Finite-Sample Auction Inference Using Transaction Prices

David M. Kaplan* Xin Liu†

December 12, 2024

Abstract

We provide finite-sample, nonparametric, uniform confidence bands for the bid distribution's quantile function in first-price, second-price, descending, and ascending auctions with symmetric independent private values, when only the transaction price (highest or second-highest bid) is observed. These bands can also be used to construct finite-sample confidence intervals for several counterfactuals and economic quantities. Even with a varying number of bidders, our bands' finite-sample coverage is exact. With a fixed number of bidders, we also derive uniform confidence bands robust to auction-level unobserved heterogeneity. This includes new bounds on the bid quantile function in terms of the transaction price quantile function. We also provide results on computation, median-unbiased quantile estimation, and pointwise quantile inference. Empirically, our new methodology is applied to timber auction data to examine heterogeneity across appraisal value and number of bidders, which helps assess the combination of symmetric independent private values and exogenous participation.

JEL classification: C57

Keywords: ascending; first-price; order statistics; uniform confidence band; unobserved heterogeneity

*Corresponding author, Department of Economics, University of Missouri, kaplandm@missouri.edu. Thanks to Ron Flores for research assistance in drafting some of the initial code and to Carlos Lamarche, CEBA seminar participants including Artem Prokhorov and Rustam Ibragimov, Yu Zhu, and especially Alyssa Carlson for helpful feedback and references.

†School of Economic Sciences, Washington State University, xin.liu1@wsu.edu

1 Introduction

We propose nonparametric uniform confidence bands with exact finite-sample coverage for the quantile function of the bid distribution in the following setting. For each auction, the data include only the number of bidders and the transaction price, which is either the highest bid (in a sealed-bid first-price or descending/Dutch auction), second-highest bid (in a sealed-bid second-price or ascending/English auction), or another order statistic. The number of bidders may vary arbitrarily across auctions. Initially, we assume symmetric independent private values (IPV) with no unobserved heterogeneity, so each bid is an iid draw from the same bid distribution. If there is a binding reserve price, then we can interpret the "bid distribution" as a truncated version of the bid distribution as in equation (1) of Haile and Tamer (2003, p. 7) or (6.20) of Athey and Haile (2007). As discussed later in this section, we extend certain results to allow auction-level unobserved heterogeneity. Overall, our results complement existing work that relaxes some of these assumptions but resorts to asymptotic inference, in contrast to our finite-sample inference.

Economically, the bid distribution may be of interest directly or else for its relationship to the value distribution and related counterfactuals and economic objects of interest. Under the above assumptions, the bid distribution equals the value distribution for ascending auctions; for example, see (1) of Athey and Haile (2002). Alternatively, under (only) the weak assumption that bids do not exceed bidders' values, equation (2) of Haile and Tamer (2003) implies the bid quantile function is a lower bound of the value quantile function; see also Proposition 1 of Grundl and Zhu (2023). The value quantile function is a direct input to various counterfactuals and quantities of economic interest; for example, see Table 1 of Andreyanov and Franguridi (2024). Uniform coverage is important for learning about the shape of the bid distribution, for assessing distributional properties or effects, and for inference on such functionals, for which pointwise confidence bands are not sufficient. In Appendix A, for several counterfactuals and economic quantities, we describe how to construct finite-sample confidence intervals based on our finite-sample uniform confidence bands, in part using bounds from Haile and Tamer (2003) and Armstrong (2013). In the main text, we focus on our main contribution: finite-sample uniform confidence bands.

Theoretically, we use the probability integral transform and related results for order statistics. For intuition, the most basic version can be used to derive a one-sided confidence interval for quantiles of continuous random variable Y with CDF F_Y as follows. The probability integral transform says $F_Y(Y) \sim \text{Unif}(0,1)$, and given iid Y_i over $i = 1, \ldots, n$, the transformed kth order statistic satisfies $F_Y(Y_{n:k}) \sim \text{Beta}(k, n+1-k)$; for example, see Wilks (1962, pp. 236–238). Thus, $\text{P}(Y_{n:k} \leq F_Y^{-1}(\tau)) = \text{P}(F_Y(Y_{n:k}) \leq \tau)$ equals the $\text{Beta}(k, n+1-k)$ CDF evaluated at τ. That is, the confidence interval $[Y_{n:k}, \infty)$ includes the τ-quantile $F_Y^{-1}(\tau)$ with this exactly known probability. This is not an asymptotic approximation but an exact finite-sample result. In our more complex setting, we apply the probability integral transform to the transaction prices and show how to (jointly) use all order statistics of the observed transaction prices (across auctions) and interpolate appropriately to get exact uniform confidence bands for the bid quantile function, even with a varying number of bidders per auction.

When each auction has the same number of bidders, we provide one-sided uniform confidence bands robust to arbitrary auction-level unobserved heterogeneity. That is, the bid distribution may differ across auctions for unobserved reasons, but bids within each auction remain iid (e.g., Athey and Haile, 2007, §6.1.2). These bands are derived using bounds on the bid quantile function in terms of the transaction price quantile function, similar in spirit to the partial identification approach of Haile and Tamer (2003), although they do not provide results for statistical inference (beyond computing bootstrap confidence bands in the empirical example). For first-price auctions, such bounds are provided in CDF form in equation (2) of Armstrong (2013); that is, he shows the bid CDF is bounded above and below by certain transformations of the transaction price CDF. Thus, we can achieve (at least) the desired coverage probability in finite samples by using the corresponding transformations of uniform confidence bands for the transaction price quantile function. We build the latter bands based on results of Goldman and Kaplan (2018a).

For second-price auctions with unobserved heterogeneity, our uniform confidence bands are based on our new bid quantile function bounds. These bounds involve the greatest convex minorant and least concave majorant of a certain beta distribution CDF that depends on the number of bidders. Whereas for first-price auctions this beta CDF is fully convex, for second-price auctions it is partly concave, so the Jensen's inequality argument from Armstrong (2013) cannot be used directly. However, Jensen's can be applied to the convex minorant, or applied in the opposite direction to the concave majorant. Our results apply to general pth-price auctions, including first-price and second-price as special cases.

We apply our new methodology to the well-known U.S. Forest Service timber auction data to assess heterogeneity across numbers of bidders and across appraisal value. Conditioning on a particular number of bidders and a particular range of appraisal values reduces the sample size considerably, so it is important that our uniform confidence bands have finite-sample coverage. Also, it is important that our bands are uniform (not pointwise) to assess whether there exists a bid quantile function that is consistent with the confidence bands from all the subsamples. When analyzing net bids (after subtracting the appraisal value), for each of four appraisal value ranges, we find little heterogeneity across numbers of bidders: besides a few exceptions in the lower tail, the bands have substantial overlap and are consistent with a single net bid quantile function for any number of bidders. Looking across appraisal value ranges, the bands are also consistent with a single net bid quantile function. We also show the value of our method's ability to combine data from auctions with different numbers of bidders: compared to the uniform confidence band for any subsample with a fixed number of bidders, the band using all such subsamples together is much more precise.

In the statistics literature, use of the probability integral transform dates back to at least Fisher (1932), Pearson (1933), and Neyman (1937). More directly, we build on the work of Goldman and Kaplan (2018a). They also manipulate joint probabilities involving transformed order statistics, but in the context of multiple testing of $H_{0r}\colon F(r) = G(r)$ across

$r \in \mathbb{R}$ for two population CDFs $F(\cdot)$ and $G(\cdot)$, as well as the corresponding one-sided and one-sample analogs. Another difference is that they assume a single sample with all order statistics observed, rather than our setting of repeated samples with only one order statistic observed. In our special case with a constant number of bidders in each auction, we derive a transformation that can be applied to a uniform confidence band (based on their work) for the transaction price quantile function. However, with a varying number of bidders, we significantly extend their technical results in ways that may be of independent interest. Also, our key theoretical result with unobserved heterogeneity (Theorem 9) is unrelated to their work.

The auctions literature covers a wide range of assumptions and objects of interest that are not nested with our paper but offer different advantages and disadvantages. The primary distinction of our work is that we characterize our new methods' coverage probability not merely asymptotically but in finite samples. Finite-sample coverage probabilities are also given by Syrgkanis, Tamer, and Ziani (2021, Thms. 10–12), for an identified set given a relaxation of IPV, but with all bids observed and (discrete) finite support assumptions. Finite-sample coverage probabilities are not provided by any work cited below. Further, we do this with only the transaction price observed, unlike all but a few of the below papers, and with either a varying number of bidders or else general auction-level unobserved heterogeneity. Other differences are noted below, too, with more closely related work described first.

