. We will carefully select k later.

We analyze the extremal buffering procedure in Section C.3.1, and the binning algorithm that results from resampling in Section C.3.2. We will argue for an arbitrary monotone allocation algorithm, and a fairly general class of virtual value functions. Taking the allocation algorithm to be surplus maximization and the virtual value function to be value or Myerson virtual value will yield the following theorem:

Lemma 29. *For each agent i, let ϕ_i be a nondecreasing virtual value function, and let $\mathbf{x}$ be an a monotone allocation algorithm. Further, for some $k \leq T/2$, let $\hat{\mathbf{x}}$ denote the k/T-buffering algorithm for $\mathbf{x}$. Then there exists a binning algorithm for $\hat{\mathbf{x}}$ obtains at least a $\frac{k}{k+1}(1-\frac{q}{1-q})(1-q)(1-2(n-1)q)$-fraction of the expected virtual surplus under $\mathbf{x}$, with $q = k/T$.*

C.3.1 Extremal Buffering

We have seen that naïvely resampling each agent's value based on their bin cannot yield a mechanism with a good welfare or revenue approximation. We show now how to transform an arbitrary allocation algorithm to guarantee that not too much virtual surplus is lost from mishandling agents with extreme quantiles. The procedure follows:

Definition 30. *Given a monotone allocation algorithm $\mathbf{x}$ and a quantile $q \in [0,1]$, the q-*buffering algorithm *for $\mathbf{x}$ and $\mathbf{x}$ on agents with quantiles transformed for each agent as follows:*

- *For any $q_i \in [0, q]$, return 0.*
- *For any $q_i \in [q, 1-q]$, return $(q_i - q)/(1-2q)$.*
- *For any $q_i \in [1-q, 1]$, return 1.*

We will prove the following approximation guarantee:

Lemma 31. *For each agent i, let $\phi_i : [0,1] \to \mathbb{R}$ be an arbitrary nonincreasing virtual value function satisfying $\int_0^1 \phi_i(q)\,dq \geq 0$. The k/T-buffering algorithm for $\mathbf{x}$ attains at least a $(1-\frac{q}{1-q})(1-q)(1-2(n-1)q)$-fraction of the virtual surplus of $\mathbf{x}$, with $q = k/T$.*

The proof of Lemma 31 will proceed in two main steps. First, we will show that applying the quantile remapping procedure in Definition 30 to a single agent i (leaving other agents' quantiles untouched) cannot reduce the virtual surplus from that agent by too much. This will follow from a natural approximation result we derive, which relates the virtual surpluses of allocation rules with inverses that are multiplicatively close. Second, we will show that subsequently applying the quantile resampling procedure to the remaining agents other than i also does not significantly reduce the expected virtual surplus from i. This will follow from the fact that the distribution of quantiles input to the base allocation algorithm is identical, conditioned on no agents having extreme quantiles.

We begin with the single-agent analysis. Note that for a single agent, the q-buffering procedure can be thought of as two composed steps. First is a *top promotion* procedure, which remaps sufficiently low quantiles to 0 while remapping the remaining quantiles to induce a uniform distribution over $[0, 1]$. Top promotion is then composed with *bottom demotion*, which performs analogous transformation, mapping high quantiles to 1 and mapping the rest of the interval to $[0, 1]$. We formalize this as follows:

Definition 32. *Given a monotone single-agent allocation rule x and quantile $\underline{q}$, the* top promotion algorithm *for x and $\underline{q}$ runs x on the agent with quantiles transformed as follows:*

- *For any $q \in [0, \underline{q}]$, return 0.*
- *For any $q \in [\underline{q}, 1]$, return $(q - \underline{q})/(1 - \underline{q})$.*

Definition 33. *Given a monotone single-agent allocation rule x and quantile $\overline{q}$, the* bottom demotion algorithm *for x and $\overline{q}$ runs x on the agent with quantiles transformed as follows:*

- *For any $q \in [0, \overline{q}]$, return $q/\overline{q}$.*
- *For any $q \in [\overline{q}, 1]$, return 1.*

The interim allocation rule faced by an agent i after applying the extremal buffering algorithm to just i is the composition of the bottom demotion algorithm for quantile $\hat{q}_i^{T-k}$ composed with the top promotion algorithm for original allocation rule x_i and quantile $\hat{q}_i^k/\hat{q}_i^{T-k}$. Consequently, we may analyze the loss from applying these two transformations separately and multiply the losses.

