THE SIMPLE EMPIRICS OF OPTIMAL ONLINE AUCTIONS

DOMINIC COEY, BRADLEY J. LARSEN, KANE SWEENEY, AND CAIO WAISMAN

ABSTRACT. We study reserve prices computed to maximize the expected profit of the seller based on historical observations of incomplete bid data typically available to the auction designer in online auctions for advertising or e-commerce. This direct approach to computing reserve prices circumvents the need to fully recover distributions of bidder valuations. We derive asymptotic results and also provide a new bound, based on the empirical Rademacher complexity, for the number of historical auction observations needed in order for revenue under the estimated reserve price to approximate revenue under the optimal reserve arbitrarily closely. We illustrate the approach with e-commerce auction data from eBay, where we examine empirically the benefit of the optimal reserve as well as the size of dataset required in practice to achieve that benefit. This simple approach to estimating reserves may be particularly useful for auction design in Big Data settings, where traditional empirical auctions methods may be costly to implement. We also demonstrate how this idea can be extended to estimate all objects necessary to implement the Myerson (1981) optimal auction.

Date: June 7, 2018.
We thank Isa Chaves, Han Hong, Dan Quint, Evan Storms, and Anthony Zhang for helpful comments. Coey and Sweeney were employees of eBay Research Labs while working on this project, and Larsen was a part-time contractor when the project was started.
Coey: Facebook, Core Data Science; coey@fb.com.
Larsen: Stanford University and NBER; bjlarsen@stanford.edu.
Sweeney: Uber, Marketplace Optimization; kane@uber.com.
Waisman: Stanford University; cwaisman@stanford.edu.

1. Introduction

Auctions are one of the main selling mechanisms used online. Not only are they employed to sell physical goods through sites such as eBay, but also to offer advertisement space online, such as through Google's DoubleClick. Display and banner ads have become a common tool to target consumers, which is conducted via real-time bidding, in which impressions are commonly sold via second-price, sealed-bid auctions that last a fraction of a second. It is estimated that over $20 billion dollars are sold via real-time bidding per year (eMarketer 2016), providing an abundance of data on bidding activity and reflecting the relevance and potential revenue gains from implementing the best auctions possible. Traditional approaches to empirical auctions would analyze such data by estimating underlying distributions of bidder valuations and then computing counterfactual mechanisms of interest, such as optimal reserve prices or other instruments of auction design. These approaches, however, can be computationally demanding, in particular in large data settings, and are unlikely to be implemented in real-time by online platforms interested in optimal auction design. We provide a more direct approach to obtaining the information required for optimal reserve prices in these online auctions. The primary virtue of the approach is that it is simple to execute—a practical approach to computing optimal reserve prices in certain environments without the effort necessary to fully estimate distributions of bidder valuations.

The online auction environments we are interested in studying typically follow a second-price-like format, where, unlike traditional ascending auctions, the highest bid *is* recorded. We show that this opens up a new and more direct approach to obtaining the information required for computing optimal reserve prices. Our starting point is the observation that seller profit is a simple-to-compute function of the reserve price and the two highest bids—and thus profit can be computed even with incomplete bidding data (i.e. where not all bids are observed) and without observing the number of bidders. Although it is true that, in theory, complex auction design may be optimal—such as the Myerson (1981) auction for asymmetric independent private values (IPV) settings—in practice, the seller's primary instrument of auction design in the real world is typically a single reserve price. We demonstrate that the optimal single reserve price can be computed using only historical observations of data on the two highest bids by simply maximizing the seller profit function with respect to the reserve price.

We derive several properties of estimated reserve prices. We prove consistency, demonstrate that the asymptotic distribution of the estimator of the optimal reserve price is non-normal, and derive its asymptotic distribution. Our interest in deriving these results is not theoretical but is instead practical: we wish to provide a clear notion of *how many previous transactions* the practicioneer needs to observe in order for it be the case that designing an auction based on estimated reserves is a good idea. In this spirit, we then provide a type of result that is relatively new to the empirical auctions literature: we derive an explicit lower bound, based on the empirical Rademacher complexity, for the number of auctions one would need to observe in order to guarantee that the revenue based on the estimated reserve price approximates the true optimal revenue arbitrarily closely. We then take a step beyond asymptotic and learning theory and examine, empirically, how many auctions one needs to observe in practice in order for it to be the case that the estimated optimal reserve price based on that data sample indeed yields more revenue than a no-reserve auction (an auction with the reserve price simply fixed at the seller's valuation).

We apply the approach to a sample of popular smartphone products sold through eBay auctions. We find that implementing the optimal reserve price in these settings would raise expected profit between 0.11–0.86% compared to an auction with a reserve price equal to the seller's value, with the precise benefit depending on the product. In the data we find that a history of at least 7–25 auctions is required in order for an auction using the estimated reserve price to outperform an auction with the reserve price set to the seller's value. These numbers are based on the median number of auctions (across many simulated draws) required for the estimated optimal reserve price to perform well. We also report the 95^{th} quantile dataset size, which ranges from 42–176. This suggests that, for these products, collecting a dataset of 176 previous auctions and estimating the optimal reserve price based on that data will, with at least 95% probability, yield a reserve price that will outperform an auction where the reserve is set to the seller's valuation. In online auctions for advertising or e-commerce such data requirements are not likely to be restrictive.

The computational simplicity of this approach can be particularly attractive given the recent advent of Big Data. Specifically, auctions for online advertising or e-commerce often contain too many records to be feasibly analyzed using traditional structural methods, which would typically involve 1) restricting to a sample of data in a flat file which can be analyzed in standard statistical software; 2) fully recovering the underlying distribution of buyer valuations through maximum likelihood estimation or other approaches, which often require searching over multi-dimensional parameter spaces; and 3) computing optimal reserve prices (e.g. Paarsch 1997; Bajari and Hortaçsu 2003). In contrast, our approach requires only searching over a single dimension at once and can be performed with a simple grid search, raising the possibility of computing optimal reserve prices directly on data storage platforms (e.g. Teradata, Hadoop, Spark) without bringing the data into more traditional analysis packages (e.g. Matlab, R, etc.).

We also demonstrate an interesting extension of our methodology to the setting of the Myerson (1981) optimal auction. If, in addition to the two highest bids, the highest bidder's identity is also known, we show how to recover bidder-specific marginal revenue curves (as defined by Bulow and Roberts 1989) in asymmetric IPV settings, enabling the Myerson (1981) optimal auction to be implemented. We use simulated data to illustrate this approach and quantify the revenue gain from optimal auction design. While the Myerson (1981) auction is not widely used in practice, auctions with "soft floors" have been used explicitly to deal with potential asymmetry between bidders.[1] However, as Zeithammer (2017) demonstrated, this auction format can actually be detrimental to sellers, in which case implementing the Myerson (1981) auction might be a preferred alternative. When there are concerns about discriminating bidders, recovering marginal revenue curves allows one to implement the symmetric auction proposed by Deb and Pai (2017), which yields the same expected revenue to the seller as the original Myerson (1981) auction.

All of our results allow for the number of bidders in each auction to be unknown, a data complication which has precluded nearly all previous empirical auction methods. When the number of bidders is unobserved, approaches relying on inverting order statistic distributions (Haile and Tamer 2003; Athey and Haile 2007; Aradillas-López et al. 2013) cannot be applied. This has particularly been an issue for online auctions, where the number of bidders is often unobserved to the econometrician.[2] Exceptions in the literature allowing for an unobserved number of bidders in second-price-like/ascending auction cases include Song (2004), Kim and Lee (2014), Platt (2017), and Freyberger and Larsen (2017), which each focus on symmetric independent (or conditionally independent) private values settings. Relative to our method, these approaches have the advantage of yielding estimates of the underlying valuation distributions, but at a higher computational cost. Our method circumvents the need for estimating valuation distributions and instead directly estimates the optimal reserve price. Also, it is important to emphasize that the specific information environments and data requirements we consider (asymmetric, possibly correlated, private values where the number of bidders is unobserved to the econometrician and bidding data is incomplete) are settings for

[1]A soft floor is a threshold that effectively determines the auction format. If the highest bid is below it, the winner pays her own bid, thereby characterizing a first-price auction. On the other hand, if at least one bid is above the soft floor, the wininng bidder's payment will be equal to the maximum of the soft floor and the second highest bid, which constitutes a second-price auction. Reserve prices in soft floor auctions are also known as "hard floors".

[2]For example, on eBay, if a bidder intends to bid in an auction but sees that the price has already exceeded her valuation, her bid will not be recorded, leading to a situation in which the number of bidders placing bids differs from the true number of bidders.

which no identification results exist in the auction literature for obtaining distributions of valuations even if that were the desired object of interest.

Our approach is related to a theoretical literature at the intersection of economics and computer science that examines *approximately optimal* auctions in a variety of cases, as surveyed by Roughgarden (2014). In particular, several studies in this literature employed the same tools from statistical learning theory we use to derive our finite sample bound on the difference between the optimal expected revenue and the one accrued from an auction using an estimated reserve price. For example, Cole and Roughgarden (2014) derive the number of auctions needed to approximate the optimal expected revenue as a function the number of bidders under the asymmetric IPV setting of Myerson (1981); under the same paradigm, but assuming that the number of bidders is either observed or drawn from a known distribution, Cesa-Bianchi et al. (2015) introduce a regret minimization algorithm to choose the reserve price; Alaei et al. (2013) extend the environment to allow for richer bidder preferences, including multi-dimensional types; Morgenstern and Roughgarden (2015) allow for several units being auctioned at the same time; Roughgarden and Wang (2016) consider the problem of choosing bidder-specific reserve prices; Kanoria and Nazerzadeh (2017) consider a dynamic environment of repeated auctions and how strategic bidder account for a seller using their past bids to learn the distribution of valuation; Jun and Pinkse (2017) utilize tools from information theory to obtain the optimal reserve price from bidding data under a symmetric IPV setting; and Balcan et al. (2018) extend the analysis to mechanisms other than auctions. Unlike these papers, we do not require bidder valuations to be independent, nor do we impose any restrictions on the number of bidders or the process behind which it is drawn. On the other hand, we do not consider regret minimization, implementation algorithms, multiple units or reserves, dynamics, or more general bidder preferences. Within this literature, the paper to which ours is most closely related is Mohri and Medina (2016), who introduced several results we leverage in this study. We reinterpret their learning framework as a sample size-dependent implementation decision from the vantage point of the seller.

Our paper also contributes to a literature in economics that provides direct inference about reserve prices or other objects of interest to auction design rather than attempting to estimate the full distribution of valuations, such as Li et al. (2003), Haile and Tamer (2003), Tang (2011), Aradillas-López et al. (2013), Coey et al. (2014), Chawla et al. (2014), and Coey et al. (2017). We also relate to other work empirical focusing specifically on online e-commerce auctions, such as Song (2004) or Platt (2017), and online ad auctions. In this paper, we consider only private values settings; in theory work, Abraham et al. (2016) model ad auctions with common values. In extending our approach to address the Myerson auction empirically, we also relate to Celis et al. (2014), which addresses non-regular distributions and Myerson's ironing in ad auctions, while we focus only on regular distributions. Finally, our work is related to the work of Ostrovsky and Schwarz (2016), where the authors experimentally varied reserve prices in position ad auctions to measure the improvement in profits from choosing different reserve prices.

2. Optimal Reserve Prices

We consider private value, second-price auctions. Values are independent and identically distributed across auctions, but within auctions we allow for arbitrary correlation in bidders' values and asymmetric valuation distributions among bidders. Let N be a random variable representing the number of potential bidders for a given auction. Let $V^{(k)}$ represent the k^{th} highest value among these N bidders.[3]

As in the entry model of Samuelson (1985), in a given auction each potential entrant observes the realization of her value and enters if her value exceeds the reserve price r. We assume that bidders do not bid above their values, and that if either of the two potential entrants with the highest values enter, they bid their

[3] As in Song (2004), we use this notation, rather than the traditional notation $V^{n:n}$, $V^{n-1:n}$, etc., because in our case the order statistics come from samples of varying sizes. That is, $V^{(k)}$ is the k^{th} highest bid among N bidders, unconditional on the realization of the random variable N, and is thus a draw from the distribution

$$F_{V^{(k)}}(v) \equiv \sum_{n} \Pr(N = n) F_{V^{n-k+1:n}}(v) \tag{1}$$

where $F_{V^{n-k+1:n}}$ is the distribution of the k^{th} highest bid conditional on n, the realization of N.

values.[4] This implies that the two highest bids, if they exist, are $V^{(1)}$ and $V^{(2)}$.[5] We will not require that the econometrician observe the number of potential or actual entrants.

Each potential entrant i has a type $t \in \{1, \ldots, T\}$. In the most general case, each bidder has his own type. All potential entrants with type t have values that are continuously distributed on some finite interval $[0, \omega_t]$, with $\bar{\omega} = \max_t \omega_t$. Let F_t denote this distribution, and f_t the corresponding density. We use F and f to denote the joint distribution and density of all potential entrants' values.

Let $j \in \{1, ..., J\}$ index a particular auction, and $V_j^{(1)}$ and $V_j^{(2)}$ represent the two highest bids in auction j. We will follow the convention and denote random variables with upper case letters and realizations of random variables with lower case letters. Let $v_0 \geqslant 0$ represent the seller's fixed value from keeping the good.

We assume the researcher observes the top two bids from auctions without a reserve price. The seller's revenue from setting a reserve price r in auction j is given by the following:[6]

$$\pi(V_j^{(1)}, V_j^{(2)}, v_0, r) = r\mathbb{1}(V_j^{(2)} < r \leqslant V_j^{(1)}) + V_j^{(2)}\mathbb{1}(r \leqslant V_j^{(2)}) + v_0\mathbb{1}(r > V_j^{(1)}). \quad (2)$$

Define expected profits as a function of the reserve as

$$p(r) \equiv \mathbb{E}\left[\pi(V_j^{(1)}, V_j^{(2)}, v_0, r)\right] \quad (3)$$

where we suppress dependence of p on v_0 for notational simplicity. We use the term *optimal reserve* to refer to a reserve price maximizing $p(r)$, and denote such a reserve by r^*. Given realizations of $V^{(1)}$ and $V^{(2)}$, one can obtain the optimal reserve price simply by maximizing (3), which we state as the following observation:

[4] This is an implication of the dynamic bidding model in Song (2004). Athey and Haile (2002) also argue, "...for many ascending auctions, a plausible ... hypothesis is that bids $V^{n-2:n}$ and below do not always reflect the full willingness to pay of losing bidders, although $V^{n-1:n}$ does (since only two bidders are active when that bid is placed)."