These features are now compared with specific papers in the auctions literature, remembering that none of these provide finite-sample coverage probability like we do. Compared with Armstrong (2013), who also assumes only the winning bid of a first-price auction is observed and allows general unobserved heterogeneity, we focus on the bid distribution rather than bidder profit and other functionals, and we also include results for second-price auctions. Menzel and Morganti (2013) also allow for only observing the transaction price (a single order statistic) but focus on asymptotic rates of convergence for nonparametric estimation of (functionals of) the valuation distribution. Compared with Aradillas-López, Gandhi, and Quint (2013), who also only require the transaction price, we only allow correlation of valuations through unobserved heterogeneity (a special case of their setting as noted on their page 494), our object of interest is the full bid distribution (instead of seller and bidder surplus), and we also have results for first-price auctions. Gimenes (2017) similarly uses only the transaction price in an ascending auction, using a quantile regression framework to incorporate observable heterogeneity, but she assumes a fixed number of bidders and no unobserved heterogeneity. Other papers that contribute results for auctions with unobserved heterogeneity require observing more than one bid per auction: for example, Freyberger and Larsen (2022) require two order statistics (but allow unknown number of bidders), Chesher and Rosen (2017) require all final bids in an ascending auction, Compiani, Haile, and Sant'Anna (2020) require all bids in a first-price auction (but endogenously model bidders' decisions to enter the auction, with the help of an instrument), Luo (2020) also requires all bids (in a first-price auction with discretely-supported unobserved heterogeneity), and Haile and Kitamura (2019) also assume all bids are observed (page C5), as do the other approaches they review. Other classic work like that of Guerre, Perrigne, and Vuong (2000) also requires all bids observed, as do more recent extensions like the quantile regression approach of Gimenes and Guerre (2022). Again, most distinctly, none of the above papers characterizes finite-sample coverage probabilities.

Paper structure Section 2 describes the setting and derives a useful identity. Section 3 provides methods and results for uniform confidence bands for the bid quantile function. Section 4 extends our methodology to varying numbers of bidders across auctions. Section 5 discusses interpretation given unobserved heterogeneity, including some new bounds for second-price auctions. Sections 6 and 7 contain empirical results and simulations. Appendix A describes how to translate our uniform confidence bands into confidence intervals for various functionals of economic and counterfactual interest. Appendix B collects the proofs.

Appendix C provides additional methods and results for pointwise inference, for quantiles and interquantile ranges of the bid distribution. Code in R (R Core Team, 2023) implementing our new methods and replicating our results is provided at `https://kaplandm.github.io/`.

2 Setting and an identity

This section describes the setting with a constant number of bidders in each auction, which is generalized to varying numbers of bidders in Section 4 and to the presence of auction-level unobserved heterogeneity in Section 5.

Consider the following setting. Each auction has n bids, with B_i drawn iid for $i = 1, \ldots, n$. The goal is to learn about this bid distribution. However, only the transaction price $W = B_{n:k}$ is observed, where $B_{n:k}$ is the kth order statistic (kth-smallest bid). (The mnemonic for W is "winning bid," which often equals the transaction price, although not in a second-price sealed-bid auction where the transaction price W is the second-highest bid.) For example, in a first-price auction, $k = n$ because the transaction price is the maximum, $W = B_{n:n}$. In a second-price auction, $k = n - 1$ because $W = B_{n:n-1}$. In each auction $j = 1, \ldots, J$, we observe (only) the transaction price W_j. Both n and J are finite; there are no asymptotic approximations.

Assumption A1. The underlying bid B has a continuous distribution; its CDF $F_B(\cdot)$ is continuous and strictly increasing over its support, with inverse $F_B^{-1}(\cdot)$ equal to the quantile function $Q_B(\cdot)$.

Assumption A2. Within each auction, bids $B_i \overset{iid}{\sim} F_B$ over $i = 1, \ldots, n$, and only the transaction price $W = B_{n:k}$ is observed.

Assumption A3. The transaction prices W_j are independent over auctions $j = 1, \ldots, J$.

Assumption A1 implies the transaction price CDF $F_W(\cdot)$ is also continuous and strictly increasing because $F_W(x) = \frac{n!}{(n-k)!(k-1)!} \int_0^{F_B(x)} t^{k-1}(1-t)^{n-k}\,\mathrm{d}t$ as in the proof of Theorem

1(i) of Athey and Haile (2002). Assumption A2 is relaxed in Section 4 to allow the number of bidders to vary across auctions. Assumptions A2 and A3 imply W_j are also identically distributed across $j = 1, \ldots, J$. This rules out dynamic mechanisms like those discussed in Section 9 of Athey and Haile (2007), for example. Under A1–A3, Theorem 1(i) of Athey and Haile (2002) establishes nonparametric identification of $F_B(\cdot)$ for second-price auctions, and their proof applies immediately to first-price auctions, too.

Applying the probability integral transform (Fisher, 1932; Neyman, 1937; Pearson, 1933) with results from Wilks (1962, pp. 236–238), we have

$$\beta_j \equiv F_B(W_j) = F_B(B_{n:k}^{(j)}) \sim F_\beta \equiv \mathrm{Beta}(k, n+1-k), \quad j = 1, \ldots, J. \tag{1}$$

Under A1 and A2 but without A3, the β_j are allowed to be dependent while following the same marginal distribution. If we additionally impose A3, then the β_j are iid across auctions $j = 1, \ldots, J$.

Proposition 1 enables us to learn the unknown bid distribution through the transaction price's distribution and the known transformation $F_\beta(\tau)$.

Proposition 1. *Under A1 and A2, for any quantile index $\tau \in (0,1)$ and any value $q \in \mathbb{R}$,*

$$\begin{aligned} Q_B(\tau) &= Q_W(F_\beta(\tau)), \\ F_B(q) &= F_\beta^{-1}(F_W(q)), \end{aligned} \tag{2}$$

where $F_\beta(\cdot)$ is the CDF of the Beta$(k, n+1-k)$ distribution.

Proposition 1 is at its core the same as other results that have appeared in different notation, going back at least to equation (5) of Athey and Haile (2002). For example, in Lemma 2 of Haile and Tamer (2003), the right-hand side of their (3) is (in their notation) the Beta$(i, n+1-i)$ CDF evaluated at ϕ, so their (4) in Lemma 2 is in our notation equivalent to $F_B(q) = F_\beta^{-1}(F_W(q))$, as in our (2). However, they assume all bids are observed, with a fixed number of bidders, and they only discuss identification and estimation, not inference. The same result is presented in (4.3)–(4.4) of Athey and Haile (2007), who mention asymptotic normality (top p. 3876) but do not further discuss inference, and certainly not finite-sample inference.

In Appendix C, using Proposition 1, we discuss confidence intervals for bid τ-quantile $Q_B(\tau)$ using existing confidence intervals for $Q_W(F_\beta(\tau))$. These can be conservative but with exactly known finite-sample coverage probability, or else achieve the desired confidence level with high-order accuracy. We also discuss confidence intervals for the general interquantile range $Q_B(\tau_2) - Q_B(\tau_1)$.

Proposition 1 also leads to natural estimation of the τ-quantile of B using an estimator of the $F_\beta(\tau)$-quantile of W. Estimators of the latter that are (nearly) median-unbiased remain (nearly) median-unbiased for the former, such as using `quantile(..., type=8)` in R (R Core Team, 2023).

3 Uniform confidence bands

We provide uniform confidence bands for the bid quantile function $Q_B(\cdot)$ with exact finite-sample coverage probability $1-\alpha$, for any fixed number of bidders n and number of auctions J. These are extended to a varying number of bidders in Section 4 and to unobserved heterogeneity in Section 5.

The arguments below rely on the joint distribution extension of (1). This allows the use of all $\{W_{J:r}\}_{r=1}^{J}$ as points on the confidence band, for specially chosen corresponding τ_r, where like $B_{n:k}$ the notation $W_{J:r}$ is the rth order statistic, here the rth-smallest winning bid among the J observed. These J points of the form $(\tau_r, W_{J:r})$ can then be connected by a stair-step shape to avoid any additional coverage error.

We justify our bands from two perspectives. The first more explicitly considers the joint distribution of $\beta_{J:r}$ (for $r = 1, \ldots, J$) based on (1), where like $W_{J:r}$ the notation $\beta_{J:r}$ is the rth order statistic among the $\beta_j \equiv F_B(W_j)$ defined in (1), and equivalently $\beta_{J:r} = F_B(W_{J:r})$. This order statistic approach is generalized in Section 4 to varying numbers of bidders. The second uses Proposition 1 to transform a band for $Q_W(\cdot)$. That approach is generalized in Section 5 to unobserved heterogeneity.

Although there are two justifications, the uniform confidence bands themselves are the same. For an upper one-sided uniform confidence band, define the lower bound function

$$\hat{g}_L(\tau;\tilde{\alpha}) = \begin{cases} -\infty & \text{for } \tau < \tau_1(1-\tilde{\alpha}) \\ W_{J:1} & \text{for } \tau_1(1-\tilde{\alpha}) \le \tau < \tau_2(1-\tilde{\alpha}) \\ \vdots & \vdots \\ W_{J:J} & \text{for } \tau_J(1-\tilde{\alpha}) \le \tau, \end{cases} \tag{3}$$

$$F_\beta(\tau_r(p)) = \xi_{p,J,r} \equiv p\text{-quantile of } \mathrm{Beta}(r, J+1-r), \quad r = 1,\ldots,J, \tag{4}$$

where $\tilde{\alpha}$ is a function of α and J from Method 1 in Goldman and Kaplan (2018a), and more precisely the $-\infty$ can be replaced by the lower bound of the support of B (often zero). More complex than Goldman and Kaplan (2018a), our $\tau_r(p)$ involves nested probability integral transforms that yield two distinct beta distributions: the $\mathrm{Beta}(r, J+1-r)$ from the order statistics $W_{J:r}$ of transaction prices across the J auctions, and the $\mathrm{Beta}(k, n+1-k)$ implicit in $F_\beta(\cdot)$ from the order statistic $B_{n:k}$ across then n bids within each auction.