We first analyze bottom demotion. While bottom demotion does not produce an allocation rule which is multiplicatively close to the original rule, it does produce one which is close in the sense that its inverse is close to the inverse of the original rule. For the objectives of both revenue and welfare, this notion of closeness also produces a multiplicative approximation for virtual surplus. We state this as a separate technical lemma, as we will make use of the idea multiple times. The main tool in this analysis is Lemma 11 from Section 4 which we restate here for convenience. Recall, a virtual value function is $\phi(\cdot)$ and has cumulative virtual curve $\Phi(q) = \int_0^q \phi(r)\, dr$.

Lemma 11. *For virtual value function $\phi(\cdot)$ and cumulative virtual value $\Phi(q) = \int_0^q \phi(r)\, dr$ satisfying $\Phi(\alpha\, q) \geq \alpha\, \Phi(q)$ for all quantiles q and $\alpha \in [0, 1]$, and any two allocation rules x_1 and x_2 that satisfy $x_1^{-1}(z) \geq x_2^{-1}(z) \geq \frac{1}{\alpha} x_1^{-1}(z)$, the virtual surpluses satisfy*

$$\mathbf{E}_q[\phi(q)\, x_2(q)] + \Phi(0)\, x_2(0) \geq \tfrac{1}{\alpha}[\mathbf{E}_q[\phi(q)\, x_1(q)] + \Phi(0)\, x_1(0)].$$

Lemma 34. *Let $\phi : [0, 1] \to \mathbb{R}$ be an arbitrary nonincreasing virtual value function. Given a monotone single-agent allocation rule x and quantile $\overline{q}$, the bottom demotion algorithm for x and $\overline{q}$ obtains at least a $\overline{q}$-fraction of the expected virtual surplus of x.*

Proof. The lemma will follow from a straightforward application of Lemma 11. For a quantile q receiving allocation $x(q)$ from the base algorithm, the quantile receiving this probability of allocation under the bottom demotion algorithm will be $\overline{q}q$. Hence, $x^{-1}(z) \geq \hat{x}^{-1}(z) = \overline{q}x^{-1}(z)$. Since ϕ is nonincreasing in q, we have that $R(q) = \int_0^q \phi(r)\, dr$ satisfies $R(\alpha q) \geq \alpha R(q)$ for all $\alpha \in [0, 1]$. Hence, Lemma 11 implies the desired result. □

We have shown that bottom demotion results in an allocation rule which has an inverse close to that of the original rule on which it is based. To derive an approximation result for the top promotion algorithm requires a more nuanced version of the same approach, based on two observations. First, the "unallocation rules", i.e., $y(q) = 1 - x(1-q)$ for allocation rule $x(q)$, satisfy the inverse-approximation condition of the lemma. Second, the virtual surplus of the unallocation rule is given by the expected virtual value plus the negative virtual surplus of the unallocation rule. Specifically $\mathbb{E}_q[\phi(q)\, x(q)] = \mathbb{E}_q[\phi(q)] + \mathbb{E}_q[(-\phi(1-q))\, y(q)]$. While virtual values for revenue always satisfy the property that rays from the origin cross the cumulative virtual value curve from below, this property does not generally hold for the negative virtual values $-\phi(1-q)$. Regularity, i.e., monotonicity of the original virtual value function, however, implies the property for negative virtual values. These observations are formally summarized in the subsequent lemma:

Lemma 35. *Let $\phi : [0,1] \to \mathbb{R}$ be an arbitrary nonincreasing virtual value function satisfying $\int_0^1 \phi(q)\, dq \geq 0$. Given a monotone single-agent allocation rule x and quantile $\underline{q}$, the top promotion algorithm for x and $\underline{q}$ obtains at least a $(1-\underline{q})$-fraction of the expected virtual surplus of x.*

Proof. Note that the expected virtual surplus from any allocation rule x is can be written as $\int_0^1 \phi(q)x(q)\, dq = \int_0^1 \phi(q)\, dq - \int_0^1 \phi(q)(1-x(q))\, dq$. Specifically, let $\hat{x}$ be the allocation rule of the top promotion algorithm, and x the allocation rule of the original algorithm. Moreover, define $\hat{y}(q) = 1 - \hat{x}(1-q)$ and $y(q) = 1 - x(1-q)$ to be the corresponding "unallocation rules." We will show that

$$\int_0^1 -\phi(1-q)\hat{y}(q)\, dq \geq (1-\underline{q}) \int_0^1 -\phi(1-q)y(q)\, dq. \tag{7}$$

Since, $\int_0^1 \phi(q)\, dq \geq 0$, this will prove that $\int_0^1 \phi(q)x(q)\, dq \geq (1-\underline{q}) \int_0^1 \phi(q)\hat{x}(q)\, dq$.