[5] This may not be true for other entry models. In Levin and Smith (1994), where bidders do not observe their values before entering, the two highest value bidders need not enter even if their values exceed the reserve price.

[6] Throughout, we use the terms *revenue* and *profit* interchangeably.

Observation 1. *For a given v_0, the seller's optimal reserve price is identified from observations of $V_j^{(1)}$ and $V_j^{(2)}$ from auctions without a reserve price.*

Importantly, this observation does not require knowledge of N nor of any bids lower than $V_j^{(1)}$ and $V_j^{(2)}$. This result holds true even if bidders are asymmetric and have correlated values. We estimate profits and the optimal reserve price as follows:

$$\hat{p}(r) \equiv \frac{1}{J} \sum_{j=1}^{J} \left[\pi(V_j^{(1)}, V_j^{(2)}, v_0, r) \right]. \tag{4}$$

$$\hat{r}_J \equiv \arg\max_r \hat{p}(r). \tag{5}$$

Remarks. Note that auctions with only one bidder—in which case, a realization of $V^{(2)}$ will be missing—should not be discarded from the sample for estimation of the optimal reserve price. Rather, these observations should be included as having a realization of zero for the missing bid, as zero is precisely the contribution such a bid would make to the seller's profit in such cases. Auctions with zero bidders can be ignored, as the seller receives zero profit from these auctions no matter what reserve is chosen.

It is important to note that our approach implicitly assumes that historical realizations of $V_j^{(1)}$ and $V_j^{(2)}$ are representative of future realizations; that is, we assume that these distributions are stable over time. In practice, these distributions may change, for example, due to changes in the underlying demand for the product. Our approach requires observations of auctions with no reserve price in order to correctly estimate the reserve price; once the estimated reserve price is put into practice, the practitioner will no longer observe the full distribution of second highest bids (as the reserve price will bind in some auctions). In practice, therefore, the practitioner may find it useful to periodically remove or reduce reserve prices in order to obtain anew data on second and first highest bids and recompute the optimal reserve price. This issue is discussed in a different context in Chawla et al. (2014), who proposed an alternative approach to auction design rather than relying on reserve prices.

On par with the rest of the empirical auctions literature, we do not model how distributions of bids could change due to bidders strategically accounting for how these bids will be used to set reserve prices in the future. As pointed out by Ostrovsky and Schwarz (2016), these dynamic concerns disappear as the market grows large.

In our empirical application, we highlight that the loss from using a non-optimal reserve price is asymmetric, in that the loss from setting too high a reserve is much larger than the loss from setting too low a reserve (see also Kim 2013 and Ostrovsky and Schwarz 2016). The practitioner may find it useful, therefore, to scale down estimated reserves by some value less than one, and then increase this scaling factor as number of observed auctions grows.[7] We do not consider this scaling approach here but believe this would be a promising modification to directly implementing estimated reserve prices.

The reserve prices we compute do not condition explicitly on auction-level heterogeneity. This can be considered either as a direct focus on homogeneous goods, where reserve prices can be estimated separately for each distinct good (as in our empirical application), or instead as a practical approach to computing reserve prices in cases where auction designers know that auction-level heterogeneity is present but cannot condition on it, and are restricted to choosing instead a single, unconditional reserve price. An example of such a setting are some display ad auction platforms that, either for technological reasons or for purposes of keeping the platform easy to understand for participants, set only a single reserve price for all ads sold in a given timeframe. Our approach can be extended to account explicitly for auction-level heterogeneity, conditioning the estimation of reserve prices on a vector of observed auction-level covariates, W_j. For example, one could use kernel smoothing to estimate the expectation in (3) conditional on W_j, use these

[7]The field experiments of Ostrovsky and Schwarz (2016) were subject to a scaling factor enforced by the auction platform because the platform feared setting too high of a reserve.

estimates to construct the objective function in (4) and then proceed as before. The underlying assumption would be that, conditional on W_j, the observed auctions are independently and identically distributed.

3. Properties of Estimated Reserves and Revenue

In this section we discuss the properties of estimated reserve prices and revenue. We begin by providing traditional asymptotic results for the estimator and discussing inference. We then provide a new bound on the number of auctions that must be observed in order for the auction revenue using the estimated reserve price to approximate the true optimal revenue arbitrarily closely. All proofs are found in the Appendix.

3.1. **Asymptotic Properties.**

We now present asymptotic results for the estimators $\hat{r}_J$ and $\hat{p}(\hat{r}_J)$. Standard M-estimation theory establishes the consistency of $\hat{r}_J$ for r^*. We apply Theorem 2.1 from Newey and McFadden (1994), whose conditions we show are satisfied.

Assumption 1. *$p(r)$ is continuous, with $p(r)$ uniquely maximized at r^*, where r^* belongs to a compact set.*

Assumption 1 is common in the econometrics literature. Note that the assumption that r belongs to a compact set is natural in our setting given that we model valuations as belonging to a bounded interval.

Theorem 1. *If Assumption 1 is satisfied, then $\hat{p}(\hat{r}_J) \xrightarrow{p} p(r^*)$ and $\hat{r}_J \xrightarrow{p} r^*$.*

The proof, given in the Appendix, consists of showing that, under Assumption 1, $\sup_{r\in[0,\overline{\omega}]} |\hat{p}(r) - p(r)| \xrightarrow{p} 0$. Hence, all the requirements from Theorem 2.1 of Newey and McFadden (1994) are satisfied, which yields Theorem 1. Having established consistency, we now derive the asymptotic distribution of the estimator $\hat{r}_J$, which is not standard. This ultimately implies that this estimator belongs to a class of estimators that converge at a cube-root rate, of which an example is the maximum score estimator proposed by Manski (1975). To demonstrate this and derive the asymptotic distribution, we show that the conditions in the main theorem of Kim and Pollard (1990), adapted below, are satisfied.

We first introduce the following notation and assumptions. Let $\tilde{\pi}(\cdot, r) \equiv \pi(\cdot, r) - \pi(\cdot, r^*)$. Let $P_R(\cdot)$ be defined as the supremum of $|\tilde{\pi}(\cdot, r)|$ over the class $\mathcal{P}_R \equiv \{\tilde{\pi}(\cdot, r) : |r - r^*| \leqslant R\}$.

Assumption 2. *The classes $\mathcal{P}_R$, for R near 0, are uniformly manageable for the envelopes P_R, and $\mathbb{E}[\tilde{\pi}(\cdot, r)]$ is twice differentiable with second derivative $-\Sigma$ at r^*.*

Theorem 2. *If Assumptions 1 and 2 are satisfied, and r^* is an interior point, then the process $J^{2/3}\frac{1}{J}\sum_{j=1}^{J} \tilde{\pi}(\cdot, r^* + \alpha J^{-1/3})$ converges in distribution to a Gaussian process $Z(\alpha)$ with continuous sample paths, expected value $-\frac{1}{2}\alpha^2\Sigma$, and covariance kernel H, where $H(\beta, \alpha) = \lim_{t\to 0} \frac{1}{t}\mathbb{E}[\tilde{\pi}(\cdot, r^* + \beta t)\tilde{\pi}(\cdot, r^* + \alpha t)]$ for any β, α in $\mathbb{R}$. Furthermore, if Z has nondegenerate increments, then $J^{1/3}(\hat{r}_J - r^*)$ converges in distribution to the random maximizer of Z.*

The slow rate of convergence of $\hat{r}_J$ (cube-root rate) suggests a strong data requirement to obtain a precise estimate of r^*. In other words, the econometrician must observe many previous auctions in order to accurately estimate the optimal reserve price. In Section 3.2 we provide an explicit bound on this data requirement and discuss how it relates to Theorem 2. We also examine the practical relevance of these theoretical results in our application in Section 4.

While estimating the optimal reserve price itself may in theory require a large amount of data, determining *how much the auction designer could gain* from optimally choosing the reserve price does not. In fact, $\hat{p}(\hat{r}_J)$ converges at a *square-root* rate to a normal distribution, which we present in the following theorem:

Theorem 3. *If the conditions from the previous theorems are satisfied, then*
$$\sqrt{J}\left[\hat{p}(\hat{r}_J) - p(r^*)\right] \xrightarrow{d} N\left(0, \mathbb{E}[\tilde{\pi}(\cdot, r^*)^2]\right).$$

To conclude our discussion of traditional asymptotic results, we briefly discuss inference. Simulating the asymptotic distribution of $\hat{r}_J$ is impractical as Σ depends upon the distributions of the two order statistics used to estimate r^*, which motivates the use of resampling methods. Abrevaya and Huang (2005) showed that the nonparametric bootstrap is not valid for this cube-root class of estimators. Alternative resampling methods that may be used in this case include subsampling (Delgado et al. 2001), m out of n bootstrap (Lee and Pun 2006), numerical bootstrap (Hong and Li 2017), and rescaled bootstrap (Cattaneo et al. 2017). Another possibility is to replace the indicator in the objective function with a smoothed estimator, which, along with further assumptions, might restore asymptotic normality and achieve faster rates of convergence, akin to the Horowitz (1992) smoothed maximum score estimator. We leave this possibility as an avenue for future research.

However, the nonparametric bootstrap can be used for inference on the object $\hat{p}(\hat{r}_J)$. Let $\hat{p}^b(r)$ be the objective function defined above calculated from a bootstrap sample of size J drawn with replacement from the original sample, and let $\hat{r}_J^b$ be the estimator calculated from this bootstrap sample. Results in Abrevaya and Huang (2005) yield $\hat{p}^b(\hat{r}_J^b) - \hat{p}^b(r^*) = O_P(J^{-2/3})$ and $\hat{r}_J^b - r^* = O_P(J^{-1/3})$, which imply that for the bootstrap analog

$$\sqrt{J}\left[\hat{p}^b(\hat{r}_J^b) - \hat{p}(\hat{r}_J)\right] = \sqrt{J}\left[\hat{p}^b(r^*) - \hat{p}(r^*)\right] + o_P(1),$$

which has the same limiting distribution as $\sqrt{J}\left[\hat{p}(\hat{r}_J) - p(r^*)\right]$ conditional on the data due to a standard result from bootstrap theory.

3.2. **Bound on Revenue Performance.** We now provide a new theoretical argument establishing a probabilistic upper bound on the quantity $p(r^*) - p(\widehat{r}_J)$, which shrinks to zero at the rate $(\log J/J)^{-1/2}$. This argument is closely related to work by Mohri and Medina (2016), although the bounds Mohri and Medina (2016) derive are less explicit than ours.[8]

[8]Mohri and Medina's (2016) bounds are stated in terms of the Rademacher complexity and pseudo-dimension of the class of possible reserve price functions mapping observables to reserves. The authors consider reserve prices that may be a function of observable covariates, which we do not do here.

Given a sample of J auctions, we refer to the seller's reserve price as being the *estimated* reserve price if the seller chooses the reserve that maximizes profit on past data, that is, in auction $J+1$ they choose the reserve price $\widehat{r}_J = \arg\max_r \widehat{p}(r)$. In expectation (over the possible bids in the $J+1^{th}$ auction), this gives a profit of $p(\widehat{r}_J)$, which, by definition, is lower than the expected profit given by the optimal reserve price, $p(r^*)$. We study the size of this difference, and how it changes as the number of observed auctions becomes large.

We state our bound in the following theorem. Its proof uses techniques from statistical learning to probabilistically bound the difference between $\widehat{p}(r)$ and $p(r)$ uniformly in r. Specifically, the *empirical Rademacher complexity* (defined in the Appendix) plays a key role in obtaining this bound. These techniques are developed in Koltchinskii (2001) and Koltchinskii and Panchenko (2002); Mohri et al. (2012) provide a textbook overview.