Consider the following example of (3) and (4). Consider $J = 2$ auctions with $n = 5$ bidders each, and they are first-price auctions, so $k = n$. From (1), $F_\beta(\cdot)$ is the CDF of the $\mathrm{Beta}(5,1)$ distribution, $F_\beta(x) = x^5$, with inverse $F_\beta^{-1}(x) = x^{1/5}$. From (4), $F_\beta(\tau_r(1-\tilde{\alpha}))$ is $\xi_{1-\tilde{\alpha},2,r}$, the $(1-\tilde{\alpha})$-quantile of $\mathrm{Beta}(r, 3-r)$, so $\tau_1(1-\tilde{\alpha}) = \xi_{1-\tilde{\alpha},2,1}^{1/5}$ and $\tau_2(1-\tilde{\alpha}) = \xi_{1-\tilde{\alpha},2,2}^{1/5}$. From (3), given $\tilde{\alpha}$, the function $\hat{g}_L(\cdot;\tilde{\alpha})$ is flat from $\tau = 0$ to $\tau = \tau_1(1-\tilde{\alpha})$ at value $-\infty$, then jumps up to value $W_{2:1}$ and is again flat from $\tau_1(1-\tilde{\alpha})$ to $\tau_2(1-\tilde{\alpha})$, and finally jumps up to value $W_{2:2}$ and is flat until $\tau = 1$. Equivalent to (8) of Goldman and Kaplan (2018a), the value $\tilde{\alpha}$ is calibrated such that $1-\alpha = \mathrm{P}(U_{2:1} \le \xi_{1-\tilde{\alpha},2,1} \text{ and } U_{2:2} \le \xi_{1-\tilde{\alpha},2,2})$, where $U_{2:1}$ and $U_{2:2}$ are the order statistics from two iid $\mathrm{Unif}(0,1)$ draws (that is, $U_{2:1}$ is the lesser of the two draws and $U_{2:2}$ the greater). In practice, simulation of the $(U_{2:1}, U_{2:2})$ distribution is avoided by using the analytic formula in Fact 6 of Goldman and Kaplan (2018a).

Similar to (3), define upper bound function

$$\hat{g}_U(\tau;\tilde{\alpha}) = \begin{cases} W_{J:1} & \text{for } \tau < \tau_1(\tilde{\alpha}) \\ \vdots & \vdots \\ W_{J:J} & \text{for } \tau_{J-1}(\tilde{\alpha}) \le \tau < \tau_J(\tilde{\alpha}) \\ \infty & \text{for } \tau_J(\tilde{\alpha}) \le \tau, \end{cases} \tag{5}$$

with the same $\tau_r(p)$ from (4) and the same $\tilde{\alpha}$ from Method 1 of Goldman and Kaplan (2018a).

This $\hat{g}_U(\cdot;\tilde{\alpha})$ can be adapted to the setting of Haile and Tamer (2003) as follows. Their Lemma 3 says that in an ascending auction, the second-highest valuation does not exceed the highest bid plus the minimum bid increment. Thus, if we interpret the latter as the second-highest order statistic and construct our $\hat{g}_U(\cdot;\tilde{\alpha})$ accordingly, it will have uniform coverage probability of at least $1-\alpha$.

3.1 First perspective: order statistics

Some intuition is now given for the argument from the first perspective, with details in the proof of Theorem 2. Consider $\hat{g}_L(\cdot;\tilde{\alpha})$. First, because of the stair-step shape of the band combined with the non-decreasing property of $Q_B(\cdot)$, the uniform coverage probability of $\hat{g}_L(\cdot;\tilde{\alpha})$ reduces to the joint coverage probability over J points. Then, taking $F_B(\cdot)$ then transforms the probability to be in terms of the $\beta_{J:r}$ order statistics. Altogether,

$$\begin{aligned} &\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}) \le Q_B(\tau) \text{ for all } \tau \in [0,1]\big) \\ &= \mathrm{P}\big(W_{J:1} \le Q_B(\tau_1(1-\tilde{\alpha})) \text{ and } \ldots \text{ and } W_{J:J} \le Q_B(\tau_J(1-\tilde{\alpha}))\big) \\ &= \mathrm{P}\big(\overbrace{F_B(W_{J:1})}^{\beta_{J:1}} \le \tau_1(1-\tilde{\alpha}) \text{ and } \ldots \text{ and } \overbrace{F_B(W_{J:J})}^{\beta_{J:J}} \le \tau_J(1-\tilde{\alpha})\big). \end{aligned}$$

We know $\beta_j \stackrel{iid}{\sim} \mathrm{Beta}(k, n+1-k)$ from (1), so we also know the joint distribution of the corresponding order statistics. Thus, we can solve for the $\tilde{\alpha}$ that sets the above probability equal to $1-\alpha$ exactly.

Theorem 2. *Let Assumptions A1–A3 hold. Consider the following one-sided uniform confidence bands for the bid quantile function* $Q_B(\cdot)$*. Lower and upper bound functions* $\hat{g}_L(\cdot;\tilde{\alpha})$ *and* $\hat{g}_U(\cdot;\tilde{\alpha})$ *are from* (3) *and* (5)*, with* $\tau_r(\cdot)$ *defined in* (4)*. Given the* $\tilde{\alpha}$ *from Method 1 of Goldman and Kaplan (2018a), finite-sample uniform coverage probability is exactly* $1-\alpha$*:*

$$\begin{aligned} &\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}) \le Q_B(\tau) \textit{ for all } \tau \in [0,1]\big) = 1-\alpha, \\ &\mathrm{P}\big(Q_B(\tau) \le \hat{g}_U(\tau;\tilde{\alpha}) \textit{ for all } \tau \in [0,1]\big) = 1-\alpha. \end{aligned} \tag{6}$$

For a two-sided uniform confidence band, the same $\hat{g}_L(\cdot;\tilde{\alpha})$ and $\hat{g}_U(\cdot;\tilde{\alpha})$ from (3) and (5) can be used, but with $\tilde{\alpha}$ calibrated to exact two-sided coverage as in Method 2 of Goldman and Kaplan (2018a). The intuition for the proof is the same as for the one-sided bands.

Theorem 3. *Let Assumptions A1–A3 hold. Consider the following two-sided uniform confidence band for the bid quantile function* $Q_B(\cdot)$*. Lower and upper bound functions* $\hat{g}_L(\cdot;\tilde{\alpha})$ *and* $\hat{g}_U(\cdot;\tilde{\alpha})$ *are from* (3) *and* (5)*, with* $\tau_r(\cdot)$ *defined in* (4)*. Given the* $\tilde{\alpha}$ *from Method 2 of Goldman and Kaplan (2018a), finite-sample uniform coverage probability is exactly* $1-\alpha$*:*

$$\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}) \le Q_B(\tau) \le \hat{g}_U(\tau;\tilde{\alpha}) \textit{ for all } \tau \in [0,1]\big) = 1-\alpha. \tag{7}$$

3.2 Second perspective: transformed confidence band

From the second perspective, given any uniform confidence band for $Q_W(\cdot)$, we can use Proposition 1 to construct a uniform confidence band for $Q_B(\cdot)$ with the same properties. Consider the band $[\hat{L}(\cdot), \hat{U}(\cdot)]$ for $Q_W(\cdot)$, where the "hats" indicate they are random functions (computed from the sample), and one-sided bands are included by allowing $\hat{L}(\cdot) = -\infty$ or $\hat{U}(\cdot) = \infty$. The same uniform coverage probability is attained by the band $[\hat{L}(F_\beta(\cdot)), \hat{U}(F_\beta(\cdot))]$ for $Q_B(\cdot)$: using $Q_W(F_\beta(\tau)) = Q_B(\tau)$ from Proposition 1,

$$\begin{aligned} &\mathrm{P}\big(\hat{L}(F_\beta(\tau)) \le Q_B(\tau) \le \hat{U}(F_\beta(\tau)) \text{ for all } \tau \in [0,1]\big) \\ &= \mathrm{P}\big(\hat{L}(F_\beta(\tau)) \le Q_W(F_\beta(\tau)) \le \hat{U}(F_\beta(\tau)) \text{ for all } \tau \in [0,1]\big) \\ &= \mathrm{P}\big(\hat{L}(t) \le Q_W(t) \le \hat{U}(t) \text{ for all } t \in [0,1]\big), \end{aligned} \tag{8}$$

where the last equality comes from the fact that $F_\beta(\cdot)$ is a continuous increasing function from $F_\beta(0) = 0$ to $F_\beta(1) = 1$.