To prove (7), note that the definition of the top promotion algorithm can be manipulated to obtain $\hat{x}^{-1}(z) = x^{-1}(z)(1-\underline{q}) + \underline{q}$. Moreover, by the definition of y and $\hat{y}$, we have $y^{-1} = 1 - x^{-1}(1-z)$ and $\hat{y}^{-1} = 1 - \hat{x}^{-1}(1-z)$. Combining these three equations yields that $y^{-1}(z) \geq \hat{y}^{-1}(z) = (1-\underline{q})y^{-1}(z)$ for all $z \in [0,1]$. Moreover, note that $-\phi(1-q)$ is decreasing in q. This implies that $R(q)/q \geq -\phi(1-q)$, where $R(q) = \int_0^q -\phi(1-q)\, dq$. We may therefore apply Lemma 11, which yields (7). □

Combining Lemmas 34 and 35 yields the following lemma:

Lemma 36. *Let $\phi : [0,1] \to \mathbb{R}$ be an arbitrary nonincreasing virtual value function satisfying $\int_0^1 \phi(q)\, dq \geq 0$, and consider an arbitrary agent i. The q-buffering algorithm for $\mathbf{x}$ and $\hat{Q}$, when applied only to agent i, attains at least a $(1 - \frac{q}{1-q})(1-q)$-fraction of the expected virtual surplus for i.*

Having derived a single-agent guarantee, we now show that applying the q-buffering algorithm to all agents at once, rather than just to one agent, yields only a small additional loss. Intuitively, for each agent, the mechanism only appears different when another agent has an extreme quantile which is promoted or demoted by the buffering algorithm. The probability of such an event can be bounded using the union bound. Formally, we have:

Proof of Lemma 31. Lemma 36 states that the virtual surplus lost from applying the q-buffering procedure to a single agent is small. We now argue that applying the algorithm to all agents at once does not incur much additional loss. We argue from the perspective of some agent i.

The key observation in our analysis is that the distribution of the quantiles of other agents is nearly unchanged by the extremal buffering algorithm. In particular, note that the probability that one or more agents other than i with quantiles set to 0 or 1 by the extremal buffering algorithm is at most $(n-1)(1-2q)$, by the union bound. Conditioned on there being no such agents, the distribution of quantiles input to the allocation algorithm remains uniform. It follows that the virtual surplus from distribution i conditioned on this event is identical to the revenue from the extremal buffering algorithm applied only to i.

In the event that there are one or more agents from populations other than i who have top quantiles (which are promoted) or bottom quantiles (which are demoted), we note that the conditional virtual surplus from population i is nonnegative. To see this, let $\tilde{x}_i$ be the interim allocation rule for agent i in the q-buffering algorithm conditioned on the event $\mathcal{E}$ that at least one agent j other than i has a quantile in $[0,q] \cup [1-q,1]$. Since $\mathbf{x}$ is a monotone function of its inputs, it must be that $\tilde{x}_i$ is nondecreasing. The expected virtual surplus from agent i conditioned on $\mathcal{E}$ is $\int_0^1 \phi_i(q)\tilde{x}_i(q)\,dq$. By assumption, $\int_0^1 \phi_i(q)\,dq \geq 0$, so it must also be the case that $\int_0^1 \phi_i(q)\tilde{x}_i(q)\,dq \geq 0$.

To conclude the proof, let $\hat{x}_i$ denote the interim allocation rule of the extremal buffering algorithm conditioned on the event $\overline{\mathcal{E}}$. The total virtual surplus from agent i is:

$$\Pr(\mathcal{E}) \int_0^1 \phi_i(q)\tilde{x}_i(q)\,dq + \Pr(\overline{\mathcal{E}}) \int_0^1 \phi_i(q)\hat{x}_i(q)\,dq$$

By the union bound, $\Pr(\overline{\mathcal{E}}) = 1 - \Pr(\mathcal{E}) \geq 1 - (n-1)(1-2q)$. By Lemma 36 and the fact that, conditioned on $\overline{\mathcal{E}}$, the distribution of quantiles i perceives from other agents is uniform implies that

$$\begin{aligned}\int_0^1 \phi_i(q)\hat{x}_i(q)\,dq \quad &\geq (1-\tfrac{q}{1-q})(1-q)\textstyle\int_0^1 \phi_i(q)x_i(q)\,dq \\ &\geq (1-\tfrac{q}{1-q})(1-q)\textstyle\int_0^1 \phi_i(q)x_i(q)\,dq.\end{aligned}$$