Theorem 4. *Let $0 \leqslant V_j^{(1)} \leqslant \overline{\omega} < \infty$. For any $\delta > 0$, with probability at least $1-\delta$ over the possible realizations of the J auctions, it holds that*

$$p(r^*) - p(\widehat{r}_J) \leqslant \overline{\omega}\left(\frac{8\sqrt{\log 2}}{J} + 4\sqrt{\frac{2+2\log J}{J}} + 6\sqrt{\frac{\log\frac{4}{\delta}}{2J}}\right).$$

Equivalently, in order for the profit from the estimated reserve to be within ε of the profit from the optimal reserve with probability at least $1-\delta$, the seller needs to have observed approximately J auctions such that the expression in the right-hand side of the expression in Theorem 4 equals ε. It can be shown that when $J = 1$ the expression is positive and that it is strictly decreasing in J, which, in principle, enables one to obtain the desired J for each (δ, ε) via a simple bisection procedure. The only assumption made on the first order statistic of bidders' values in this theorem is that it is non-negative and bounded above. Bidders' values may be arbitrarily correlated with each other and may be drawn from asymmetric marginal distributions. Note also that the right-hand side only depends on the value distribution through its upper bound $\overline{\omega}$, and not, for example, its variance. It is also important to note that the bound in Theorem 4 relates to the optimal reserve,

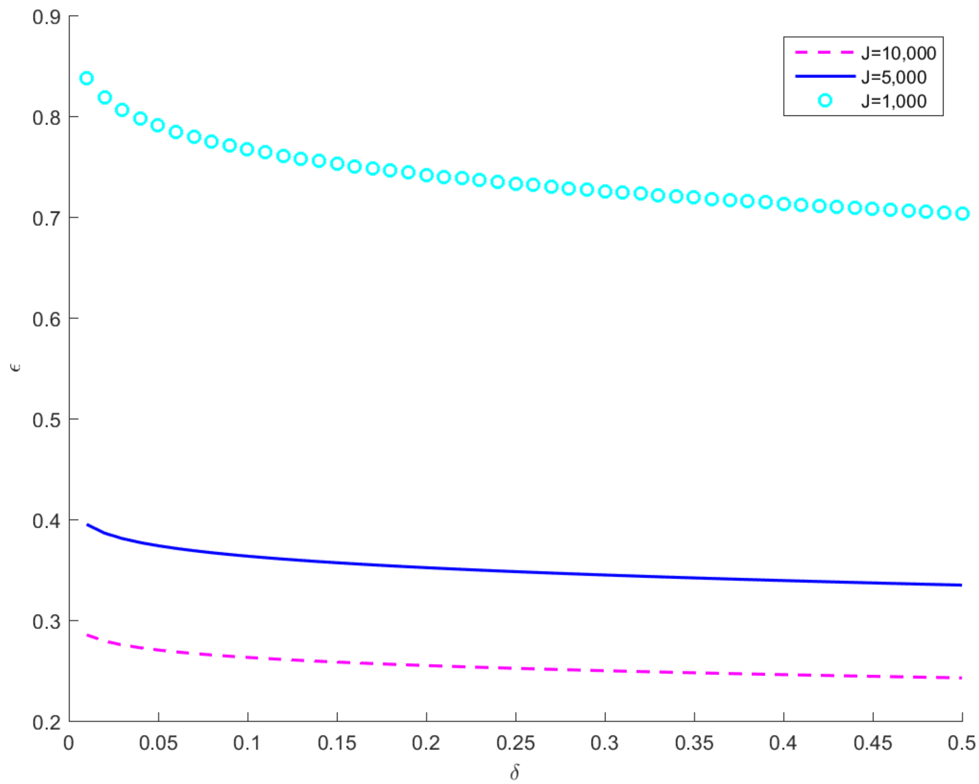


FIGURE 1. Iso-data curves in (δ, ε) space

Notes: Figure displays combinations of (δ, ε) that can be achieved given a fixed sample size J using the bound implied by Theorem 4. Top line represents $J = 1,000$, middle line represents $J = 5,000$, and bottom line represents $J = 10,000$.

r^*, and the true expected profit function, $p(\cdot)$, without requiring these objects to be known.

We illustrate the implications of Theorem 4 in Figure 1 below. For this illustration, we normalize $\overline{\omega} = 1$, and thus revenue is in units of fractions of the maximum willingness to pay. We then plot "iso-data" curves in (δ, ε) space, where each curve represents the possible combinations of ε and δ that are possible given a fixed history of observed auctions. In this figure, a curve located further to the southwest is preferable, as it represents a closer approximation to the true optimal revenue (i.e. a smaller ε) with a higher probability (i.e. a lower δ). The top line represents a sample size of $J = 1,000$, the middle line represents $J = 5,000$, and the bottom line represents $J = 10,000$. The middle line suggests that with a history of 5,000 auction realizations, one could guarantee a payoff within 0.348 (units of the maximum willingness to pay) of the optimal profit with probability 0.975; or, with the same size history, one could guarantee a payoff within 0.3447 of the optimal profit with probability 0.70. The larger sample, $J = 10,000$, can guarantee a payoff that is much closer to the optimal profit. Each iso-data curve is relatively flat in the δ dimension, reflecting the fact that the sample size requirements are more stringent for achieving a given level of ε closeness to the optimal profit, and are less stringent for achieving an improvement in δ (i.e. in the probability with which the revenue is reached).

We now relate the explicit bound obtained in this subsection to the asymptotic results obtained in the previous subsection. Theorem 2 implies that $p(\hat{r}_J) - p(r^*) = O_p(J^{-2/3})$.[9] By definition, therefore, for any $\delta > 0$, there exists an $M > 0$ such that $\Pr(J^{2/3}|p(\hat{r}_J) - p(r^*)| < M) \geqslant 1 - \delta$. The fact that the convergence result in Theorem 2 is achieved at a $J^{-2/3}$ rate implies that the bound in Theorem 4 is conservative, as it is expressed as a function of $(\log J/J)^{-1/2}$. However, Theorem 2 does not allow one to explicitly compute the number of auctions required in order to approach the optimal revenue with a given probability; it simply states that such an M exists. The advantage of the bound in Theorem 4, on the other hand, is that it is explicit, allowing one to directly compute a conservative estimate of the number of auctions J which must be observed for estimated reserve prices to perform well without requiring knowledge of $p(\cdot)$ or r^*.

In our application in Section 4, we take a step beyond asymptotic and learning theory and demonstrate that each of these theoretical results may be quite conservative in practice, as we find that, even with a very small dataset of previously observed auctions, revenue based on the estimated optimal reserve price can exceed that of revenue in an auction where the reserve price is ignored (and simply set to the seller's valuation).

[9] Performing a second-order Taylor expansion yields $p(\hat{r}_J) - p(r^*) = p''(\tilde{r})(\hat{r}_J - r^*)^2 = O_p(J^{-2/3})$, where $\tilde{r}$ is an intermediate value between $\hat{r}_J$ and r^*.

4. Computing Optimal Reserve Prices in E-commerce Auctions

We apply our methodology to a dataset of eBay auctions selling commodity products, which we define as those products which are cataloged in one of several commercially available product catalogs. Examples of commodity products include "Microsoft Xbox One, 500 GB Black Console", "Chanel No.5 3.4oz Women's Eau de Parfum", and "The Sopranos - The Complete Series (DVD, 2009)". We will refer to each distinct product as a "product" or "product-category." Within each product, the items sold are relatively homogeneous. For this exercise, we select popular smartphone products listed through auctions in 2013. We consider only auctions with no reserve price; specifically, we only include auctions for which the start price was less than or equal to \$0.99, the default start price recommendation on eBay. We omit auctions in which the highest bid is in the top 1% of all highest bids for that product and limit to products that are auctioned at least 500 times in our sample.

Table 1 shows summary statistics at the product level. There are 18 distinct products in our sample, and each product category contains 961 auctions on average, with the number of auctions ranging from 516 to 2,525. As reserve prices increase revenue only when they lie between the highest and second highest bids, the size of the gap between these bids is of particular interest. This gap ranges from \$8.41 (8.7% of the mean second highest bid) for product 2 to \$42 for product 18 (17.7% of the mean second highest bid). This gap suggests that reserve prices may be able to increase bids somewhat over a no-reserve auction.

We estimate optimal reserves separately in each product category. We do this exercise separately for the case where the seller's outside option (v_0) for the good is set to zero and where v_0 is set to half of the average second highest bid. Setting $v_0 = 0$ will yield the reserve price that will maximize the expected payment of the winning bidder to the seller ignoring the seller's outside option, which may be the quantity that the auction platform is most interested in maximizing, as platform fees are typically proportional to this payment. Setting v_0 to some fraction of

TABLE 1. Product-Level Descriptive Statistics

		Highest Bid		Second Highest Bid	
Product	# Obs	Average($)	Std Dev ($)	Average ($)	Std Dev ($)
1	805	78.82	46.83	67.73	33.59
2	657	55.82	25.15	47.41	23.23
3	1309	239.02	108.41	219.80	101.87
4	594	211.78	123.20	185.56	81.98
5	877	235.99	93.59	213.46	86.74
6	1286	257.49	121.87	234.76	107.28
7	728	215.04	92.21	190.06	80.39
8	839	230.17	100.72	205.38	90.69
9	2525	199.06	92.55	181.10	76.90
10	2465	139.26	58.48	125.56	54.69
11	712	143.11	64.07	128.22	55.68
12	733	208.91	93.24	188.25	76.20
13	570	159.53	62.46	142.29	59.40
14	680	172.18	107.12	149.91	73.52
15	733	120.28	60.35	105.26	54.31
16	572	121.25	65.54	104.10	56.63
17	701	127.28	59.13	114.64	55.62
18	516	350.80	156.78	308.68	138.60

Notes: Table displays, for each product, the number of auctions recorded and the average and standard deviation of the first and second highest bids.

the expected second highest bid is perhaps a more realistic representation of the seller's perceived value of keeping the good.

In Figure 2, we select one specific product in our sample, product #18, which is the "Apple iPhone 5 16GB Black-Slate Verizon". In the figure we plot expected profit as a function of the reserve price given the empirical distribution of the first and second highest bids. Panel A considers the setting where $v_0 = 0$ and panel

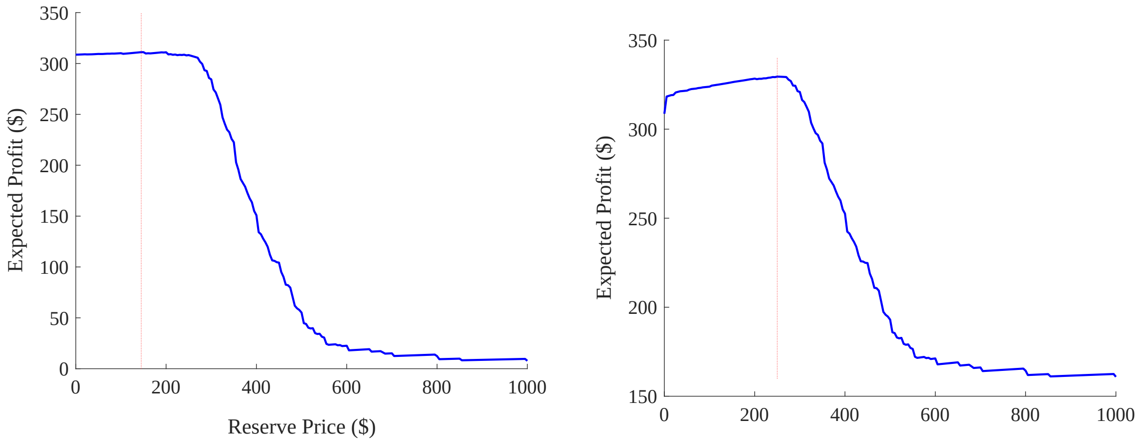
B considers the setting where v_0 is half the average second order statistic. Unsurprisingly for a product supplied elastically on other online or offline platforms, the figure shows that there is a sharp drop-off in profit for reserves beyond a certain point (about \$150 in panel A and about \$275 in panel B). This large drop off illustrates a point also discussed in Ostrovsky and Schwarz (2016) and Kim (2013): the loss from setting a non-optimal reserve price is asymmetric, such that overshooting the optimal reserve has a much larger loss in magnitude than undershooting it.

The vertical line represents the optimal reserve price computed using the first and second highest bid distributions. The gain from the optimal reserve over a reserve of $r = v_0$ is 0.81% in panel A and 0.86% in panel B. Figures 3 and 4 display the same measure for each product separately. In Figure 3, when $v_0 = 0$, the gain from using the estimated reserve price over setting $r = v_0$ ranges from 0.27% to 0.81%, depending on the product. Similarly, in Figure 4, when v_0 is set to half the average second highest bid, the gain from using the estimated reserve price ranges from 0.11–0.86%.

We now turn to the question of how close optimal reserve prices will be to those estimated using a finite history of first and second-highest bids. The theoretical guarantee of Theorem 4 assures us that estimated reserve prices will eventually perform close to optimally. We assess this feature through a simulation exercise.

For each product, we draw 1,000 sequences, each of length 250, at random with replacement from the empirical distribution of all auctions observed for that product over the sample period. Within each sequence, we then estimate the reserve price suggested by our approach using only the first τ observations in the sequence, doing so separately for each $\tau \in \{1, ..., 250\}$. Thus, we begin with only 1 historical auction observation, then 2, then 3, and so on, for each drawn sequence. Next, at each of these estimated reserve prices, using the *full sample* of historical observations for the product, we compute the expected profit the seller would receive from using this computed reserve price. Therefore, for this exercise

FIGURE 2. Profit Under Different Reserve Prices



(A) $v_0 = 0$

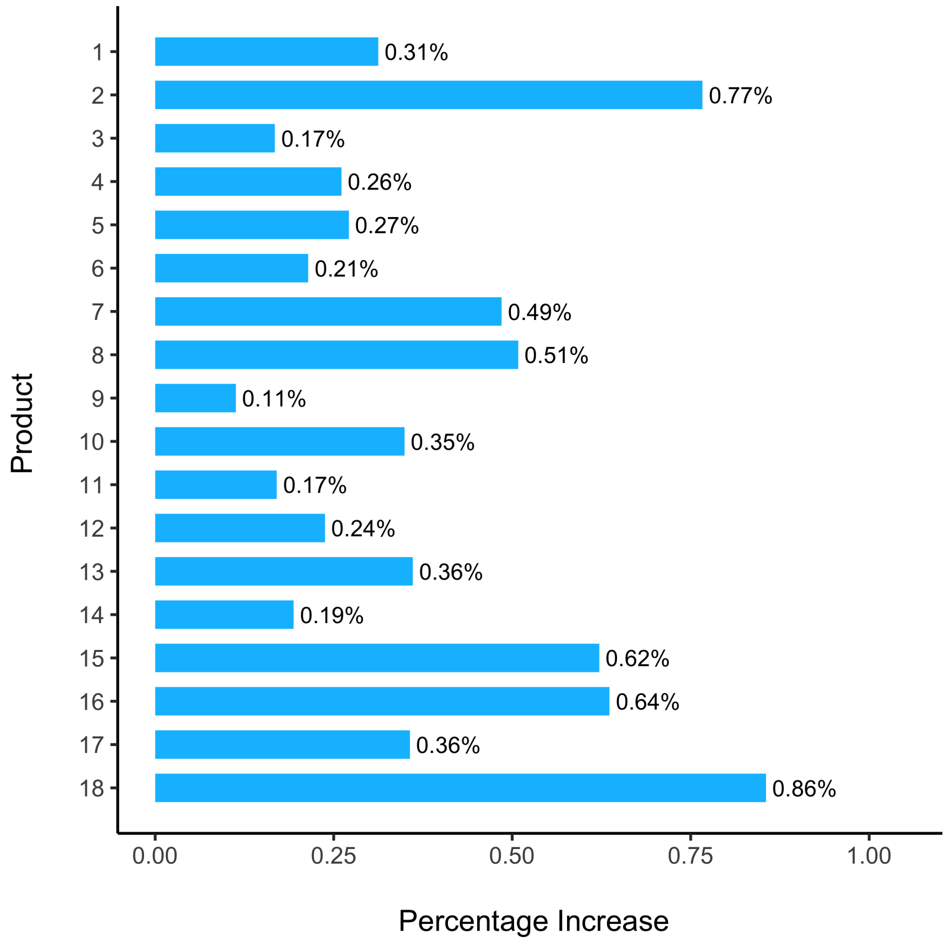
(B) $v_0 = 0.5 * E(V^{(2)})$

Notes: Seller expected profit as a function of the reserve price, given the empirical distribution of first and second highest bids, for Apple iPhone 5 16GB Black-Slate Verizon. Panel A sets $v_0 = 0$ and panel B sets v_0 to half the average second highest bid. Vertical line displays location of optimal reserve price.

we treat the empirical distribution of auctions in our sample as representing the "true" distribution of first and second highest bids, and we treat sellers as only having information on a history of τ auctions drawn at random from the full empirical distribution.