Theorem 4. *Under Assumptions A1–A3, with $F_\beta(\cdot)$ the CDF of the Beta$(k, n+1-k)$ distribution as in* (1)*, the band $[\hat{L}(F_\beta(\cdot)), \hat{U}(F_\beta(\cdot))]$ for $Q_B(\cdot)$ has the same uniform coverage probability as the band $[\hat{L}(\cdot), \hat{U}(\cdot)]$ for $Q_W(\cdot)$. Similarly, the band $[F_\beta^{-1}(\hat{L}(\cdot)), F_\beta^{-1}(\hat{U}(\cdot))]$ for $F_B(\cdot)$ has the same uniform coverage probability as the band $[\hat{L}(\cdot), \hat{U}(\cdot)]$ for $F_W(\cdot)$.*

The final argument is to show that our uniform confidence bands in Theorems 2 and 3 are transformations of uniform confidence bands for $Q_W(\cdot)$. These latter bands follow readily from results of Goldman and Kaplan (2018a) and can be computed easily using their code. They did not formally state these bands because it seems they thought such bands were proposed by Buja and Rolke (2006, §5.1) and Aldor-Noiman, Brown, Buja, Rolke, and Stine (2013, §2.1), which are closely related but instead "null bands" for the purpose of goodness-of-fit testing based on Q–Q plots.

The uniform confidence bands for $Q_W(\cdot)$ based on iid draws of $W_j \overset{iid}{\sim} F_W$ are as follows. Using the notation of (4),

$$\hat{w}_L(\tau; \tilde{\alpha}) = \begin{cases} -\infty & \text{for } \tau < \xi_{1-\tilde{\alpha},J,1} \\ W_{J:1} & \text{for } \xi_{1-\tilde{\alpha},J,1} \le \tau < \xi_{1-\tilde{\alpha},J,2} \\ \vdots & \vdots \\ W_{J:J} & \text{for } \xi_{1-\tilde{\alpha},J,J} \le \tau, \end{cases} \tag{9}$$

$$\hat{w}_U(\tau; \tilde{\alpha}) = \begin{cases} W_{J:1} & \text{for } \tau < \xi_{\tilde{\alpha},J,1} \\ \vdots & \vdots \\ W_{J:J} & \text{for } \xi_{\tilde{\alpha},J,J-1} \le \tau < \xi_{\tilde{\alpha},J,J} \\ \infty & \text{for } \xi_{\tilde{\alpha},J,J} \le \tau. \end{cases} \tag{10}$$

One-sided bands use Method 1 of Goldman and Kaplan (2018a) to get $\tilde{\alpha}$ (a function of confidence level $1-\alpha$ and sample size J, which in their notation is n), whereas two-sided bands use the $\tilde{\alpha}$ from their Method 2.

Theorem 5. *Given $W_j \overset{iid}{\sim} F_W$ for $j = 1, \ldots, J$ and continuous, strictly-increasing CDF $F_W(\cdot)$, the uniform confidence bands in* (9) *and* (10) *have exact finite-sample coverage probability.*

For completion, we show that the bands from Theorems 2 and 3 are equivalent to applying the transformation from Theorem 4 to the corresponding band for $Q_W(\cdot)$ from Theorem 5. In fact, this is immediate from the definitions: given $F_\beta(\tau_r(p)) = \xi_{p,J,r}$ in (4), $\hat{g}_L(\tau; \tilde{\alpha}) = \hat{w}_L(F_\beta(\tau); \tilde{\alpha})$, and similarly $\hat{g}_U(\tau; \tilde{\alpha}) = \hat{w}_U(F_\beta(\tau); \tilde{\alpha})$. For example, $\xi_{1-\tilde{\alpha},J,1} \leq F_\beta(\tau) < \xi_{1-\tilde{\alpha},J,2}$ if and only if $F_\beta^{-1}(\xi_{1-\tilde{\alpha},J,1}) \leq \tau < F_\beta^{-1}(\xi_{1-\tilde{\alpha},J,2})$, which is equivalent to $\tau_1(1-\tilde{\alpha}) \leq \tau < \tau_2(1-\tilde{\alpha})$ as in (3).

Theorem 5 does not depend on Assumption A2, so it applies to some settings even with unobserved heterogeneity, which we use in Corollary 11 in Section 5.

4 Generalization to varying numbers of bidders

In this section, we extend our setting and results to allow the number of bids to vary over auctions. Computationally, we cannot use the values of $\tilde{\alpha}$ calibrated by Goldman and Kaplan (2018a) and instead need to simulate such values ourselves. Theoretically, up to an arbitrarily small simulation error, our confidence bands still achieve exact finite-sample uniform coverage probability. There is no corollary of Proposition 1 because it does not make sense to talk of "the" marginal distribution of W: the distribution of W_j depends on the number of bidders in auction j, which varies over $j = 1, \ldots, J$. Still, we can use the order statistics $W_{J:r}$ to construct confidence intervals and bands.

4.1 Setting

Consider the following generalization of our setting. There are now n_j bids in auction j, $j = 1, \ldots, J$. We observe the transaction price $W_j \equiv B_{n_j:k_j}$, with $k_j = n_j$ for first-price auctions or $k_j = n_j - 1$ for second-price. (The type of auction may also vary across j, although this seems uncommon.)

We maintain A1 and A3 but replace A2 with the more general A4.

Assumption A4. Within each auction $j = 1, \ldots, J$, the bids $B_i^{(j)}$ are iid over $i = 1, \ldots, n_j$.

Assumptions A3 and A4 imply the W_j are independent but not identically distributed (inid), and similarly the β_j are now inid:

$$\beta_j \equiv F_B(W_j) = F_B(B_{n_j:k_j}^{(j)}) \overset{inid}{\sim} \text{Beta}(k_j, n_j + 1 - k_j), \quad j = 1, \ldots, J. \tag{11}$$

The Bapat–Beg theorem (Bapat and Beg, 1989) provides the theoretical formula for the joint distribution of $\beta_{J:r}$, $r = 1, \ldots, J$, but it is computationally intractable, as noted by Glueck, Karimpour-Fard, Mandel, Hunter, and Muller (2008).

Fortunately, given (11), it is straightforward to simulate the joint distribution of the $\beta_{J:r}$ over $r = 1, \ldots, J$. First, we simulate many sets of independent $\{\beta_j\}_{j=1}^J$ with each β_j drawn from the known distribution in (11). Second, we sort each set of $\{\beta_j\}_{j=1}^J$ to get corresponding order statistics $\{\beta_{J:r}\}_{r=1}^J$. Third, the simulated probability of an event is the proportion of sets in which the event occurs. The simulation error can be made arbitrarily small by simulating a large enough number of sets.

4.2 Uniform confidence bands

Our one-sided uniform confidence bands have lower and upper functions with the same structure as in (3) and (5), but with different τ_r and $\tilde{\alpha}$. Generalizing (4),

$$\tau_r(p) \equiv p\text{-quantile of } \beta_{J:r}. \tag{12}$$

Previously, we had $\beta_j \overset{iid}{\sim} \text{Beta}(k, n + 1 - k)$ with CDF $F_\beta(\cdot)$, so the probability integral transform gave the analytic form $F_\beta(\beta_{J:r}) \sim \text{Beta}(r, J + 1 - r)$, implying that the p-quantile of $\beta_{J:r}$ is $F_\beta^{-1}(\cdot)$ applied to the p-quantile of $\text{Beta}(r, J + 1 - r)$, as in (4). Although (12) lacks such a formula, the right-hand side is still easy to simulate. The values $\tilde{\alpha}_L$ and $\tilde{\alpha}_U$

respectively solve

$$1-\alpha = \mathrm{P}(\beta_{J:1} \le \tau_1(1-\tilde{\alpha}_L) \text{ and } \ldots \text{ and } \beta_{J:J} \le \tau_J(1-\tilde{\alpha}_L)), \tag{13}$$

$$1-\alpha = \mathrm{P}(\beta_{J:1} \ge \tau_1(\tilde{\alpha}_U) \text{ and } \ldots \text{ and } \beta_{J:J} \ge \tau_J(\tilde{\alpha}_U)). \tag{14}$$

In practice, (13) and (14) can be simulated following the steps in Section 4.1, as implemented in our provided code that computes uniform confidence bands.