Since we have shown that $\int_0^1 \phi_i(q)\hat{x}_i(q)\,dq \geq 0$, we can combine the above to obtain:

$$\begin{aligned}&\Pr(\mathcal{E}) \int_0^1 \phi_i(q)\tilde{x}_i(q)\,dq + \Pr(\overline{\mathcal{E}}) \int_0^1 \phi_i(q)\hat{x}_i(q)\,dq \\ \geq\ &(1-\frac{q}{1-q})(1-q)(1-2(n-1)q) \int_0^1 \phi_i(q)x_i(q)\,dq.\end{aligned}$$

Summing over agents proves the lemma. □

Since we have reasoned about abstract virtual surplus, which could be value or Myerson virtual value, we obtain revenue and welfare approximation results for the extremal buffering algorithm.

C.3.2 Approximately Optimal Resampling

We previously observed that naïvely resampling each agent's quantile from their bin could drastically reduce the welfare or revenue of an algorithm. Driving this loss were agents with extreme quantiles: if the base algorithm's welfare was driven primarily by allocating rare but high-valued agents while rejecting all other quantiles, it is very unlikely that the resampling procedure will give these high-valued agents priority.

Applying the k/T-buffering procedure for some k before resampling solves exactly this issue. After applying the k/T-buffering procedure, resampling does not change the way the mechanism treats the top k and bottom k bins, as agents in those bins are already treated identically. Consequently, any loss from resampling must occur in the central quantiles of the distribution. For decreasing virtual value functions, such loss cannot be too high. In what follows, we formally define our resampling procedure, and then prove its performance guarantee.

Definition 37. *The* resampling algorithm *for allocation algorithm* $\mathbf{x}$ *allocates according to the following randomized procedure:*

- *For each agent i:*
  - *Compute j such that i's quantile $q_i \in [(j-1)/T, j/T]$.*
  - *Resample a quantile uniformly from $[(j-1)/T, j/T]$.*
- *Runs $\mathbf{x}$ on the resampled quantiles.*

Lemma 38. *For each agent i, let ϕ_i be a nondecreasing virtual value function, and let $\mathbf{x}$ be an a monotone allocation algorithm. Further, for some $k \leq T/2$, let $\hat{\mathbf{x}}$ denote the k/T-buffering algorithm for $\mathbf{x}$. Then the resampling algorithm for $\hat{\mathbf{x}}$ obtains at least a $\frac{k}{k+1}(1-\frac{q}{1-q})(1-q)(1-2(n-1)q)$ the expected virtual surplus under $\mathbf{x}$, with $q = k/T$.*

Since the resampling algorithm is a surrogate binning mechanism, Lemma 29 immediately follows. To prove Lemma 38, we first rephrase a useful lemma from Roughgarden and Schrijvers (2016), which characterizes the relationship between the virtual surplus of resampling algorithms and their base algorithms. Informally, the result states that the revenue from bin-based resampling for an allocation rule $\mathbf{x}$ is the same as if $\mathbf{x}$ was run on the binned revenue curve for those same bins. Formally:

Lemma 39 (Roughgarden and Schrijvers, 2016)**.** *For every agent i, let ϕ_i be a virtual value function for each agent, with cumulative virtual surplus $R_i = \int_0^q \phi_i(r)\,dr$. For any allocation algorithm $\mathbf{x}$, let $\overline{\mathbf{x}}$ denote the resampling algorithm based on $\mathbf{x}$. Then for each agent i we have:*

$$\mathbb{E}[-\overline{x}_i'(q_i)R_i(q_i)] = \mathbb{E}[-\overline{x}_i'(q_i)\overline{R}_i(q_i)] = \mathbb{E}[-x_i'(q_i)\overline{R}_i(q_i)],$$

where $\overline{R}_i(q_i)$ is a piecewise linear approximation to R_i given by:

$$\overline{R}_i(q) = T(q - \tfrac{j-1}{T})R_i(\tfrac{j}{T}) + T(\tfrac{j}{T} - q)R_i(\tfrac{j-1}{T}) \text{ for } q \in [\tfrac{j-1}{T}, \tfrac{j}{T}].$$