Figure 5 shows the results of this exercise for the same product as in Figure 2. The quantities on the y-axis are expressed as a fraction of the expected profit that would be generated by using the optimal reserve price computed on the full sample of observations for the product (in the case of the Apple iPhone 5 16GB Black-Slate Verizon, there are 516 total observations, as shown in Table 1). The three non-solid lines in Figure 5 are the estimated expected profits (relative to the full sample optimal profit) computed along three of these randomly drawn sequences, progressively incorporating more observations in the sequence in computing the reserve price, as described in the preceding paragraph. The solid red line is an average across all such paths.

FIGURE 3. Revenue Increase from Optimal Reserve Price When $v_0 = 0$

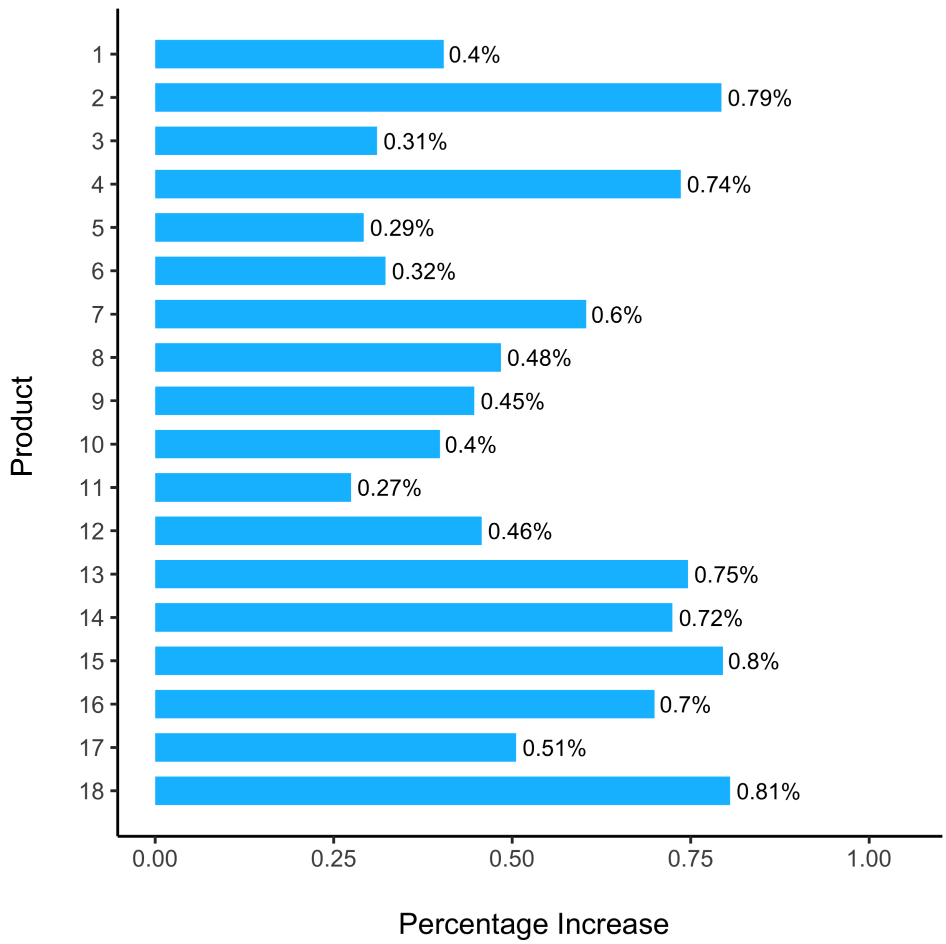


Notes: Figure displays, for each product, the percentage increase in revenue from an auction using the optimal computed reserve price vs. an auction with $r = v_0$, where $v_0 = 0$.

As expected, given Theorem 4, the non-solid and solid lines in Figure 5 do indeed converge to the optimal expected profit level. In the initial phases, estimated reserve prices can be seriously suboptimal, even compared to setting a reserve of $r = v_0$. However, convergence to the optimal level appears to occur quite quickly.

Table 2 displays results from this simulation exercise separately for each product for the two cases we consider for v_0. For each product category, the $q = 0.50$ column displays the median number of auctions required for the computed reserve price to yield expected profit higher than an auction with $r = v_0$. The median

FIGURE 4. Revenue Increase from Optimal Reserve Price When $v_0 = 0.5 * E(V^{(2)})$

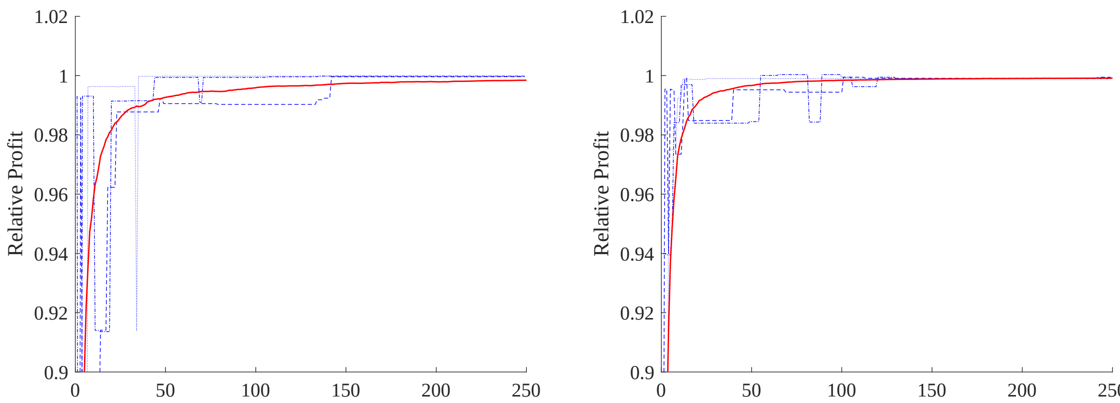


Notes: Figure displays, for each product, the percentage increase in revenue from an auction using the optimal computed reserve price vs. an auction with $r = v_0$, where v_0 is set to half the average second highest bid.

is taken across the 1,000 simulated samples for that product. The $q = 0.05$ and $q = 0.95$ columns represent, respectively, the 0.05 and 0.95 quantiles rather than the median.

For the example product considered in Figures 2 and 5, across all simulated sample paths, the median number of auctions required for the computed reserve price to yield a higher expected profit than simply setting $r = v_0$ is 12 when $v_0 = 0$. In other words, a history of only 12 prior auctions may be sufficient for a seller to

FIGURE 5. Expected Profit From Computed Reserve Prices Using Different Numbers of Observed Auctions



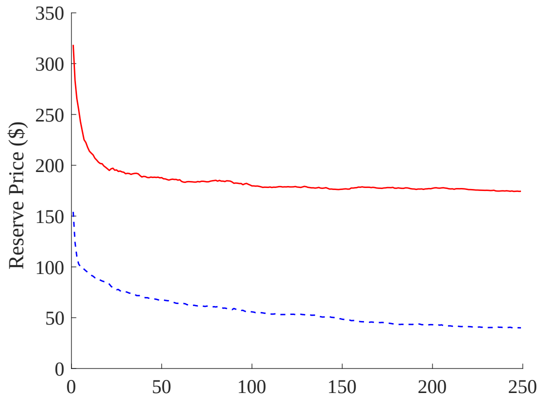
(A) $v_0 = 0$ (B) $v_0 = 0.5 * E(V^{(2)})$

Notes: Expected profit (as a fraction of the optimal expected profit) as a function of the number of auctions observed, where simulations are conducted by drawing sequences of auctions from the empirical distribution of first and second highest bids from auctions for Apple iPhone 5 16GB Black-Slate Verizon and computing the estimated expected profit progressively adding each auction at a time. Non-solid lines represent expected profits estimates from this exercise for three such randomly drawn sequences. Solid red line represents expected profits, averaged over all of the 1,000 drawn samples. Panel A sets $v_0 = 0$ and panel B sets v_0 to half the average second highest bid.

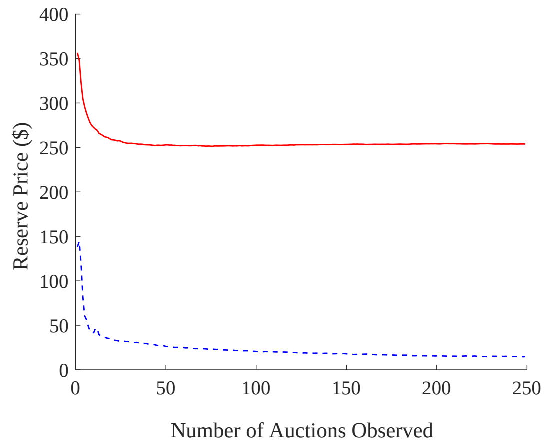
benefit by computing the optimal reserve price based on this historical data. We find that the 0.05 and 0.95 quantiles of the number of auctions required for the computed reserve price to outperform $r = v_0$ are 1 and 120, respectively. Thus, observing 120 auctions of this particular product should be sufficient to ensure a positive gain from using a computed reserve price. When v_0 is instead set to half of the average second highest bid, even fewer auctions are required to achieve at gain over the $r = v_0$ case: the median number of auctions required is 9, and the 0.05 and 0.95 quantiles are 1 and 42, respectively.

The results for other products are similar, with the median number of auctions required falling between 7 and 25 and the 0.95 quantile falling between 42 and 176, with fewer auctions required to achieve a positive gain in the case where the seller's outside option is higher. Given the large data settings in which we see our approach being the most beneficial, these results are encouraging, as many

FIGURE 6. Computed Reserve Prices Using Different Numbers of Observed Auctions



(A) $v_0 = 0$



(B) $v_0 = 0.5 * E(V^{(2)})$

Notes: Figure shows average (solid red line) and standard deviation (dashed blue line) of estimated reserve prices over the simulated sequence of auctions drawn from the empirical distribution of first and second highest bids from auctions for Apple iPhone 5 16GB Black-Slate Verizon. Panel A sets $v_0 = 0$ and panel B sets v_0 to half the average second highest bid.

popular products accumulate long histories of sales rather quickly. These results also suggest that, if a mechanism designer concerned with changes in demand wishes to "reset" or "update" the optimal reserve price based on recently observed auctions, doing so may not be very costly, as it may only require a small sample of recent auctions.

We now turn again to our example product and examine how the computed reserve prices change as the length of the auction sequence used to compute the reserve price increases. Figure 6 explains the poor initial performance of estimated reserve prices. It shows the mean (solid red line) and standard deviation (dashed blue line) of reserve prices for the Apple iPhone 5 16GB Black-Slate Verizon given the number of auctions used to compute the reserve price. Reserve prices converge to the optimal reserve while their variance falls, and both appear to be monotonic. The large standard deviation in reserve prices in the initial auctions implies that the reserve prices recommended by optimization on a short history of auctions are often too high. As seen in Figure 2, this can lead to sharply reduced profit.

TABLE 2. Number of Auctions Required for Computed Reserve Price to Outperform $r = v_0$ Auction

	$v_0 = 0$			$v_0 = 0.5 * E(V^{(2)})$		
Product	q=0.05	q=0.50	q=0.95	q=0.05	q=0.50	q=0.95
1	2	21	136	1	11	68
2	2	15	106	1	8	74
3	2	15	131	2	23	135
4	2	14	97	1	14	79
5	2	20	159	2	18	141
6	1	14	138	2	15	88
7	1	14	92	1	11	59
8	2	15	164	1	10	48
9	2	14	69	1	14	85
10	2	19	141	1	14	79
11	2	17	176	1	11	57
12	2	25	121	1	12	93
13	2	16	93	1	10	55
14	1	12	70	1	12	71
15	1	13	93	1	10	65
16	1	16	110	1	7	58
17	2	18	129	1	12	72
18	1	12	120	1	9	42

Notes: For each product, $q = 0.50$ column displays the median number of auctions required for the computed reserve price to yield expected profit higher than an auction with $r = v_0$. The median is taken across $1,000$ simulated samples drawn with replacement from the full set of observations for that product. The $q = 0.05$ and $q = 0.95$ columns represent, respectively, the 0.05 and 0.95 quantiles rather than the median. The first three columns consider $v_0 = 0$ and the last three consider $v_0 = 0.5 * E(V^{(2)})$.

5. MYERSON OPTIMAL AUCTION CASE

We now shift our focus slightly and restrict our attention to asymmetric *independent* (rather than correlated) private values, to focus on how our approach can be extended to the setting considered by Myerson (1981). We demonstrate that our approach to estimating reserves can be extended to estimate the objects required for implementing the optimal auction (rather than just a single reserve price) in the case considered by Myerson (1981), that of asymmetric bidders with independent private values. In this case, we assume that, in addition to observing the top two bids from auctions without a reserve price, the researcher also observes the identity of the highest bidder (the bidder's type, t); this is a standard assumption in asymmetric auctions (Athey and Haile 2007).