Theorem 6. *Let Assumptions A1, A3, and A4 hold. Consider the following one-sided uniform confidence bands for the bid quantile function* $Q_B(\cdot)$*. Lower and upper bound functions* $\hat{g}_L(\cdot;\tilde{\alpha}_L)$ *and* $\hat{g}_U(\cdot;\tilde{\alpha}_U)$ *are from* (3) *and* (5)*, with* $\tau_r(\cdot)$ *defined in* (12)*,* $\tilde{\alpha}_L$ *satisfying* (13)*, and* $\tilde{\alpha}_U$ *satisfying* (14)*. Then, the finite-sample uniform coverage probability is exactly* $1-\alpha$*:*

$$\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}_L) \le Q_B(\tau) \textit{ for all } \tau \in [0,1]\big) = \mathrm{P}\big(Q_B(\tau) \le \hat{g}_U(\tau;\tilde{\alpha}_U) \textit{ for all } \tau \in [0,1]\big) = 1-\alpha.$$

We also construct a two-sided uniform confidence band, where the lower and upper bound functions are the same as above but with value $\tilde{\alpha}_T$ solving

$$\mathrm{P}(\tau_1(\tilde{\alpha}_T) \le \beta_{J:1} \le \tau_1(1-\tilde{\alpha}_T) \text{ and } \ldots \text{ and } \tau_J(\tilde{\alpha}_T) \le \beta_{J:J} \le \tau_J(1-\tilde{\alpha}_T)) = 1-\alpha. \tag{15}$$

Theorem 7. *Let Assumptions A1, A3, and A4 hold. Consider the following two-sided uniform confidence band for the bid quantile function* $Q_B(\cdot)$*. Lower and upper bound functions* $\hat{g}_L(\cdot;\tilde{\alpha}_T)$ *and* $\hat{g}_U(\cdot;\tilde{\alpha}_T)$ *are from* (3) *and* (5)*, with* $\tau_r(\cdot)$ *defined in* (12) *and* $\tilde{\alpha}_T$ *satisfying* (15)*. Then, the finite-sample uniform coverage probability is exactly* $1-\alpha$*:*

$$\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}_T) \le Q_B(\tau) \le \hat{g}_U(\tau;\tilde{\alpha}_T) \textit{ for all } \tau \in [0,1]\big) = 1-\alpha.$$

4.3 Median-unbiased quantile estimation

The relation in (12) also leads to a finite-sample median-unbiased estimator of certain quantiles of B.

Theorem 8. *Let Assumptions A1, A3, and A4 hold. Then, for each $r = 1, \ldots, J$, $W_{J:r}$ is a finite-sample median-unbiased estimator of the $[\text{median}(\beta_{J:r})]$-quantile of the bid distribution.*

4.4 Alternative uniform bands with faster computation

In practice, at least for larger J, it may be desirable to more quickly solve for $\tilde{\alpha}$ from (13), (14), or (15). When searching for $\tilde{\alpha}$, these rely on the $\tau_r(\cdot)$ from (12) that must also be simulated. Below, we suggest an alternative $\tau_r(\cdot)$ that computes more quickly and maintains exact $1-\alpha$ uniform coverage probability, but with the cost of having less even marginal coverage probabilities.

Consider the following "average CDF" of the β_j. Letting $F_{\beta_j}(\cdot)$ be the CDF of $\beta_j \sim \text{Beta}(k_j, n_j + 1 - k_j)$ from (11), define

$$\bar{F}_\beta(x) \equiv \frac{1}{J}\sum_{j=1}^{J} F_{\beta_j}(x).$$

In the special case of Section 3 with a constant number of bidders $n_j = n$, the β_j are iid, so $\bar{F}_\beta(\cdot) = F_\beta(\cdot)$, the $\text{Beta}(k, n+1-k)$ CDF from (1). In that case, $\bar{F}_\beta(\beta_{J:r}) \sim \text{Beta}(r, J+1-r)$, so the p-quantile of $\beta_{J:r}$ is $\bar{F}_\beta^{-1}(\cdot)$ applied to the p-quantile of $\text{Beta}(r, J+1-r)$; this is the $\tau_r(p)$ defined in (12). With varying n_j, we can still replace (12) with

$$\tau_r(p) = \bar{F}_\beta^{-1}\big(p\text{-quantile of Beta}(r, J+1-r)\big).$$

This does not yield the exact p-quantile of $\beta_{J:r}$, so the *pointwise* coverage probabilities are not identical. That is, for example, $\text{P}(\beta_{J:1} \le \tau_1(1-\tilde{\alpha})) \ne \text{P}(\beta_{J:2} \le \tau_2(1-\tilde{\alpha}))$, so the pointwise coverage probability $\text{P}(\hat{g}_L(\tau;\tilde{\alpha}) \le Q_B(\tau))$ differs across τ more than before. However, in the sense of Theorems 6 and 7, the *uniform* coverage probability remains exact by choosing $\tilde{\alpha}_L$, $\tilde{\alpha}_U$, or $\tilde{\alpha}_T$ as in (13), (14), or (15), respectively. Although determining $\tilde{\alpha}$ still requires simulation, using $\bar{F}_\beta(\cdot)$ does not, which makes computation faster.

5 Auction-level unobserved heterogeneity

We now provide uniform confidence bands for $Q_B(\cdot)$ under the following type of unobserved heterogeneity. Auction-level variable U is unobserved in the data (or unused) but observed by bidders, with U_j sampled iid across auctions $j = 1, \ldots, J$. This U can arbitrarily affect the bid distribution, violating Assumption A2. Instead, we assume A1 and A2 hold for the conditional distribution of B given U, but interest remains in the unconditional distribution of B and its quantile function $Q_B(\cdot)$. We assume n bidders throughout this section.

The unconditional bid distribution remains of economic interest even with unobserved heterogeneity, as argued in other papers. For example, also with unobserved heterogeneity, Armstrong (2013) derives bounds on economically relevant features of the unconditional value distribution. Specifically, in Section 2.1 he considers the mean, which (among other interpretations) is the expected "buyer" surplus from counterfactually allocating the good by a random lottery, and in Section 2.2 he considers the expectation of the highest bidder's value, which is the expected total surplus. For details and other economic motivation, see the beginnings of Sections 2.1 and 2.2 as well as Section 2.4 of Armstrong (2013). Also, for first-price auctions with unobserved heterogeneity, Armstrong (2013) provides bounds for the mean value and mean highest value in terms of the unconditional mean bid $\mathrm{E}(B)$ and mean highest bid $\mathrm{E}(B_{n:n})$. Confidence intervals for both of those can be derived from our uniform confidence bands for $Q_B(\cdot)$ that we develop below.

5.1 Bounds on the bid distribution

First, we provide new bounds for the bid distribution under unobserved heterogeneity for general k. Specifically, the unconditional bid CDF is bounded by transformations of the observed transaction price's CDF, and similarly for the quantile functions. In the special case $k = n$ (first-price), our bounds match those of Armstrong (2013).

Our approach uses bounds of $F_\beta(\cdot)$ to which Jensen's inequality can be applied. With

$k = n$ for first-price auctions, $F_\beta(\cdot)$ is convex, so Jensen's inequality applies directly, as shown by Armstrong (2013). However, with $k < n$, $F_\beta(\cdot)$ is not convex, so Jensen's inequality does not directly apply. Instead, we apply Jensen's to the greatest convex minorant of $F_\beta(\cdot)$, which we denote $g(\cdot)$. That is, $g(\cdot)$ is convex, and $g(\cdot) \le F_\beta(\cdot)$ ("minorant"), and it is weakly above any other function with those two properties. (Recall $F_\beta \colon [0,1] \mapsto [0,1]$, so $g \colon [0,1] \mapsto [0,1]$, too.) Conditional on $U = u$, Proposition 1 gives $F_\beta\big(F_{B|U}(q \mid U = u)\big) = F_{W|U}(q \mid U = u)$, so

$$\begin{aligned} F_W(q) &= \mathrm{E}[\overbrace{F_{W|U}(q \mid U)}^{\text{use Proposition 1}}] = \mathrm{E}\big[\overbrace{F_\beta\big(F_{B|U}(q \mid U)\big)}^{\text{use } F_\beta(\cdot) \ge g(\cdot)}\big] \\ &\ge \overbrace{\mathrm{E}\big[g\big(F_{B|U}(q \mid U)\big)\big]}^{\text{use Jensen's}} \ge g\big(\mathrm{E}[F_{B|U}(q \mid U)]\big) = g\big(F_B(q)\big). \end{aligned} \tag{16}$$

Letting $h(\cdot)$ be the least concave majorant of $F_\beta(\cdot)$ yields a bound in the other direction: the first two equalities are the same as above, and then

$$\mathrm{E}\big[\overbrace{F_\beta\big(F_{B|U}(q \mid U)\big)}^{\text{use } F_\beta(\cdot) \le h(\cdot)}\big] \le \overbrace{\mathrm{E}\big[h\big(F_{B|U}(q \mid U)\big)\big]}^{\text{use Jensen's}} \le h\big(\mathrm{E}[F_{B|U}(q \mid U)]\big) = h\big(F_B(q)\big). \tag{17}$$

Theorem 9 formally states the bounds, after which Corollary 10 characterizes the $g(\cdot)$ and $h(\cdot)$ functions in more detail.