Proof of Lemma 38. Let $\hat{\mathbf{x}}$ denote the allocation rule of the k/T-buffering algorithm, and $\overline{\mathbf{x}}$ that of the resampling algorithm. Define

$$\hat{R}_i(q_i) = \begin{cases} R_i(q_i) & \text{for } q_i \in [\frac{k}{T}, \frac{T-k}{T}] \\ \frac{Tq_i}{k} R_i(\frac{k}{T}) & \text{for } q_i \in [0, \frac{k}{T}] \\ \frac{T(1-q_i)}{k} R_i(\frac{T-k}{T}) & \text{for } q_i \in [\frac{T-k}{T}, 1]. \end{cases}$$

That is, $\hat{R}_i$ is equal to R_i except on $[0, \frac{k}{T}]$ and $[\frac{T-k}{T}, 1]$, where it is a linear interpolation between the values of the R_i at the endpoints of those intervals.

We will argue the following sequence of inequalities for each agent:

$$\begin{aligned}\mathbb{E}[-\overline{x}_i'(q_i)R_i(q_i)] &= \mathbb{E}[-\hat{x}_i'(q_i)\overline{R}_i(q_i)] \\ &\geq \tfrac{k}{k+1}\mathbb{E}[-\hat{x}_i'(q_i)\hat{R}_i(q_i)] \\ &= \tfrac{k}{k+1}\mathbb{E}[-\hat{x}_i'(q_i)R_i(q_i)].\end{aligned}$$

The first equality follows from Lemma 39. We will argue the second and third lines shortly. Since the first and last expressions in the above chain are the respective virtual surpluses of the resampling and buffering algorithms, respectively, the result will follow.

To see that $\overline{R}_i(q_i) \geq \frac{k}{k+1}\hat{R}_i(q_i)$, note that $\overline{R}_i(q_i) = \hat{R}_i(q_i)$ for all quantiles in $[0, \hat{q}_i^k] \cup [\hat{q}_i^{T-k}, 1]$. Otherwise, consider $q_i \in [\frac{j}{T}, \frac{j+1}{T}]$ for $j \in \{k, \ldots, T-k-1\}$. Assume without loss of generality that $R_i(\frac{j}{T}) \leq R_i(\frac{j+1}{T})$; a symmetric argument will apply to the case where $R_i(\frac{j}{T}) \geq R_i(\frac{j+1}{T})$. The concavity of R_i implies that for all $q \in [\frac{j}{T}, \frac{j+1}{T}]$, $R_i(q) \leq \frac{Tq}{j}R_i(\frac{j}{T})$. Moreover, note that $\overline{R}_i(q) \geq R_i(\frac{j}{T})$ for all $q \in [\frac{j}{T}, \frac{j+1}{T}]$ Since $\frac{Tq}{j} \leq \frac{k+1}{k}$, it follows that $\overline{R}_i(q_i) \geq \frac{k}{k+1}R_i(q_i) = \frac{k}{k+1}\hat{R}_i(q_i)$ for all $q_i \in [\frac{k}{T}, \frac{T-k}{T}]$, and $\overline{R}_i(q_i) = \hat{R}_i(q_i)$ elsewhere.

Finally, we argue that $\mathbb{E}[-\hat{x}_i'(q_i)\hat{R}_i(q_i)] = E[-\hat{x}_i'(q_i)R_i(q_i)]$. Note that since $\hat{R}_i(q_i) = R_i(q_i)$ for $q_i \in [\frac{k}{T}, \frac{T-k}{T}]$, it follows that

$$\int_{\frac{k}{T}}^{\frac{T-k}{T}} -\hat{x}_i(q_i)R_i(q_i)\,dq_i = \int_{\frac{k}{T}}^{\frac{T-k}{T}} -\hat{x}_i(q_i)\hat{R}_i(q_i)\,dq_i.$$

To prove the claim, notice that $\hat{x}_i'(q_i) = 0$ on $[0, \frac{k}{T}]$ and $[\frac{T-k}{T}, 1]$. Hence,

$$\int_0^{\frac{k}{T}} -\hat{x}_i'(q_i)R_i(q_i)\,dq_i = \int_0^{\frac{k}{T}} -\hat{x}_i'(q_i)\hat{R}_i(q_i)\,dq_i.$$

and

$$\int_{\frac{T-k}{T}}^1 -\hat{x}_i'(q_i)R_i(q_i)\,dq_i = \int_{\frac{T-k}{T}}^1 -\hat{x}_i'(q_i)\hat{R}_i(q_i)\,dq_i.$$