Following Bulow and Roberts (1989), we denote the *marginal revenue* curve of bidder i with type t as $MR_t(\cdot)$. This marginal revenue curve has the same interpretation as in a standard monopoly problem: it is the additional revenue that a monopoly seller would expect to obtain from the buyer, when facing this buyer in isolation (in which case *demand* is given by $1 - F_t(v)$) by raising the price a small amount. We assume here that F_t is *regular*, i.e. that $\psi_t(v) \equiv v - \frac{1-F_t(v)}{f_t(v)}$ is increasing, in which case $MR_t(v) = \psi_t(v)$. In the Myerson (1981) optimal auction, all bidders report their values and the seller assigns the good to the bidder with the highest marginal revenue, as long as it exceeds v_0. The winning bidder pays the least value he could report that would still result in his winning the object, given all other reported values. That is, if i is the winning bidder and i has type t, i pays $MR_t^{-1}\left(\max\left\{v_0, \max_{s \neq t}\{MR_s(v_s)\}\right\}\right)$. To implement this auction, it is necessary to know the marginal revenue functions of each bidder.

5.1. **Identifying and Estimating Marginal Revenue Curves.** We now describe how $MR_t(\cdot)$ can be identified. We first describe the concept of type-specific reserve prices. The feature that the seller assigns the good to the bidder with the highest marginal revenue as long as it exceeds v_0 implies that each bidder faces a type-specific reserve price, which we denote r_t^*, given by $r_t^* = MR_t^{-1}(v_0)$. Continuing with the standard monopoly pricing analogy from Bulow and Roberts (1989) (see also Krishna (2009), p. 67–73), the seller's value, v_0, can be thought of as the *marginal cost* of selling the good, and thus the seller facing demand by a buyer of type t would choose a price r_t^* such that $MR_t(r_t^*) = v_0$.

Now consider a seller designing a second-price auction in which the highest value bidder was known to be of type t. We will refer to such a second-price auction as a *type-t auction*. Let $i_{\max} \equiv \arg\max_{i\in\{1,\dots,\overline{n}\}} V_i$ and denote i's type as $t(i)$. The condition that the highest value bidder being of type t is then denoted $t(i_{\max}) = t$. Using similar notation to (3) above, the expected profit of a seller in a type-t auction is given by

$$p_t(r_t) \equiv \mathbb{E}\left[\pi(V_j^{(1)}, V_j^{(2)}, v_0, r_t) | t(i_{\max}) = t\right] \tag{6}$$

The optimal reserve price in this setting can be obtained simply by maximizing (6) over r_t. As highlighted above, arguments in Bulow and Roberts (1989) imply that, in terms of marginal revenues, a seller designing such an auction should choose a reserve price such that r_t^* such that $MR_t(r_t^*) = v_0$, equivalent to the bidder-specific reserve prices in the Myerson auction. Therefore, maximizing (6) over r_t at different values of v_0 will *trace out* the inverse marginal revenue curve. That is, it will yield $r_t^*(v_0) = MR_t^{-1}(v_0)$, and hence will identify the MR_t curve.

We state this result as the following observation:

Observation 2. *For any type t, the marginal revenue curve $MR_t(\cdot)$ is identified from observations of $V_j^{(1)}$ and $V_j^{(2)}$ from auctions without a reserve price, in which the winner was of type t.*

To estimate $MR_t(\cdot)$, we limit to the J_t auctions in which a type-t bidder was the winner, and estimate the following for a given value of v_0:

$$\hat{p}_{t,J_t}(r_t, v_0) \equiv \frac{1}{J_t}\sum_{j=1}^{J_t}\left[\pi(V_j^{(1)}, V_j^{(2)}, v_0, r)\right] \tag{7}$$

$$\hat{r}_{t,J_t}(v_0) \equiv \arg\max_r \hat{p}_{t,J_t}(r, v_0) \tag{8}$$

where we explicitly state the dependence of the procedure on v_0 in this case. Repeating this process for a range of values for v_0 yields $\hat{MR}_t^{-1}(\cdot)$.

Once marginal revenues have been estimated, the Myerson auction can be implemented for a given set of reported values by determining which bidder has the highest marginal revenue and determining that bidder's payment as the lowest value that bidder could have reported and still won.

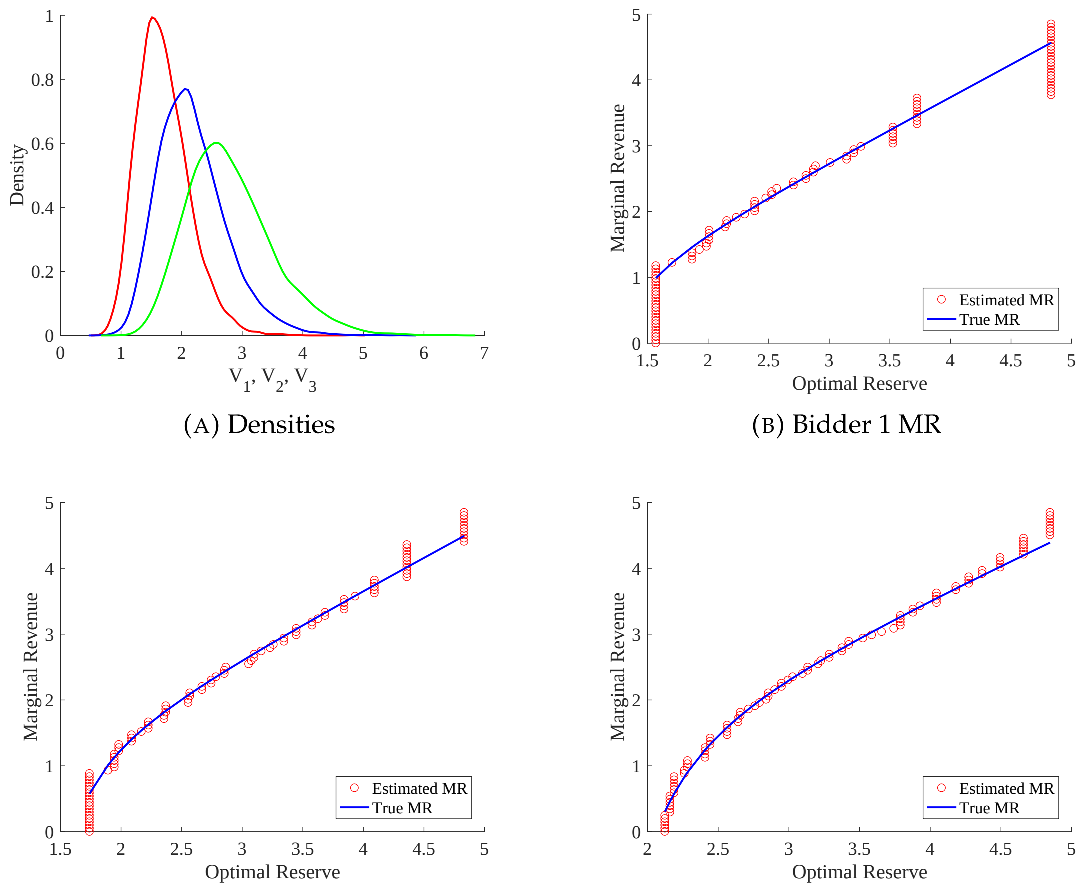
5.2. **Monte Carlo Simulations: Implementing Myerson's Optimal Auction.** We now provide Monte Carlo simulations illustrating the approach in the Myerson case. In these simulations we work with three bidders, although nowhere in the estimation do we require the researcher to know the number of bidders in any given auction; we assume for each auction the researcher observes the top two bids and the identity of the winning bidder.[10] We choose each bidder's values from a different lognormal distribution: $V_1 \sim LOGN(.5, .25)$, $V_2 \sim LOGN(.75, .25)$, and $V_3 \sim LOGN(1, .25)$. We simulate 10,000 auctions. Kernel density estimates of the simulated draws are displayed in Panel A of Figure 7.

To illustrate the simplicity of this method and its potential for use even in Big Data settings, we take a very coarse approach to optimization and estimation, and still find that it performs quite well. We begin by evaluating (6) on a grid of possible reserve prices for each bidder and choose the maximizing reserve prices. We perform this step at different values on a grid of v_0, yielding estimates of bidder-specific marginal revenue curves. Panels B–D of Figure 7 displays, for each bidder, the estimated marginal revenue curve and the true underlying marginal revenue curve, demonstrating that the estimates coincide quite well with the truth. We then implement the Myerson (1981) optimal auction using the same simulated data draws, estimating marginal revenue through linear interpolation of the marginal revenues at grid points.

Figure 8 displays the estimated seller revenue under three different auction designs. The horizontal axis represents the seller's valuation for the good (v_0). The flat, dashed (blue) line is the seller's expected revenue in a no-reserve auction.

[10]Alternatively, one could assume that, for each bidder i, the researcher knows the probability that i is the winning bidder and knows quantiles of the marginal distributions of the first and second order statistics from auctions in which i was the winner.

FIGURE 7. Densities of Bidder Valuations and Estimated Marginal Revenue Functions



(A) Densities

(B) Bidder 1 MR

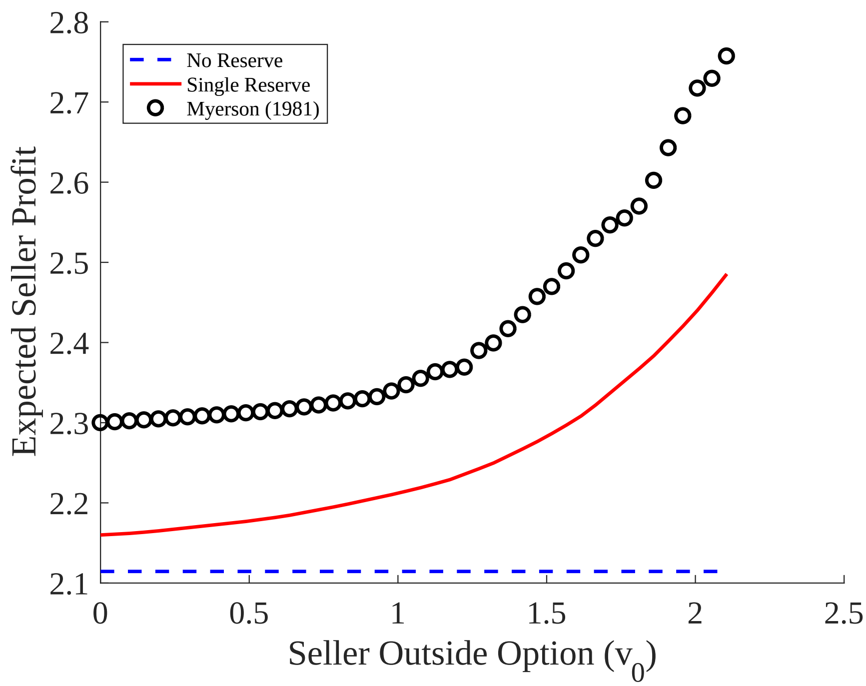
(C) Bidder 2 MR

(D) Bidder 3 MR

Notes: Panel A displays Bidder densities for simulations, where $V_1 \sim LOGN(.5, .25)$, $V_2 \sim LOGN(.75, .25)$, and $V_3 \sim LOGN(1, .25)$. Panels B-D display estimated marginal revenue curves and true marginal revenue curves for bidders in Monte Carlo simulations.

The solid (red) line is the seller's expected revenue from choosing a single reserve price. The final, highest line (marked with black circles) is the expected revenue in the Myerson (1981) auction. Figure 8 demonstrates that the gains from implementing a Myerson's optimal auction (as opposed to a no-reserve auction) can be nearly twice the gains of implementing a single reserve price, especially at lower values of the seller's outside option.

FIGURE 8. Expected Seller Revenue



Notes: Expected seller revenue, as a function of seller's valuation of the good, v_0, under an absolute (no-reserve) auction, auction with a single optimal reserve, or Myerson's optimal auction, from Monte Carlo simulation exercise.

6. CONCLUSION

We study a computationally simple approach for estimating the single optimal reserve price in asymmetric, correlated private values settings and for estimating bidder-specific marginal revenue curves in asymmetric, independent private values settings (in order to implement the Myerson (1981) optimal auction). The approach applies to settings with incomplete bidding data, where only the top two bids are observed, and where the number of bidders is unknown. These data requirements are frequently met in online (advertising or e-commerce) settings. We also derive a new bound on the number of auction records one needs to observe in order for realized revenue based on estimated reserve prices to approximate the optimal revenue arbitrarily closely. We illustrate the approach using eBay data for cataloged products, and illustrate that revenue would increase if optimal reserve prices were employed in practice. We examine the empirical relevance of our theoretical results and find that approximately only 7–25 auctions need to be recorded prior to estimating reserve prices in order for the estimated reserve price to outperform revenue in a no-reserve auction (or an auction with the reserve set to the seller's valuation). In simulated data, we demonstrate the simplicity of our approach to estimating marginal revenue curves and find that it performs well. We also illustrate the implementation of Myerson's optimal auction and the revenue gain that this auction design entails.

While the approach abstracts away from a number of information settings or real-world details (such as common values or inter-auction dynamics), we believe the virtue of the approach is its simplicity, providing a tractable approach to computing reserve prices even in large, unwieldy datasets where typical computationally demanding empirical auction approaches would be infeasible.

References

Abraham, I., Athey, S., Babaioff, M., and Grubb, M. (2016). Peaches, lemons, and cookies: Designing auction markets with dispersed information. Working paper, Stanford University.

Abrevaya, J. and Huang, J. (2005). On the bootstrap of the maximum score estimator. *Econometrica*, 73(4):1175–1204.

Alaei, S., Fu, H., Haghpanah, N., and Hartline, J. (2013). The simple economics of approximately optimal auctions. In *Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 628–637. IEEE.

Aradillas-López, A., Gandhi, A., and Quint, D. (2013). Identification and inference in ascending auctions with correlated private values. *Econometrica*, 81(2):489–534.

Athey, S. and Haile, P. A. (2002). Identification of standard auction models. *Econometrica*, 70(6):2107–2140.

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

Bajari, P. and Hortaçsu, A. (2003). The winner's curse, reserve prices, and endogenous entry: Empirical insights from eBay auctions. *RAND Journal of Economics*, 34(2):329–355.

Balcan, M.-F., Sandholm, T., and Vitercik, E. (2018). A general theory of sample complexity for multi-item profit maximization. Working paper, Carnegie Mellon University.