Theorem 9. *Let Assumptions A1–A3 hold when replacing F_B with the conditional distribution of B given U, and assume the U_j are iid over auctions $j = 1, \ldots, J$. Assume $n \ge 3$ bidders. Let $F_\beta(\cdot)$ and $F'_\beta(\cdot)$ respectively denote the CDF and PDF of the Beta$(k, n+1-k)$ distribution as in Proposition 1. Over the shared domain $[0,1]$, let $g(\cdot)$ be the greatest convex minorant of $F_\beta(\cdot)$, and let $h(\cdot)$ be the least concave majorant of $F_\beta(\cdot)$, with inverse functions $g^{-1}(\cdot)$ and $h^{-1}(\cdot)$. Then, for any $1 \le k \le n$,*

$$h^{-1}\big(F_W(q)\big) \le F_B(q) \le g^{-1}\big(F_W(q)\big) \quad \textit{for all } q \in \mathbb{R},$$

$$Q_W\big(g(\tau)\big) \le Q_B(\tau) \le Q_W\big(h(\tau)\big) \quad \textit{for all } \tau \in [0,1].$$

Corollary 10. *Theorem 9 implies the following.*

(i) *For $k = n$ (first-price auctions), $g(x) = F_\beta(x) = x^n$ and $h(x) = x$, so $g^{-1}(p) = p^{1/n}$ and $h^{-1}(p) = p$. Thus, $F_W(q) \le F_B(q) \le [F_W(q)]^{1/n}$ for all $q \in \mathbb{R}$, and $Q_W(\tau^n) \le Q_B(\tau) \le Q_W(\tau)$ for all $\tau \in [0,1]$.*

(ii) *For $k < n$ (and $k > 1$), including second-price auctions ($k = n-1$), define tangency point t_{1n} to solve $F'_\beta(t_{1n}) = [1 - F_\beta(t_{1n})]/(1 - t_{1n})$, and define tangency point t_{2n} to solve $F'_\beta(t_{2n}) = F_\beta(t_{2n})/t_{2n}$. Then,*

$$g(x) = \begin{cases} F_\beta(x) & \text{if } x \le t_{1n}, \\ F_\beta(t_{1n}) + (x - t_{1n})\frac{1-F_\beta(t_{1n})}{1-t_{1n}} & \text{if } x \ge t_{1n}, \end{cases}$$

$$h(x) = \begin{cases} x\frac{F_\beta(t_{2n})}{t_{2n}} & \text{if } x \le t_{2n}, \\ F_\beta(x) & \text{if } x \ge t_{2n}. \end{cases}$$

Corollary 10(i) shows that our bounds include the first-price CDF bounds of Armstrong (2013) as a special case.

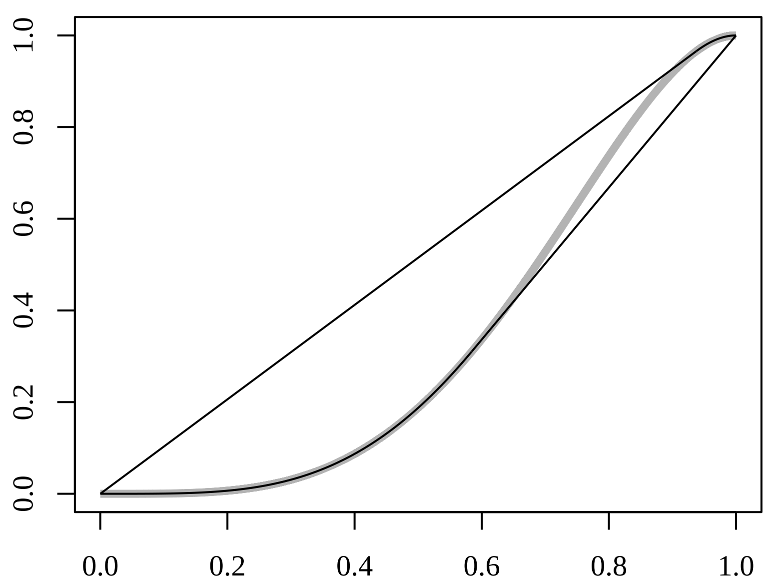

Figure 1: Illustration of Corollary 10(ii), $n = 5$ bidders, $k = 4$ (second-price).

Corollary 10(ii) is illustrated in Figure 1. The thick gray line is $F_\beta(\cdot)$, here with $n = 5$ and $k = 4$ (second-price auction with five bidders). The lower black line is $g(\cdot)$, and the upper black line is $h(\cdot)$. For second-price auctions ($k = n-1$), the tangency points t_{1n} and t_{2n} both increase toward one as n increases. For example, with $n \in \{3, \ldots, 8\}$, the corresponding t_{1n} values are respectively $\{0.250, 0.462, 0.584, 0.662, 0.715, 0.754\}$,[1] and the corresponding t_{2n} values are respectively $\{0.750, 0.889, 0.937, 0.960, 0.972, 0.980\}$.[2] Thus, generally $g(\cdot)$ is much closer to $F_\beta(\cdot)$ than is $h(\cdot)$, so the bounds depending on $g(\cdot)$ (the upper bound of $F_B(\cdot)$ and lower bound of $Q_B(\cdot)$) are tighter than the bounds depending on $h(\cdot)$.

5.2 Uniform confidence bands

Corollary 11 uses our quantile function bounds to derive one-sided uniform confidence bands that remain valid with unobserved heterogeneity. That is, although coverage probability may not exactly equal the desired $1-\alpha$, it is $1-\alpha$ or higher in finite samples. In practice, one-sided finite-sample $1-\alpha$ uniform confidence bands for $F_W(\cdot)$ can be computed with the user-provided `distcomp` command in Stata (Kaplan, 2019) or the `GK_dist_inf.R` R code provided by Goldman and Kaplan (2018a); switching the axes yields the corresponding uniform confidence band for $Q_W(\cdot)$. This is then transformed per Theorem 9, yielding a one-sided uniform confidence band for (a bound for) $Q_B(\cdot)$. An example of this computation is in our provided empirical code.

The finite-sample validity of the uniform confidence bands for $Q_W(\cdot)$ is from Theorem 5, which requires iid transaction prices W_j. This is implied by the assumptions of Theorem 9. Specifically, this is implied by the combination of iid U_j and the conditional version of Assumption A2. Without the fixed number of bidders n as in A2, the W_j would be independent but not identical, so Theorem 5 could not be applied.

Corollary 11. *Theorem 9 implies the following for uniform confidence bands. Confidence functions $\hat{w}_L(\cdot;\tilde{\alpha})$ and $\hat{w}_U(\cdot;\tilde{\alpha})$ are from (9) and (10), which are one-sided uniform confidence bands for the transaction price quantile function. For any $1 \le k \le n$, the following*

[1] `for (n in 3:8) print(uniroot(function(t1) dbeta(t1,n-1,2)-(1-pbeta(t1,n-1,2))/(1-t1), c(0.01,0.99))$root)`

[2] `for (n in 3:8) print(uniroot(function(t2) dbeta(t2,n-1,2)-pbeta(t2,n-1,2)/t2, c(0.01, 0.99))$root)`

transformations have at least $1-\alpha$ finite-sample uniform coverage probability:

$$1-\alpha \leq \mathrm{P}\big(\hat{w}_L(g(\tau);\tilde{\alpha}) \leq Q_B(\tau) \text{ for all } \tau \in [0,1]\big),$$
$$1-\alpha \leq \mathrm{P}\big(\hat{w}_U(h(\tau);\tilde{\alpha}) \geq Q_B(\tau) \text{ for all } \tau \in [0,1]\big).$$

Similarly, if instead $\hat{w}_L(\cdot;\tilde{\alpha})$ and $\hat{w}_U(\cdot;\tilde{\alpha})$ jointly form a two-sided uniform confidence band for the transaction price quantile function, then the transformed two-sided uniform confidence band for $Q_B(\cdot)$ satisfies

$$1-\alpha \leq \mathrm{P}\big(\hat{w}_L(g(\tau);\tilde{\alpha}) \leq Q_B(\tau) \leq \hat{w}_U(h(\tau);\tilde{\alpha}) \text{ for all } \tau \in [0,1]\big).$$

6 Empirical illustration

To illustrate our new methodology, we use the U.S. Forest Service timber auction data from Aradillas-López, Gandhi, and Quint (2013), who in their Section 4.1 provide many details about the data related to common-value uncertainty, private values, contract expiration horizon, and other considerations. Although they find some evidence of correlated values, we focus on our confidence bands that assume independence, while also comparing with our confidence bands that are robust to auction-level unobserved heterogeneity, which can generate unconditionally correlated values. (To be clear: we still assume independent *private* values, that is, independent values conditional on the unobserved heterogeneity.) As described by Aradillas-López, Gandhi, and Quint (2013, pp. 503–4), their data have been filtered to include auctions between 1982 and 1990 in Zone 2 of Region 6 (i.e., western part of Washington and Oregon) with between two and 11 bidders (see their footnote 26). Also, only "scaled sales" are included, so the bids are in dollars per thousand board-feet of actually harvested timber. Like them (p. 503), given that these are English (ascending) auctions, we treat the transaction price as the order statistic $B_{n:n-1}$, which is not exactly true but a close approximation. The number of bidders n_j is observed in each auction j.

We apply our methodology to examine heterogeneity across auctions with different numbers of bidders and different appraisal values, and to learn about the underlying bid distributions based only on the transaction prices. Heterogeneity across the number of bidders n relates to the concept of exogenous participation or exogenous variation in n, which says that the value distribution does not depend on n; for example, see Definition 2.2 of Athey and Haile (2007) or Definition 1 of Aradillas-López, Gandhi, and Quint (2013). Exogenous participation is an important and common assumption (Guerre, Perrigne, and Vuong, 2009, §3.1), so it is valuable to assess whenever possible. Exogenous participation combined with symmetric IPV yields a testable implication: in our setting, that the bid distribution implied by the transaction price distribution is the same given any number of bidders n; see (the proof of) Theorem 1(ii)(b) of Athey and Haile (2002). This is exactly what our uniform confidence bands assess when constructed separately for each n. If the bands do not overlap, then either exogenous participation or symmetric IPV or both are violated; for example, there could be symmetric IPV but with endogenous participation (Athey and Haile, 2007, §6.3), or there could be exogenous participation but with affiliated or common values (Athey and Haile, 2007, §2.1). To facilitate comparison across auctions with different appraisal values, we consider the "net bid" after subtracting the appraisal value. The net bids are all positive.