This proves the lemma. □

C.3.3 Proof of Theorem 10

All that remains to prove the approximation result is to compose our lemmas and select a value for the parameter k. First, note the binning algorithm's surrogate values are the same for the intervals $[0, \frac{1}{T}], \ldots, [\frac{k-1}{T}, \frac{k}{T}]$ and the intervals $[\frac{T-k}{T}, \frac{T-k+1}{T}], \ldots, [\frac{T-1}{T}, 1]$. It follows that welfare and revenue loss from applying Theorem 28 are $\rho(k)$ and $\eta(k, n)$, respectively. Moreover, by lemma 38, the resampling algorithm attains a $\frac{k}{k+1}(1 - \frac{k/T}{1-k/T})(1-k/T)(1-2(n-1)k/T)$-fraction of the revenue and welfare of $\hat{\mathbf{x}}$. Composing these lemmas and setting $k = (n/T)^{\frac{2}{3}} \log^{1/3}(n/T)$ for welfare and $k = (n/T)^{\frac{2}{5}} \log^{1/5}(n/T)$ for revenue yields the theorem.

D Sample Complexity in General Feasibility Environments

In this appendix, we show how to use polynomially many truthfully sampled values to estimate the revenue of the k-unit, T-buyer auction for all k from 1 to $T-1$ simultaneously. We assume the value distribution is regular, but may have unbounded support. Formally, we prove the following guarantee:

Theorem 40. *Let $R^* = \max_q R(q)$ be the* monopoly revenue*. For any $\delta, \gamma \in (0,1)$, $O(T^6\delta^{-4}\log(T/\gamma)$ samples suffice to estimate the expected revenue of a k-unit, T-bidder auction up to additive error δR^* for all $k \in \{1, \ldots, T-1\}$ simultaneously with probability at least $1-\gamma$.*

Combined with Proposition 12, Theorem 40 gives a sample complexity result as follows. Since $\sum_i R_i^*$ is at most n times the revenue of the optimal mechanism, the additive propagation of error guarantee of Theorem 15 also implies a multiplicative $(1-n\epsilon)$ guarantee, which can be composed with the multiplicative guarantee of Theorem 10. The result, restated from Section 5.3, is:

Theorem 19. *Consider an arbitrary feasibility environment. Assume agents have regular (possibly unbounded value distributions). Then polynomially many sampled value profiles and polynomial run-time suffice to estimate a mechanism which obtains at least a $(1-\epsilon)$ fraction of the revenue of the optimal mechanism with probability at least $1-\epsilon$.*

We first outline the high-level strategy for proving Theorem 40. First, for any $j \in \{0, \ldots, T\}$, let P_j denote the expected revenue of a j-unit auction with T agents, and let ψ^k denote the expected kth order statistic of the virtual value distribution. Then we may write: $\psi^k = P_k - P_{k-1}$. It follows that to estimate ψ^k with additive error δR^*, it suffices to estimate P_k and P_{k-1} with error $2\delta R^*$.

To estimate the k-unit revenue P_k, we will estimate the revenue contribution from a single agent, $\mathbb{E}_q[p(q)]$. Note that from an agent's perspective, playing in a k-unit auction is equivalent to facing a posted price distributed according to $q_{k:T-1}$), where $q_{k:T-1}$ denotes the kth lowest order statistic of $T-1$ $U[0,1]$ random variables. Note that $q_{k:T-1}$ is distributed according to $\text{Beta}(k, T-k)$. Let $f_{k:T-1}$ denote the density of $q_{k:T-1}$, and let $R = v(q)q$ denote the price-posting revenue curve. We have:

Lemma 41. *For any $k \in \{1, \ldots, T-1\}$, $P_k = T\int_0^1 f_{k:T-1}(q)R(q)\,dq$.*

In what follows, we will show how to estimate $R(q)$ for all $q \in [1/T^2, 1-1/T^2]$. We will further show that the loss from misestimating $R(q)$ on $[0, 1/T^2] \cup [1-1/T^2, 1]$ is minimal. This will immediately imply Theorem 40.

D.1 Estimation on a Grid

To create a skeleton for our estimated revenue curve, we first estimate $R(q)$ for $q \in \{1/K, \ldots, 1-1/K\}$, for some large K to be determined later. The concavity of R will imply that the rest of the revenue can be estimated with low error via interpolation.