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

Cattaneo, M. D., Jansson, M., and Nagasawa, K. (2017). Bootstrap-based inference for cube root consistent estimators. Working paper, University of California, Berkeley.

Celis, L. E., Lewis, G., Mobius, M., and Nazerzadeh, H. (2014). Buy-it-now or take-a-chance: Price discrimination through randomization auctions. *Management Science*, 60(12):2927–2948.

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

Chawla, S., Hartline, J., and Nekipelov, D. (2014). Mechanism design for data science. Working paper, Northwestern University.

Coey, D., Larsen, B., and Sweeney, K. (2014). The bidder exclusion effect. NBER Working Paper 20523.

Coey, D., Larsen, B., Sweeney, K., and Waisman, C. (2017). Ascending auctions with bidder asymmetries. *Quantitative Economics*, 8(1):181–200.

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

Deb, R. and Pai, M. M. (2017). Discrimination via symmetric auctions. *American Economic Journal: Microeconomics*, 9(1):275–314.

Delgado, M. A., Rodriguez-Poo, J. M., and Wolf, M. (2001). Subsampling inference in cube root asymptotics with an application to Manski's maximum score estimator. *Economic Letters*, 73(2):241–250.

eMarketer (2016). More than two-thirds of US digital display ad spending is programmatic. `https://www.emarketer.com/Article/More-Than-Two-Thirds-of-US-Digital-Display-Ad-Spending-Programmatic/1013789`. Last accessed on March 16, 2018.

Freyberger, J. and Larsen, B. (2017). Identification in ascending auctions, with an application to digital rights management. NBER Working Paper No. 23569.

Haile, P. A. and Tamer, E. (2003). Inference with an incomplete model of English auctions. *Journal of Political Economy*, 111(1):1–51.

Hong, H. and Li, J. (2017). The numerical delta method and bootstrap. Working paper, Stanford University.

Horowitz, J. L. (1992). A smoothed maximum score estimator for the binary response model. *Econometrica*, 60(3):505–531.

Jun, S. J. and Pinkse, J. (2017). An information-theoretic approach to partially identified auction models. Working paper, Pennsylvania State University.

Kanoria, Y. and Nazerzadeh, H. (2017). Dynamic reserve prices for repeated auctions: Learning from bids. Working paper, Columbia University.

Kim, D.-H. (2013). Optimal choice of a reserve price under uncertainty. *International Journal of Industrial Organization*, 31(5):587–602.

Kim, J. and Pollard, D. (1990). Cube root asymptotics. *The Annals of Statistics*, 18(1):191–219.

Kim, K. I. and Lee, J. (2014). Nonparametric estimation and testing of the symmetric IPV framework with unknown number of bidders. Working paper, University of Minnesota.

Koltchinskii, V. (2001). Rademacher penalties and structural risk minimization. *IEEE Transactions on Information Theory*, 47(5):1902–1914.

Koltchinskii, V. and Panchenko, D. (2002). Empirical margin distributions and bounding the generalization error of combined classifiers. *Annals of Statistics*, 30(1):1–50.

Krishna, V. (2009). *Auction Theory*. Academic press.

Lee, S. M. S. and Pun, M. C. (2006). On m out of n bootstrapping for nonstandard M-estimation with nuisance parameters. *Journal of the American Statistical Association*, 101(475):1185–1197.

Levin, D. and Smith, J. L. (1994). Equilibrium in auctions with entry. *American Economic Review*, 84(3):585–599.

Li, T., Perrigne, I., and Vuong, Q. (2003). Semiparametric estimation of the optimal reserve price in first-price auctions. *Journal of Business & Economic Statistics*, 21(1):53–64.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3(3):205–228.

Mohri, M. and Medina, A. M. (2016). Learning algorithms for second-price auctions with reserve. *Journal of Machine Learning Research*, 17(74):1–25.

Mohri, M., Rostamizadeh, A., and Talwalkar, A. (2012). *Foundations of Machine Learning*. MIT press.

Morgenstern, J. and Roughgarden, J. (2015). The pseudo-dimension of near optimal auctions. In *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS 2015)*, pages 136–144. NIPS.

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

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In Engle, R. F. and McFadden, D., editors, *Handbook of Econometrics vol 4*. Elsevier.

Ostrovsky, M. and Schwarz, M. (2016). Reserve prices in internet advertising auctions: A field experiment. Working paper, Stanford University.

Paarsch, H. J. (1997). Deriving an estimate of the optimal reserve price: An application to British Columbian timber sales. *Journal of Econometrics*, 78(1):333–357.

Platt, B. C. (2017). Inferring ascending auction participation from observed bidders. *International Journal of Industrial Organization*, 54:65–88.

Roughgarden, T. (2014). Approximately optimal mechanism design: Motivation, examples, and lessons learned. *ACM SIGEcom Exchanges*, 13(2):4–20.

Roughgarden, T. and Wang, J. R. (2016). Minimizing regret with multiple reserves. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC '16)*, pages 601–616. ACM.

Samuelson, W. F. (1985). Competitive bidding with entry costs. *Economics Letters*, 17(1):53–57.

Song, U. (2004). Nonparametric estimation of an eBay auction model with an unknown number of bidders. Working paper, University of British Columbia.

Tang, X. (2011). Bounds on revenue distributions in counterfactual auctions with reserve prices. *RAND Journal of Economics*, 42(1):175–203.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer.

Zeithammer, R. (2017). The futility of soft floors in auctions. Working paper, UCLA.

APPENDIX A. PROOFS

A.1. **Proof of Theorem 1.**

Proof. We will show that $\hat{p}(r)$ converges uniformly in probability to $p(r)$. Note that $\pi(V_j^{(1)}, V_j^{(2)}, v_0, r)$ can alternatively be written as $\max\{r, V_j^{(2)}\} + (v_0 - r)\mathbb{1}(V_j^{(1)} < r)$ (see Aradillas-López et al. (2013)). For simplicity, assume that $v_0 = 0$ and let $\hat{p}_1(r) = \frac{1}{J}\sum_{j=1}^{J} \max\left\{r, V_j^{(2)}\right\}$ and $\hat{p}_2(r) = -\frac{1}{J}\sum_{j=1}^{J} r\mathbb{1}\left\{V_j^{(1)} < r\right\}$, so that $\hat{p}(r) = \hat{p}_1(r) + \hat{p}_2(r)$. Notice that $\hat{p}_1(r)$ is Lipschitz continuous because, for any r_1 and r_2, it follows that $|\hat{p}_1(r_1) - \hat{p}_1(r_2)| \leqslant |r_1 - r_2|$. Furthermore, for any r, it follows by the law of large numbers that $\hat{p}_1(r) \xrightarrow{p} p_1(r)$. Thus, we can invoke Lemma 2.9 in Newey and McFadden (1994) to obtain $\sup_r |\hat{p}_1(r) - p_1(r)| \xrightarrow{p} 0$. Finally, it is straightforward to check that the function $f(x, r) = -r\mathbb{1}\{x \leqslant r\}$ belongs to a VC subgraph class (see, for example, van der Vaart and Wellner (1996)), which guarantees uniform convergence of $\hat{p}_2(\cdot)$. Consequently, we have $\sup_r |\hat{p}(r) - p(r)| \xrightarrow{p} 0$, which guarantees that $\hat{r}_J \xrightarrow{p} r^*$ and $\hat{p}(\hat{r}_J) \xrightarrow{p} p(r^*)$. □

A.2. **Proof of Theorem 2.**

Proof. Let $\tilde{p}(r) = \frac{1}{J}\sum_{j=1}^{J} \tilde{\pi}(\xi_j, r)$. Throughout the proof we will use r_1 and r_2 such that, without loss of generality, $r_1 > r_2 > r^*$. Throughout this proof, we will denote the joint density of the two highest bids as $f_{1,2}(\cdot,\cdot)$ and the marginals as $f_1(\cdot)$ and $f_2(\cdot)$.

By the main theorem of Kim and Pollard (1990), if Assumptions 1 and 2 are satisfied, and r^* is an interior point, and if the following conditions hold

(1) $H(\beta, \alpha) = \lim_{t\to 0} \frac{1}{t}\mathbb{E}[\tilde{\pi}(\cdot, r^* + \beta t)\tilde{\pi}(\cdot, r^* + \alpha t)]$ exists for each β, α in $\mathbb{R}$ and
$$\lim_{t\to 0} \frac{1}{t}\mathbb{E}[\tilde{\pi}(\cdot, r^* + \alpha t)^2 \mathbb{1}\{|\tilde{\pi}(\cdot, r^* + \alpha t)| > \epsilon/t\}] = 0$$
for each $\epsilon > 0$ and α in $\mathbb{R}$;

(2) $\mathbb{E}[P_R^2] = O(R)$ as $R \to 0$ and for each $\epsilon > 0$ there is a constant K such that $\mathbb{E}[P_R^2 \mathbb{1}\{P_R > K\}] < \epsilon R$ for R near 0;

(3) $\mathbb{E}[|\tilde{\pi}(\cdot, r_1) - \tilde{\pi}(\cdot, r_2)|] = O(|r_1 - r_2|)$ near r^*;

Then the process $J^{2/3}\frac{1}{J}\sum_{j=1}^{J}\tilde{\pi}(\cdot, r^* + \alpha J^{-1/3})$ converges in distribution to a Gaussian process $Z(\alpha)$ with continuous sample paths, expected value $-\frac{1}{2}\alpha^2\Sigma$, and covariance kernel H, where $H(\beta, \alpha) = \lim_{t\to 0}\frac{1}{t}\mathbb{E}[\tilde{\pi}(\cdot, r^* + \beta t)\tilde{\pi}(\cdot, r^* + \alpha t)]$ for any β, α in $\mathbb{R}$. Furthermore, if Z has nondegenerate increments, then $J^{1/3}(\hat{r}_J - r^*)$ converges in distribution to the random maximizer of Z. We now prove that conditions (1)–(3) above are satisfied.

To establish condition (1), we first characterize the limiting behavior of $\frac{1}{t}\mathbb{E}[\tilde{\pi}(\cdot, r^* + \beta t)\tilde{\pi}(\cdot, r^* + \alpha t)]$ as $t \to 0$. By the definition of $\tilde{\pi}(\cdot, r)$, this amounts to studying the behavior of four different terms, which we conduct separately below. Let $r_1 = r^* + \alpha t$ and $r_2 = r^* + \beta t$. First, we consider the term

$$h_1 \equiv \left(\max\{V^{(2)}, r_1\} - \max\{V^{(2)}, r^*\}\right)\left(\max\{V^{(2)}, r_2\} - \max\{V^{(2)}, r^*\}\right).$$

Notice that when $V^{(2)} > r_2$ then $h_1 = 0$; when $V^{(2)} < r^*$ then $h_1 = (r_1 - r^*)(r_2 - r^*)$; and when $r^* < V^{(2)} < r_2$ then $h_1 = (r_1 - V^{(2)})(r_2 - V^{(2)})$. Therefore,

$$\begin{aligned}
&\frac{1}{t}\mathbb{E}\left[\left(\max\{V^{(2)}, r_1\} - \max\{V^{(2)}, r^*\}\right)\left(\max\{V^{(2)}, r_2\} - \max\{V^{(2)}, r^*\}\right)\right] \\
&= \frac{1}{t}\Big\{(r_1 - r^*)(r_2 - r^*)Pr(V^{(2)} < r^*) \\
&\qquad + \mathbb{E}\left[(r_1 - V^{(2)})(r_2 - V^{(2)})|r^* < V^{(2)} < r_2\right]Pr(r^* < V^{(2)} < r_2)\Big\} \\
&= \frac{1}{t}\left\{\alpha\beta t^2\int_0^{r^*} f_2(u)du + r_1 r_2\int_{r^*}^{r_2} f_2(u)du - (r_1 + r_2)\int_{r^*}^{r_2} u f_2(u)du + \int_{r^*}^{r_2} u^2 f_2(u)du\right\} \\
&= \frac{1}{t}\Big\{\alpha\beta t^2 F_2(r^*) + \left[(r^*)^2 + (\alpha+\beta)t + \alpha\beta t^2\right][f_2(r^*)(r_2 - r^*) + o(r_2 - r^*)] \\
&\qquad - [2r^* + (\alpha+\beta)t][r^* f_2(r^*)(r_2 - r^*) + o(r_2 - r^*)] + \left[(r^*)^2 f_2(r^*)(r_2 - r^*) + o(r_2 - r^*)\right]\Big\} \\
&= \frac{1}{t}\left\{\alpha\beta t^2 F_2(r^*) + \alpha\beta t^2 f_2(r^*)\beta t + o(t)\right\} \\
&= \frac{1}{t}o(t) = o(1).
\end{aligned}$$

The second term we consider is

$$h_2 \equiv \left(r^* \mathbb{1}\{V^{(1)} < r^*\} - r_2 \mathbb{1}\{V^{(1)} < r_2\}\right) \left(\max\{V^{(2)}, r_1\} - \max\{V^{(2)}, r^*\}\right).$$