We use the method in Theorem 3 when the number of bidders is constant and the method in Theorem 7 for pooled analysis including varying numbers of bidders. We use 10^4 beta draws for the full pooled data analysis of $J = 733$ auctions (observations) and 10^5 draws otherwise. We plot two-sided uniform confidence bands with a 90% confidence level. When multiple bands are shown in the same plot, we have used a Bonferroni adjustment to ensure joint coverage over all bands; for example, with four simultaneous bands, each has uniform confidence level $1 - (0.1/4) = 0.975$. Code in R (R Core Team, 2023) to replicate our results is provided.

Total computation time is under four minutes. For all the bands with a fixed number of bidders, including those accounting for unobserved heterogeneity, total computation time is under one second because no simulation is required. For the bands in Figure 4 with a varying number of bidders, computation time is around 40 seconds per band, so 3.5 minutes for the whole figure.

Figure 2 assesses heterogeneity in net bid distributions across auctions with different numbers of bidders, conditioning on the appraisal value. The four plots correspond to data for auctions in four separate ranges of appraisal value: $[10, 30]$, $[30, 50]$, $[50, 70]$, and $[70, 90]$. (These four ranges together account for 66% of all observations.) Within each plot, uniform confidence bands are constructed separately for each number of bidders ($n = 2, \ldots, 11$), using a Bonferroni-adjusted confidence level to ensure joint coverage across n (as well as uniformly across τ). For example, the $n = 5$ band in the bottom-left plot uses data only for auctions with appraisal value in the range $[50, 70]$ and five bidders. Because we restrict to a constant number of bidders for each band, we use the bands described in Theorem 3 that compute nearly instantaneously.

Besides the far left tails and a few encroaching bands, all four plots in Figure 2 show substantial overlap across all n. That is, there are possible quantile functions that fit within all ten bands in each plot, consistent with homogeneity across n. Although there is probably some heterogeneity, this suggests that the statistical cost of ignoring heterogeneity may be outweighed by the benefit of increased precision by pooling the data across $n = 2, \ldots, 11$, as in the next plot.

Figure 3 takes a brief detour from our main analysis to illustrate our results for unobserved heterogeneity. We show some examples of the heterogeneity-robust uniform confidence bands described in Corollary 11 based on finite-sample one-sided uniform confidence bands for $Q_W(\cdot)$, computed using the R code from Goldman and Kaplan (2018a). As the majorant/minorant graph of Figure 1 suggests, the upper bound for the bid quantile function is significantly affected, whereas the lower bound requires much less change in order to provide robustness to unobserved heterogeneity.

Figure 4 (left) extends Figure 2 by assessing heterogeneity in net bid distributions across

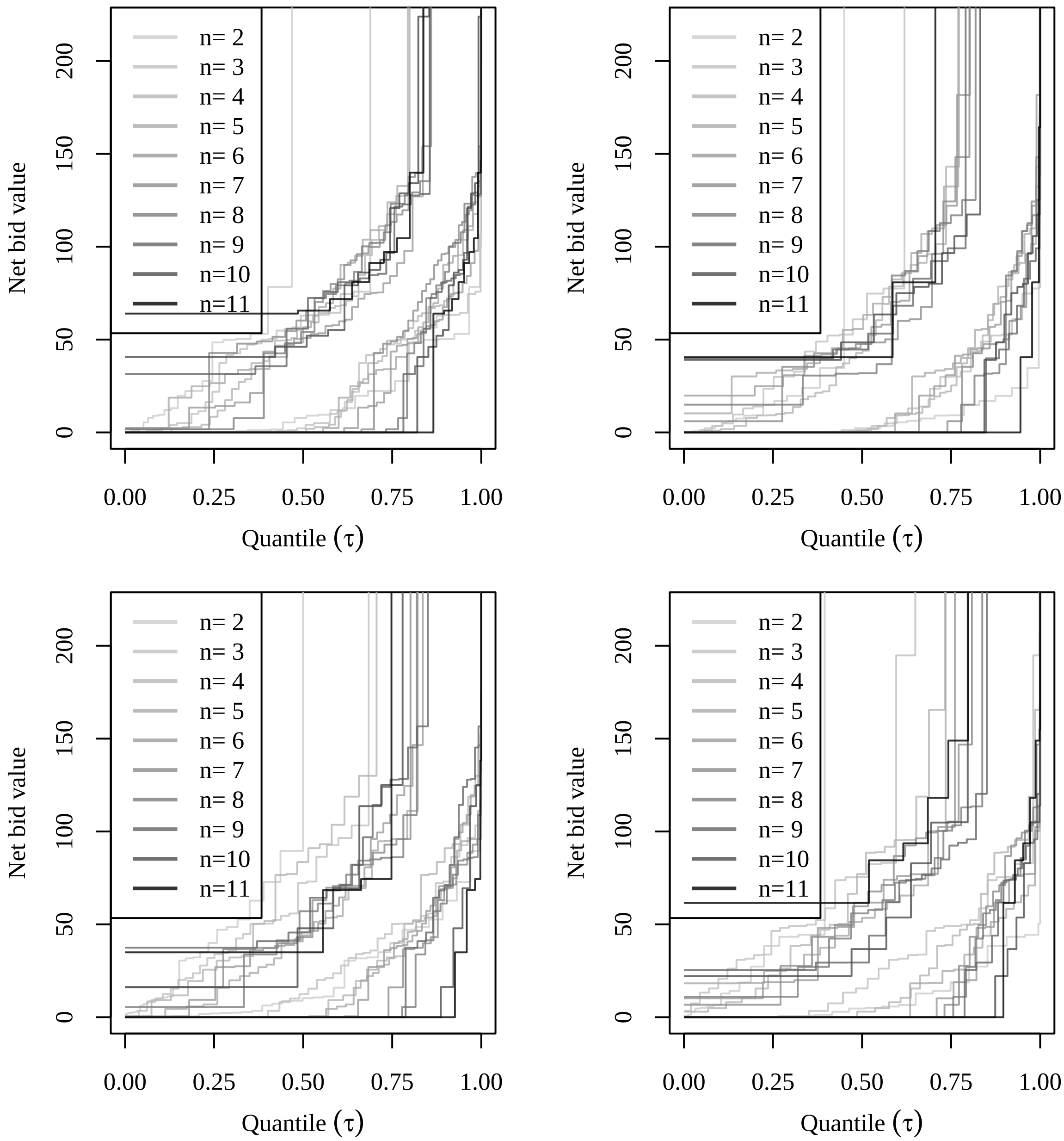

Figure 2: Two-sided uniform confidence bands, Bonferroni-corrected to have joint (over $n = 2, \ldots, 11$) 90% coverage within each plot, for appraisal values in the range $[10, 30]$ in the top left, $[30, 50]$ in top right, $[50, 70]$ in bottom left, and $[70, 90]$ in bottom right.

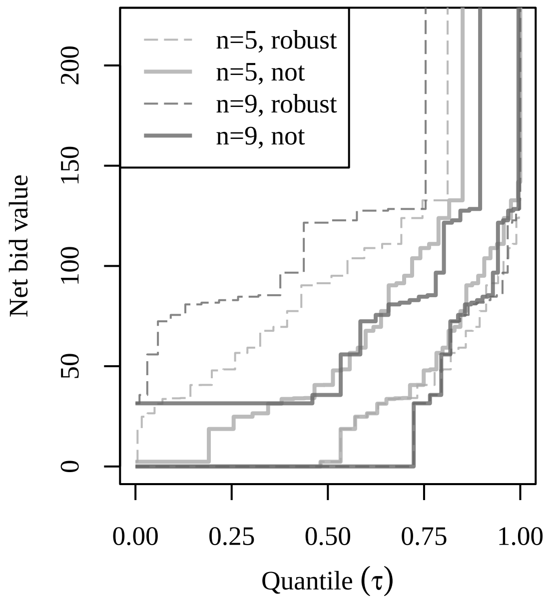

Figure 3: One-sided 90% uniform confidence bands by n, appraisal in $[10, 30]$.