Lemma 42. *Let $R^* = \max_q R(q)$. For any $1 > \delta > 0$ and γ, $O(K^2\delta^{-2}\log(K/\gamma))$ samples suffice to guarantee that $K\hat{v}_j/j \in [R(j/K) - \delta R^*, R(j/K) + \delta R^*]$ for all $j \in \{1, \ldots, K-1\}$ simultaneously with probability at least $1-\gamma$.*

Consider drawing $N = Km - 1$ samples, for some positive integer m. Note that the jmth smallest sample has mean j/K. Let $\hat{v}_j$ denote the value of this sample. We will use $\hat{v}_j$ as an estimator for $v(j/K)$, and $K\hat{v}_j/j$ as an estimator for $R(j/K) = Tv(j/K)/j$. The proof will proceed in two steps. First, we will use a Chernoff bound to show that with high probability, $q(\hat{v}_j)$ is close to j/K. We will then use the concavity of R to show that $\hat{v}_j$ is close to $v(j/K)$.

Lemma 43. *For any $1 > \delta > 0$ and γ, $Km = O(K\delta^{-2}\log(1/\gamma))$ samples suffice to guarantee that $q_j \in [(1-\delta)j/K, (1+\delta)j/K]$ with probability at least $1-\gamma$.*

Proof. We now bound the probability of significantly misestimating $q(\hat{v}_j)$. Let $q_j = q(\hat{v}_j)$. Note that for any $\delta \in (0,1)$, the number of samples with quantile at most $(1+\delta)j/K$ is the sum of N iid Bernoulli random variables with mean $(1+\delta)j/K$. Moreover, note that $q_j > (1+\delta)j/K$ only if at most $jm-1$ samples overall have quantile at most $(1+\delta)j/K$. Chernoff then gives us that

$$\Pr[q_j \geq (1+\delta)j/K] \leq e^{-\left(1-\frac{(jm-1)K}{(1+\delta)Nj}\right)^2 \frac{(1+\delta)Nj}{2K}} = e^{-\Omega(\delta^2 mj)}$$

Similarly, the the number of samples with quantile at most $(1-\delta)j/K$ is the sum of N iid Bernoullis with mean $(1-\delta)j/K$. We have $q_j < (1-\delta)j/K$ only if at least jm samples have quantile at most $(1-\delta)j/K$. Chernoff then gives us:

$$\Pr[q_j \leq (1-\delta)j/K] \leq e^{-\Omega(\delta^2 mj)}$$

It follows that as $m = O(\delta^{-2}j^{-1}\log(1/\gamma))$ suffices for suffices for $q_j \in [(1-\delta)j/K, (1+\delta)j/K]$ with probability at least $1-\gamma$. This bound is worst when $j=1$. Hence $mK = O(K\delta^{-2}\log(1/\gamma))$ samples suffice overall. □

We next show that if q_j is close to j/K, then $K\hat{v}_j/j$ will be a close estimate of $R(j/K)$.

Lemma 44. *Assume $q_j \in [(1-\delta)j/K, (1+\delta)j/K]$. Then $K\hat{v}_j/j \in [(1-\delta K)R(j/K), (1+\delta K)R(j/K)]$.*

Proof. We will show that $\hat{v}_j \in [(1-\delta K)v(j/K), (1+\delta K)v(j/K)]$. The result follows from multiplying by K/j. We first argue the case where $q_j < j/K$. That is, $\hat{v}_j \geq v(j/K)$. Concavity of the revenue curve implies that $R(q_j) \leq \frac{1-q_j}{1-j/K}R(j/K)$. In the event that $q_j \geq (1-\delta)j/K$, we have:

$$\hat{v}_j q_j \leq \frac{1-q_j}{1-j/K}\frac{v(j/K)j}{K} \leq (1+\delta K)v(j/K).$$

Dividing the inequality by q_j and using the fact that $q_j \geq (1-\delta)j/K$ yields:

$$\hat{v}_j \leq \frac{(1-(1-\delta)j/K)}{(1-j/K)(1-\delta)}v(j/K).$$

A symmetric argument applies when $q_j > j/K$. By concavity, we have $R(q_j) \geq \frac{1-q_j}{1-j/K}R(j/K)$. In the event that $q_j \leq (1+\delta)j/K$, we have

$$v_j \geq \frac{(1-(1+\delta)j/K)}{(1-j/K)(1+\delta)}v(j/K) \geq (1-\delta K)v(j/K).$$

□

Lemma 42 follows from the above results by applying the union bound.