When $V^{(1)} < r^*$, $h_2 = (r^* - r_2)(r_1 - r^*)$; when $r^* < V^{(1)} < r_2$ and $V^{(2)} < r^*$, $h_2 = -r_2(r_1 - r^*)$; and when $r^* < V^{(1)} < r_2$ and $r^* < V^{(2)} < V^{(1)}$, $h_2 = -r_2(r_1 - V^{(2)})$. Hence,

$$\begin{aligned}
&\frac{1}{t}\mathbb{E}\left[\left(r^* \mathbb{1}\{V^{(1)} < r^*\} - r_2 \mathbb{1}\{V^{(1)} < r_2\}\right) \left(\max\{V^{(2)}, r_1\} - \max\{V^{(2)}, r^*\}\right)\right] \\
&= \frac{1}{t}\Big\{(r^* - r_2)(r_1 - r^*)Pr(V^{(1)} < r^*) - r_2(r_1 - r^*)Pr(V^{(2)} < r^* < V^{(1)} < r_2) \\
&\quad - r_2 r_1 Pr(r^* < V^{(2)} < V^{(1)} < r_2) \\
&\quad + \mathbb{E}[V^{(2)} | r^* < V^{(2)} < V^{(1)} < r_2] Pr(r^* < V^{(2)} < V^{(1)} < r_2)\Big\} \\
&= \frac{1}{t}\Big\{-\alpha\beta t^2 F_1(r^*) - r_2 \alpha t \int_{r^*}^{r_2}\int_0^{r^*} f_{1,2}(u,v)dudv - r_1 r_2 \int_{r^*}^{r_2}\int_{r^*}^{v} f_{1,2}(u,v)dudv \\
&\quad + r_2 \int_{r^*}^{r_2}\int_{r^*}^{v} u f_{1,2}(u,v)dudv\Big\} \\
&= \frac{1}{t}\Big\{-\alpha\beta t^2 F_1(r^*) - r_2 \alpha t \int_{r^*}^{r_2} F_v(r^*, v)dv - r_1 r_2 \int_{r^*}^{r_2} [F_v(v,v) - F_v(r^*, v)]dv \\
&\quad + r_2 \int_{r^*}^{r_2} [F_{1,2}(v,v) - F_{1,2}(v, r^*)]dv\Big\} \\
&= \frac{1}{t}\Big\{-\alpha\beta t^2 F_1(r^*) - r_2 \alpha t[\beta t F_v(r^*, r^*) + o(t)] - r_1 r_2 o(t) + r_2 o(t)\Big\} \\
&= \frac{1}{t} o(t) = o(1).
\end{aligned}$$

The third term is

$$h_3 \equiv \left(r^* \mathbb{1}\{V^{(1)} < r^*\} - r_1 \mathbb{1}\{V^{(1)} < r_1\}\right) \left(\max\{V^{(2)}, r_2\} - \max\{V^{(2)}, r^*\}\right).$$

The term $h_3 = 0$ when $V^{(1)} > r_1$ or $V^{(2)} > r_2$; $h_3 = -r_1(r_2 - r^*)$ when $V^{(2)} < r^* < V^{(1)} < r_1$; $h_3 = -r_1(r_2 - V^{(2)})$ when $r^* < V^{(1)} < r_1$ and $r^* < V^{(2)} < r_2$; and $h_3 = (r^* - r_1)(r_2 - r^*)$ when $V^{(1)} < r^*$. Consequently,

$$\frac{1}{t}\mathbb{E}\left[\left(r^* \mathbb{1}\{V^{(1)} < r^*\} - r_1 \mathbb{1}\{V^{(1)} < r_1\}\right) \left(\max\{V^{(2)}, r_2\} - \max\{V^{(2)}, r^*\}\right)\right]$$

$$
\begin{aligned}
&= \frac{1}{t}\Big\{-r_1(r_2-r^*)Pr(V^{(2)} < r^* < V^{(1)} < r_1) \\
&\quad -r_1\left(r_2 - \mathbb{E}[V^{(2)}|r^* < V^{(1)} < r_1, r^* < V^{(2)} < r_2]\right)Pr(r^* < V^{(1)} < r_1, r^* < V^{(2)} < r_2)\Big\} \\
&= \frac{1}{t}\Big\{-r_1\beta t\int_{r^*}^{r_1}\int_0^{r^*} f_{1,2}(u,v)dudv - \alpha\beta t^2 F_1(r^*) - r_1r_2\int_{r^*}^{r_2}\int_{r^*}^{v} f_{1,2}(u,v)dudv \\
&\quad -r_1r_2\int_{r_2}^{r_1}\int_{r^*}^{r_2} f_{1,2}(u,v)dudv \\
&\quad +r_1\int_{r^*}^{r_2}\int_{r^*}^{v} uf_{1,2}(u,v)dudv + r_1\int_{r_2}^{r_1}\int_{r^*}^{r_2} uf_{1,2}(u,v)dudv\Big\} \\
&= \frac{1}{t}\Big\{-r_1\beta t\int_{r^*}^{r_1} F_v(r^*,v)dv - \alpha\beta t^2F_1(r^*) - r_1r_2\int_{r^*}^{r_2}[F_v(v,v)-F_v(r^*,v)]dv \\
&\quad -r_1r_2\int_{r_2}^{r_1} f_{1,2}(r^*,v)(r_2-r^*)dv \\
&\quad +r_1\int_{r^*}^{r_2}[F_{1,2}(v,v)-F_{1,2}(v,r^*)]dv + r_1r^*\int_{r_2}^{r_1} f_{1,2}(r^*,v)(r_2-r^*)dv + o[(r_2-r^*)^2]\Big\} \\
&= \frac{1}{t}\Big\{-r_1\alpha\beta t^2F_v(r^*,r^*) - \alpha\beta t^2F_1(r^*) - r_1r_2 o(r_1-r^*) \\
&\quad +r_1 o(r_2-r^*) - r_1\beta^2t^2\int_{r_2}^{r_1} f_{1,2}(r^*,v)dv + o[(r_2-r^*)^2]\Big\} \\
&= \frac{1}{t}o(t) = o(1).
\end{aligned}
$$

The fourth and final term is

$$
h_4 \equiv \left(r^*\mathbb{1}\{V^{(1)} < r^*\} - r_2\mathbb{1}\{V^{(1)} < r_2\}\right)\left(r^*\mathbb{1}\{V^{(1)} < r^*\} - r_1\mathbb{1}\{V^{(1)} < r_1\}\right).
$$

The term $h_4 = 0$ if $V^{(1)} > r_2$; $h_4 = r_1r_2$ if $r^* < V^{(1)} < r_2$; and $h_4 = (r^* - r_2)(r^* - r_1)$ if $V^{(1)} < r^*$. Thus,

$$
\begin{aligned}
&\frac{1}{t}\mathbb{E}\left[\left(r^*\mathbb{1}\{V^{(1)} < r^*\} - r_2\mathbb{1}\{V^{(1)} < r_2\}\right)\left(r^*\mathbb{1}\{V^{(1)} < r^*\} - r_1\mathbb{1}\{V^{(1)} < r_1\}\right)\right] \\
&= \frac{1}{t}\left\{r_1r_2Pr(r^* < V^{(1)} < r_2) + (r^*-r_2)(r^*-r_1)Pr(V^{(1)} < r^*)\right\} \\
&= \frac{1}{t}\left\{r_1r_2\int_{r^*}^{r_2} f_1(v)dv + \alpha\beta t^2F_1(r^*)\right\} \\
&= \frac{1}{t}\left\{r_1r_2[f_1(r^*)(r_2-r^*) + o(r_2-r^*)] + \alpha\beta t^2F_1(r^*)\right\}
\end{aligned}
$$

$$\frac{1}{t}\left\{(r^*)^2 f_1(r^*)\beta t + o(t)\right\}$$
$$= (r^*)^2 f_1(r^*)\beta + o(1).$$

These four results show that the limit $H(\beta, \alpha)$ is well defined for each β, α in $\mathbb{R}$, which establishes the first part of (1). For the second part of (1), notice that $|\tilde{\pi}(\cdot, r^* + \alpha t)|$ is bounded for any t, which means that there exists a $\underline{t} < \infty$ such that the indicator would be 0 for all $t < \underline{t}$, establishing the result.

We now establish (2). Let $R > 0$ and $\check{r}$ be the maximizer of $\tilde{\pi}(\cdot, r)$ such that $|r - r^*| < R$. We first need to show that $\mathbb{E}[\tilde{\pi}(\cdot, \check{r})^2] = O(R)$. We will split the analysis into three terms. First, notice that for the first term:

$$(\max\{V^{(2)}, \check{r}\} - \max\{V^{(2)}, r^*\})^2 = (|\max\{V^{(2)}, \check{r}\} - \max\{V^{(2)}, r^*\}|)^2 \quad \leqslant (|\check{r} - r^*|)^2 < R^2$$

which implies that its expected value is $O(R^2)$. Moving to the next two terms we will assume that $\check{r} > r^*$, as the calculations for the opposite case are analogous. For the second one,

$$(r^*\mathbb{1}\{V^{(1)} < r^*\} - \check{r}\mathbb{1}\{V^{(1)} < \check{r}\})^2 = (r^*)^2\mathbb{1}\{V^{(1)} < r^*\} + \check{r}^2\mathbb{1}\{V^{(1)} < \check{r}\})^2 - 2r^*\check{r}\mathbb{1}\{V^{(1)} < r^*\}$$

so taking expectations yields:

$$(r^*)^2\int_0^{r^*} f_1(v)dv - r^*\check{r}\int_0^{r^*} f_1(v)dv + +\check{r}^2\int_{r^*}^{\check{r}} f_1(v)dv + \check{r}^2\int_0^{r^*} f_1(v)dv - r^*\check{r}\int_0^{r^*} f_1(v)dv$$
$$= (\check{r} - r^*)^2 F_1(r^*) + \check{r}^2[f_1(r^*)(\check{r} - r^*) + o(\check{r} - r^*)]$$
$$= O(\check{r} - r^*) + o(\check{r} - r^*) = O(\check{r} - r^*) < O(R).$$

The third and last term is given by:

$$\mathbb{E}\left[(\max\{V^{(2)}, \check{r}\} - \max\{V^{(2)}, r^*\})(r^*\mathbb{1}\{V^{(1)} < r^*\} - \check{r}\mathbb{1}\{V^{(1)} < \check{r}\})\right]$$
$$= \mathbb{E}\left[(\max\{V^{(2)}, r^*\}r^*\mathbb{1}\{V^{(1)} < r^*\}\right] - \mathbb{E}\left[(\max\{V^{(2)}, \check{r}\}\check{r}\mathbb{1}\{V^{(1)} < \check{r}\}\right]$$
$$+ \mathbb{E}\left[(\max\{V^{(2)}, r^*\}\check{r}\mathbb{1}\{V^{(1)} < \check{r}\}\right] - \mathbb{E}\left[(\max\{V^{(2)}, \check{r}\}r^*\mathbb{1}\{V^{(1)} < r^*\}\right]$$

$$
\begin{aligned}
&= (r^*)^2 \int_0^{r^*} \int_0^{v} f_{1,2}(u,v)dudv - \check{r}^2 \int_0^{\check{r}} \int_0^{v} f_{1,2}(u,v)dudv \\
&\qquad + \check{r} \int_0^{\check{r}} \int_0^{v} \max\{u, r^*\} f_{1,2}(u,v)dudv - r^*\check{r} \int_0^{r^*} \int_0^{v} f_{1,2}(u,v)dudv \\
&= [(r^*)^2 - \check{r}^2] \int_0^{r^*} \int_0^{v} f_{1,2}(u,v)dudv - \check{r}^2 \int_{r^*}^{\check{r}} \int_0^{v} f_{1,2}(u,v)dudv \\
&\qquad + \check{r} \int_{r^*}^{\check{r}} \int_0^{v} \max\{u, r^*\} f_{1,2}(u,v)dudv \\
&= (r^* - \check{r})(r^* + \check{r}) F_{1,2}(r^*, r^*) - \check{r}^2 \int_{r^*}^{\check{r}} F_v(v,v)dv + r^*\check{r} \int_{r^*}^{\check{r}} \int_0^{r^*} f_{1,2}(u,v)dudv \\
&\qquad + \check{r} \int_{r^*}^{\check{r}} \int_{r^*}^{v} u f_{1,2}(u,v)dudv \\
&= O(r^* - \check{r}) - \check{r}(\check{r} - r^*)[F_v(r^*, r^*)(r^* - \check{r}) + o(r^* - \check{r})] + \check{r} \int_{r^*}^{\check{r}} G(v)dv \\
&= O(r^* - \check{r}) + \check{r}[G(r^*)(\check{r} - r^*) + o(\check{r} - r^*)] = O(\check{r} - r^*) < O(R),
\end{aligned}
$$

which, along with the results for the previous two terms, establishes the first part of condition (2). The second part of condition (2) follows directly from the integrability of $\tilde{\pi}(\cdot, \check{r})^2$.

To verify that (3) holds, notice that:

$$
\begin{aligned}
&|\tilde{\pi}(\xi_j, r_1) - \tilde{\pi}(\xi_j, r_2)| \\
&= |\pi(\xi_j, r_1) - \pi(\xi_j, r_2)| \\
&= \left| \max\left\{V_j^{(2)}, r_1\right\} - r_1 \mathbb{1}\left\{V_j^{(1)} < r_1\right\} - \max\left\{V_j^{(2)}, r_2\right\} + r_2 \mathbb{1}\left\{V_j^{(1)} < r_2\right\} \right| \\
&\leqslant \left| \max\left\{V_j^{(2)}, r_1\right\} - \max\left\{V_j^{(2)}, r_2\right\} \right| + \left| r_2 \mathbb{1}\left\{V_j^{(1)} < r_2\right\} - r_1 \mathbb{1}\left\{V_j^{(1)} < r_1\right\} \right| \\
&\leqslant |r_2 - r_1| + \mathbb{1}\left\{V_j^{(1)} < r_2\right\} |r_2 - r_1| + r_1 \left| \mathbb{1}\left\{V_j^{(1)} < r_2\right\} - \mathbb{1}\left\{V_j^{(1)} < r_1\right\} \right|
\end{aligned}
$$

Taking the expectation, we obtain

$$
\begin{aligned}
\mathbb{E}\left|\tilde{\pi}(\xi_j, r_1) - \tilde{\pi}(\xi_j, r_2)\right| &\leqslant |r_2 - r_1| + |r_2 - r_1| Pr(V_j^{(1)} < r_2) + r_1 Pr(r_2 < V_j^{(1)} < r_1) \\
&= O(|r_2 - r_1|) + f_1(r_2)(r_1 - r_2) + o(r_1 - r_2) = O(|r_2 - r_1|).
\end{aligned}
$$

This establishes (3).