Proof of Lemma 42. Lemmas 43 and 44 imply that $O(K^2\delta^{-2}\log(K/\delta))$ samples suffice to guarantee that $K\hat{v}_j/j \in [(1-\delta)R(j/K), (1+\delta)R(j/K)]$ with probability at least $1-\gamma/K$. Applying the union bound, it follows that $O(K^2\delta^{-2}\log(K/\delta))$ samples suffice to guarantee that this guarantee holds for all $j \in \{1,\ldots,K-1\}$ simultaneously with probability at least $1-\gamma$. Noting that $R(j/K) \leq R^*$ implies the lemma. □

D.2 Estimating The Interior Revenue Curve

In the previous section, we showed how to estimate $R(j/K)$ up to an additive δR^*. Pick some $q \in [(j-1)/K, j/K]$ with $j \in \{2, \ldots, K-1\}$. We may linearly interpolate between our estimates of $R((j-1)/K)$ and $R(j/K)$ to estimate $R(q)$. More formally, let $\hat{R}_j$ denote the estimator for $R(j/K)$ analyzed in the previous section. We will estimate $R(q)$ as

$$\hat{R}(q) = (q - \tfrac{j-1}{K})K\hat{R}_j + (\tfrac{j}{K} - q)K\hat{R}_{j-1}$$

We will use concavity of R to bound the error from this estimate.

Lemma 45. *Assume that for all $j \in \{1, \ldots, K-1\}$, $K\hat{v}_j/j \in [R(j/K) - \delta R^*, R(j/K) + \delta R^*]$ for all $j \in \{1, \ldots, K-1\}$. Then for all $q \in [1/K, 1-1/K]$:*

$$\hat{R}(q) \in [R(q) - (\delta + K^{-1})R^*, R(q) + \delta R^*]$$

Proof. We bound the overestimation and underestimation error in turn. First, not that since R is concave, it must be that $R(q)$ lies above the line between $((j-1)/K, \hat{R}_{j-1} - \delta R^*)$ and $(j/K, \hat{R}_j - \delta R^*)$. It follows that $\hat{R}(q)$ can only overestimate by at most δR^*. Next, note that a lower bound on $\hat{R}(q)$ is

$$\hat{R}(q) \geq (q - \tfrac{j-1}{K})KR(\tfrac{j-1}{K}) + (\tfrac{j}{K} - q)KR(\tfrac{j}{K}) - \delta R^*$$

In other words, the worst-case underestimation is δR^*, plus the worst-case underestimation of the piecewise linear curve through the points $(j/K, R(j/K))$ for all $j \in \{0, K\}$. Using the facts that R is concave, $R(0) = 0$, and $R(1) = 0$, this underestimation can be shown to be $O(R^*/K)$. The lemma follows.

□

D.3 Proof of Theorem 40

We have shown how to estimate the revenue curve to low additive error for quantiles bounded away from 0 and 1. We now use this estimator to prove the main sample complexity result of this appendix: that $O(T^6\delta^{-4}\log(T/\delta))$ samples suffice to estimate estimate ψ^k for all $k \in \{1, \ldots, T-1\}$ to within additive error δR^* with probability at least $1-\gamma$. To do so, we will consider estimating P_k for $k \in \{1, \ldots, T-1\}$ as $T\int_{\delta/T^2}^{1-\delta/T^2} f_{k:T-1}(q)\hat{R}(q)\,dq$.

First we show that ignoring the revenue contribution from the intervals $[0, \delta/T^2]$ and $[1 - \delta/T^2, 1]$ cannot hurt our estimate by much. Let $F_{k:T-1}$ denote the CDF of the kth lowest order statistic of $T-1$ $U[0,1]$ random variables. Then using properties of the Beta distribution, we have that $F_{k:T-1}(\delta/T^2) \leq \delta/T$ and $1 - F_{k:T-1}(1-\delta/T^2) \leq \delta/T$ for all $k \in \{1, \ldots, T-1\}$. Since for $q \in [0, \delta/T^2] \cup [1-\delta/T^2, 1]$, $R(q) \leq R^*$, it follows from Lemma 41 that

$$P_k - T\int_{\delta/T^2}^{1-\delta/T^2} f_{k:T-1}(q)R(q)\,dq \leq \delta R^*$$

Now assume that for all $q \in [\delta/T^2, 1-\delta/T^2]$, $|\hat{R}(q) - R(q)| \leq \delta T^{-1}R^*$. If this is the case, then we have

$$\left| T\int_{\delta/T^2}^{1-\delta/T^2} f_{k:T-1}(q)R(q)\,dq \leq \delta R^* - T\int_{\delta/T^2}^{1-\delta/T^2} f_{k:T-1}(q)\hat{R}(q)\,dq \leq \delta R^* \right| \leq \delta R^*.$$

Choosing $K = T^2/\delta$ in Lemma 45 yields the result.