Finally, we derive Σ. Notice that:

$$\begin{aligned}
\mathbb{E}[\tilde{\pi}(\cdot,r)] &= \mathbb{E}[\pi(\cdot,r)] - \mathbb{E}[\pi(\cdot,r^*)] \\
&= \int_0^{\bar{\omega}} \max\{r, V^{(2)}\} f_2(u)du - \int_0^{\bar{\omega}} r\mathbb{1}\{V^{(1)} < r\} f_1(v)dv - \mathbb{E}[\pi(\cdot,r^*)] \\
&= r\int_0^r f_2(u)du + \int_r^{\bar{\omega}} u f_2(u)du - r\int_0^r f_1(v)dv - \mathbb{E}[\pi(\cdot,r^*)]
\end{aligned}$$

Differentiating this expression with respect to r, we obtain

$$\begin{aligned}
\frac{\partial \mathbb{E}[\tilde{\pi}(\cdot,r)]}{\partial r} &= \int_0^r f_2(u)du + rf_2(r) - rf_2(r) - \int_0^r f_1(v)dv - rf_1(r) \\
&= \int_0^r f_2(u)du - \int_0^r f_1(v)dv - rf_1(r)
\end{aligned}$$

Taking the second derivative and evaluating at r^*, we obtain

$$\Sigma = -f_2(r^*) + 2f_1(r^*) + r^* f_1'(r^*).$$

□

A.3. **Proof of Theorem 3.**

Proof. This theorem follows directly from Theorem 2. First, notice that $\hat{p}(\hat{r}_J) - p(r^*) = \frac{1}{J}\sum_{j=1}^J \tilde{\pi}(\xi_j, \hat{r}_J) + \hat{p}(r^*) - p(r^*)$. Since $\hat{r}_J - r^* = O_P(J^{-1/3})$ and $\frac{1}{J}\sum_{j=1}^J \tilde{\pi}(\cdot, r^* + \alpha J^{-1/3}) = O_P(J^{-2/3})$, the term $\frac{1}{J}\sum_{j=1}^J \tilde{\pi}(\xi_j, \hat{r}_J)$ is also $O_P(J^{-2/3})$. A simple application of the Central Limit Theorem establishes the result. □

A.4. **Proof of Theorem 4.** We will use slightly different notation in this proof than elsewhere in the paper, letting $S_J = (z_1, \ldots, z_J)$ be a fixed sample of size J and denoting quantities estimated on this sample by a subscript S_J.

We start with the definition of *empirical Rademacher complexity*:

Definition. *Let G be a family of functions from Z to $[a, b]$, and $S_J = (z_1, \ldots, z_J)$ a fixed sample of size J with elements in Z. Then the empirical Rademacher complexity of G with* *respect to S_J is defined as:*

$$\widehat{\mathfrak{R}}_{S_J}(G) = E_{\boldsymbol{\sigma}}\left[\sup_{g\in G}\frac{1}{J}\sum_{j=1}^{J}\sigma_j g(z_j)\right]$$

where $\boldsymbol{\sigma} = (\sigma_1,\ldots,\sigma_J)$, and each σ_j is an independent uniform random variable with values in $\{-1,+1\}$.

We state a useful preliminary result.

Theorem 5. *Let G be a family of functions mapping Z to $[0,1]$. Then for any $\delta > 0$, with probability at least $1-\delta$, for all $g \in G$:*

$$E[g(z)] - \frac{1}{J}\sum_{j=1}^{J} g(z_j) \leqslant 2\widehat{\mathfrak{R}}_{S_J}(G) + 3\sqrt{\frac{\log\frac{2}{\delta}}{2J}}$$

Proof. See Mohri et al. (2012), Theorem 3.1. □

This result can be straightforwardly adapted to obtain a two-sided bound.

Corollary 1. *Let G be a family of functions mapping Z to $[0,1]$. Then for any $\delta > 0$, with probability at least $1-\delta$, for all $g \in G$:*

$$\left|E[g(z)] - \frac{1}{J}\sum_{j=1}^{J} g(z_j)\right| \leqslant 2\widehat{\mathfrak{R}}_{S_J}(G) + 3\sqrt{\frac{\log\frac{4}{\delta}}{2J}}$$

Proof. Applying Theorem 5 above to $G' = \{-g+1 : g \in G\}$ and noting that $\widehat{\mathfrak{R}}_{S_J}(G) = \widehat{\mathfrak{R}}_{S_J}(G')$, we obtain the result that for any $\delta/2 > 0$, with probability at least $1-\delta/2$, for all $g \in G$:

$$\frac{1}{J}\sum_{j=1}^{J} g(z_j) - E[g(z)] \leqslant 2\widehat{\mathfrak{R}}_{S_J}(G) + 3\sqrt{\frac{\log\frac{4}{\delta}}{2J}}.$$

Theorem 5 also implies that for any $\delta/2 > 0$, with probability at least $1-\delta/2$, for all $g \in G$:

$$E[g(z)] - \frac{1}{J}\sum_{j=1}^{J} g(z_j) \leqslant 2\widehat{\mathfrak{R}}_{S_J}(G) + 3\sqrt{\frac{\log\frac{4}{\delta}}{2J}}.$$

Combining these two results and applying the union bound gives the desired result. □

Now for simplicity let $v_0 = 0$ so that $\pi(V_j^{(1)}, V_j^{(2)}, r) = r\mathbb{1}(V_j^{(2)} < r \leqslant V_j^{(1)}) + V_j^{(2)}\mathbb{1}(r \leqslant V_j^{(2)})$ and $G \equiv \{\pi(\cdot,\cdot,r) : r \in [0,\overline{\omega}]\}$. We can now prove an upper bound on $p(r^*) - p(\hat{r}_{S_J})$ in terms of the empirical Rademacher complexity of G.

Lemma 1. *Let $0 \leqslant V_j^{(1)} \leqslant \overline{\omega} < \infty$. For any $\delta > 0$, with probability at least $1-\delta$, it holds that $p(r^*) - p(\hat{r}_{S_J}) \leqslant 4\hat{\mathfrak{R}}_{S_J}(G) + 6\overline{\omega}\sqrt{\frac{\log\frac{4}{\delta}}{2J}}$.*

Proof.

$$\begin{aligned} p(r^*) - p(\hat{r}_{S_J}) &= p(r^*) - \hat{p}_{S_J}(\hat{r}_{S_J}) + \hat{p}_{S_J}(\hat{r}_{S_J}) - p(\hat{r}_{S_J}) \\ &\leqslant p(r^*) - \hat{p}_{S_J}(r^*) + \hat{p}_{S_J}(\hat{r}_{S_J}) - p(\hat{r}_{S_J}). \\ &\leqslant 2 \sup_{r\in[0,\overline{\omega}]} |p(r) - \hat{p}_{S_J}(r)|. \end{aligned}$$

The first inequality follows because $\hat{r}_{S_J}$ maximizes $\hat{p}_{S_J}$ by definition. Applying Corollary 1 with $z_j = (V_j^{(1)}, V_j^{(2)})$ we have that for any $\delta > 0$, with probability at least $1-\delta$:

$$\sup_{r\in[0,\overline{\omega}]} \left| \frac{1}{\overline{\omega}} E[\pi(V_j^{(1)}, V_j^{(2)}, r)] - \frac{1}{J\overline{\omega}} \sum_{j=1}^{J} \pi(V_j^{(1)}, V_j^{(2)}, r) \right| \leqslant \frac{2}{\overline{\omega}} \hat{\mathfrak{R}}_{S_J}(G) + 3\sqrt{\frac{\log\frac{4}{\delta}}{2J}},$$

or equivalently,

$$\sup_{r\in[0,\overline{\omega}]} |p(r) - \hat{p}_{S_J}(r)| \leqslant 2\hat{\mathfrak{R}}_{S_J}(G) + 3\overline{\omega}\sqrt{\frac{\log\frac{4}{\delta}}{2J}}.$$

Therefore for any $\delta > 0$, with probability at least $1-\delta$:

$$p(r^*) - p(\hat{r}_{S_J}) \leqslant 4\hat{\mathfrak{R}}_{S_J}(G) + 6\overline{\omega}\sqrt{\frac{\log\frac{4}{\delta}}{2J}}.$$

□

Following Mohri and Medina (2016), define $\pi_1(V_j^{(1)}, V_j^{(2)}, r) = V_j^{(2)}\mathbb{1}(r \leqslant V_j^{(2)}) + r\mathbb{1}(V_j^{(2)} < r \leqslant V_j^{(1)}) + V_j^{(1)}\mathbb{1}(V_j^{(1)} < r)$ and $\pi_2(V_j^{(1)}, r) = -V_j^{(1)}\mathbb{1}(V_j^{(1)} < r)$, so that

$\pi(V_j^{(1)}, V_j^{(2)}, r) = \pi_1(V_j^{(1)}, V_j^{(2)}, r) + \pi_2(V_j^{(1)}, r)$. Define also $G_1 = \{\pi_1(\cdot,\cdot,r) : r \in [0,\overline{\omega}]\}$ and $G_2 = \{\pi_2(\cdot, r) : r \in [0,\overline{\omega}]\}$. The following lemma is useful:

Lemma 2. *Let H be a set of functions mapping $\mathcal{X}$ to $\mathbb{R}$ and let $\Psi_1, \ldots, \Psi_J$ be μ-Lipschitz functions for some $\mu > 0$. Then for any sample S_J of J points $x_1, \ldots, x_J \in \mathcal{X}$, the following inequality holds:*

$$\frac{1}{J} E_\sigma \left[\sup_{h \in H} \sum_{j=1}^{J} \sigma_j (\Psi_j \circ h)(x_j) \right] \leqslant \frac{\mu}{J} E_\sigma \left[\sup_{h \in H} \sum_{j=1}^{J} \sigma_j h(x_j) \right].$$

Proof. See Lemma 14 in Mohri and Medina (2016). □

We now find an upper bound for the right hand side of Lemma 1, which is not expressed in terms of Rademacher complexity and which makes the asymptotic behavior of the term $p(r^*) - p(\widehat{r}_{S_J})$ clear. This will lead to Theorem 4.

Lemma 3. *Let $0 \leqslant V_i^{(1)} \leqslant \overline{\omega} < \infty$. Then $\widehat{\mathfrak{R}}_{S_J}(G) \leqslant \overline{\omega}\sqrt{\frac{2}{J\pi}} + o(J^{-1/2})$.*

Proof. Note that $\widehat{\mathfrak{R}}_{S_J}(G) \leqslant \widehat{\mathfrak{R}}_{S_J}(G_1) + \widehat{\mathfrak{R}}_{S_J}(G_2)$, as the supremum of a sum is less than the sum of suprema. We give upper bounds on both of these terms. For the first term, we have:

$$\begin{aligned}
\widehat{\mathfrak{R}}_{S_J}(G_1) &\equiv E_\sigma \left[\sup_{r \in [0,\overline{\omega}]} \frac{1}{J} \sum_{j=1}^{J} \sigma_j [V_j^{(2)} \mathbb{1}(r \leqslant V_j^{(2)}) + r\mathbb{1}(V_j^{(2)} < r \leqslant V_j^{(1)}) + V_j^{(1)} \mathbb{1}(V_j^{(1)} < r)] \right] \\
&\leqslant \frac{1}{J} E_\sigma \left[\sup_{r \in [0,\overline{\omega}]} \sum_{j=1}^{J} \sigma_j r \right] \\
&= \frac{1}{J} E_\sigma \left[\sup_{r \in \{0,\overline{\omega}\}} \sum_{j=1}^{J} \sigma_i r \right] \\
&\leqslant \frac{2\overline{\omega}\sqrt{\log 2}}{J}
\end{aligned}$$

The first inequality follows from applying Lemma 2 with $\Psi_j(x) \equiv \pi_1(V_j^{(1)}, V_j^{(2)}, x)$ and $h(x) \equiv x$, and the observation that the functions $\Psi_j(r)$ are $1-$Lipschitz for all j. The equality follows because the supremum will always be attained at $r = 0$ (if

$\sum_{j=1}^{J} \sigma_i \leqslant 0$) or at $r = \overline{\omega}$ (if $\sum_{j=1}^{J} \sigma_j > 0$). The final inequality is an application of Massart's lemma (see for e.g. Mohri et al. (2012)).

For the second term, we have

$$\begin{aligned}
\widehat{\mathfrak{R}}_{S_J}(G_2) &\equiv E_\sigma\left[\sup_{r\in[0,\overline{\omega}]} \frac{1}{J}\sum_{j=1}^{J} -\sigma_j V_j^{(1)} \mathbb{1}(V_j^{(1)} < r)\right] \\
&\leqslant \frac{\overline{\omega}}{J} E_\sigma\left[\sup_{r\in[0,\overline{\omega}]} \sum_{j=1}^{J} -\sigma_j \mathbb{1}(V_j^{(1)} < r)\right] \\
&= \frac{\overline{\omega}}{J} E_\sigma\left[\sup_{r\in[0,\overline{\omega}]} \sum_{j=1}^{J} \sigma_j \mathbb{1}(V_j^{(1)} < r)\right] \\
&\leqslant \overline{\omega}\sqrt{\frac{2+2\log J}{J}}.
\end{aligned}$$

The first inequality follows from applying Lemma 2 with $\Psi_j(x) \equiv V_j^{(1)} x$ and $h(x) \equiv \mathbb{1}(V_j^{(1)} < x)$, noting that $\Psi_j(x)$ are $\overline{\omega}$-Lipschitz for all j. The equality follows because the distributions of σ_j and $-\sigma_j$ are identical. Finally, the last inequality follows from Massart's lemma (see Proposition 2 in Mohri and Medina (2016)).

Putting the bounds on $\widehat{\mathfrak{R}}_{S_J}(G_1)$ and $\widehat{\mathfrak{R}}_{S_J}(G_2)$ together, we have:

$$\widehat{\mathfrak{R}}_{S_J}(G) \leqslant \frac{2\overline{\omega}\sqrt{\log 2}}{J} + \overline{\omega}\sqrt{\frac{2+2\log J}{J}}.$$

□

This leads immediately to Theorem 4:

Theorem. *Let $0 \leqslant V_j^{(1)} \leqslant \overline{\omega} < \infty$. For any $\delta > 0$, with probability at least $1-\delta$ over the possible realizations of S_J, it holds that*

$$p(r^*) - p(\widehat{r}_{S_J}) \leqslant \overline{\omega}\left(\frac{8\sqrt{\log 2}}{J} + 4\sqrt{\frac{2+2\log J}{J}} + 6\sqrt{\frac{\log\frac{4}{\delta}}{2J}}\right).$$

Proof. Combine Lemmas 1 and 3. □