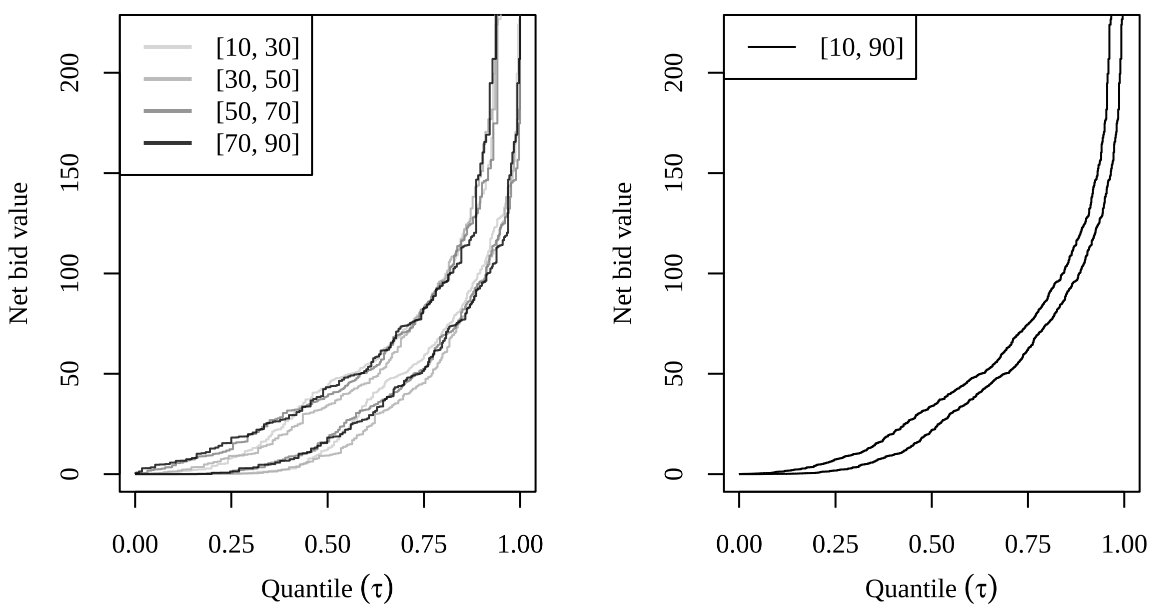

Figure 4: Two-sided uniform confidence bands. Left: separately by appraisal value range, Bonferroni-corrected to have joint 90% coverage. Right: pooling all appraisal values in $[10, 90]$, 90% confidence level.

auctions with different appraisal values, pooling across $n = 2, \ldots, 11$. Uniform confidence bands are constructed separately for each appraisal value range, using a Bonferroni-adjusted confidence level to ensure joint coverage across the four appraisal ranges (as well as uniformly across τ). Because each band now includes auctions with varying numbers of bidders, we use the method described in Theorem 7. Compared with each individual band in Figure 2, each band in Figure 4 (left) is much more precise, showing the advantage of allowing a varying number of bidders. Further, in Figure 2, bands based on a small number of bidders lack precision in the upper quartile (or more) of the distribution, while bands based on a larger n tend to lack precision in the lower part of the distribution. In contrast, the bands in Figure 4 (left) are informative across almost the whole distribution, because they use the smaller n observations to learn more about the lower tail and use the larger n observations to learn more about the upper tail. Besides improved precision, the pooled-n bands again show substantial overlap, suggesting that heterogeneity across appraisal value is small and may be outweighed by the precision benefit of pooling the data further.

Figure 4 (right) shows the two-sided 90% uniform confidence band when pooling across all n and all four appraisal value ranges. It shows the improved precision from using more data, with a relatively tight band across almost all quantiles.

Finally, we use our pooled uniform confidence band to compute some finite-sample confidence intervals for the counterfactual optimal reserve price. These are based on the bounds derived by Haile and Tamer (2003), described in our Appendix A.1. We use two possibilities for the seller's value v_0.

First, we compute a confidence interval assuming the seller's value equals the appraisal price. Here, we reparameterize the reserve price as a premium above v_0. Given the ascending auction format and our subtracting v_0 from the bids, we interpret our band as being for the quantile value function minus v_0; switching the axes yields a uniform confidence band for the shifted value CDF, $F_V(\cdot - v_0)$. Thus, our lower and upper bound functions for $F_V(\cdot)$ are $\hat{F}_L(\cdot - v_0)$ and $\hat{F}_U(\cdot - v_0)$, because $\mathrm{P}(V - v_0 \leq v - v_0) = \mathrm{P}(V \leq v) = F_V(v)$. Let

p denote a generic reserve price, and let $x \equiv p - v_0$. Writing (22)–(24) in terms of x, $\hat{\pi}_1(x) = x[1 - \hat{F}_U(x)]$, $\hat{x}_1^* = \arg\max_x \hat{\pi}_1(x)$, $\hat{\pi}_1^* = \hat{\pi}_1(\hat{x}_1^*)$, $\hat{\pi}_2(x) = x[1 - \hat{F}_L(x)]$, and

$$\hat{x}_L = \sup\{x < \hat{x}_1^* : \hat{\pi}_2(x) \leq \hat{\pi}_1^*\}, \quad \hat{x}_U = \inf\{x > \hat{x}_1^* : \hat{\pi}_2(x) \leq \hat{\pi}_1^*\}. \tag{18}$$

That is, given $\mathrm{P}(\hat{F}_L(\cdot) \leq F_V(\cdot - v_0) \leq \hat{F}_U(\cdot)) \geq 1 - \alpha$ in finite-samples, we get a finite-sample $1 - \alpha$ confidence interval for the optimal reserve price $p^* = x^* + v_0$:

$$\mathrm{P}(\hat{x}_L \leq x^* \leq \hat{x}_U) = \mathrm{P}(\hat{x}_L + v_0 \leq p^* \leq \hat{x}_U + v_0) \geq 1 - \alpha. \tag{19}$$

Empirically, our two-sided 90% confidence interval is

$$[\hat{x}_L, \hat{x}_U] = [29.42, 100.20]. \tag{20}$$

For example, if $v_0 = 10$ is both the appraisal value and seller's value, then the interval for p^* is $[39.42, 110.20]$.

Second, as commonly done, we take the seller's value as $v_0 = 0$. Now we follow (22)–(24) directly, noting as above that our uniform confidence band for value CDF $F_V(\cdot)$ consists of the functions $\hat{F}_L(\cdot - v_0)$ and $\hat{F}_U(\cdot - v_0)$, where $\hat{F}_L(\cdot)$ and $\hat{F}_U(\cdot)$ form the band for the net value CDF of $V - v_0$. Conditional on an appraisal value of 10, we compute a two-sided finite-sample 90% confidence interval for p^* as

$$[\hat{p}_L, \hat{p}_U] = [32.45, 102.53]. \tag{21}$$

This confidence interval assuming seller value $v_0 = 0$ is shifted down compared to the confidence interval for p^* with appraisal 10 but seller value 10.

7 Simulations

Our uniform confidence bands have exact finite-sample coverage probability, up to arbitrarily small simulation error, so the following simulations serve primarily as a "sanity check."

Secondarily, they show the great computational advantage of using the pre-simulated $\tilde{\alpha}$ from Goldman and Kaplan (2018a) when the number of bidders is constant across auctions. Code in R (R Core Team, 2023) to replicate our results is provided.

The first data-generating process (DGP) has the same number of bidders n in each of J auctions. The bids are $B_i \overset{iid}{\sim} \text{Unif}(0,1)$. The observed transaction price is the second-highest bid, $B_{n:n-1}$. The uniform distribution is without loss of generality because for any $Q_B(\cdot)$ satisfying A1, $V_{J:r} \le \tau \iff Q_B(V_{J:r}) \le Q_B(\tau)$. That is, to simulate bids from a general distribution, we would draw $U_i \overset{iid}{\sim} \text{Unif}(0,1)$ and take $B_i = Q_B(U_i)$ with transaction prices $W \equiv B_{n:n-1} = Q_B(U_{n:n-1}) = Q_B(V)$ with $V \equiv U_{n:n-1}$, so the order statistics (over auctions) satisfy $W_{J:r} = Q_B(V_{J:r})$. Thus, checking whether $W_{J:r}$ covers $Q_B(\tau)$ is numerically identical to checking whether $V_{J:r}$ covers τ.

Table 1: Coverage probability of 90% uniform confidence bands.

		Pre-simulated			Fully simulated			Average CDF		
n	J	2-sided	Lower	Upper	2-sided	Lower	Upper	2-sided	Lower	Upper
4	20	0.902	0.907	0.903	0.899	0.900	0.896	0.898	0.899	0.897
4	50	0.902	0.909	0.910	0.899	0.899	0.899	0.899	0.900	0.900
4	100	0.908	0.910	0.913	0.904	0.902	0.906	0.904	0.903	0.906
7	20	0.906	0.910	0.910	0.903	0.903	0.903	0.902	0.903	0.902
7	50	0.903	0.908	0.912	0.900	0.899	0.904	0.900	0.900	0.903
7	100	0.904	0.909	0.909	0.900	0.900	0.900	0.900	0.900	0.902
10	20	0.904	0.908	0.909	0.902	0.901	0.900	0.901	0.901	0.902
10	50	0.907	0.909	0.910	0.903	0.901	0.903	0.903	0.901	0.902
10	100	0.907	0.911	0.909	0.903	0.904	0.901	0.904	0.904	0.904

10,000 replications, 100,000 draws of $\{\beta_j\}_{j=1}^{J}$.

Table 1 shows the uniform coverage probability of our three uniform confidence bands: those from Theorems 2 and 3 in Section 3 using pre-simulated $\tilde{\alpha}$ values provided by Goldman and Kaplan (2018a); those from Theorems 6 and 7 in Section 4.2, where the simulated beta random variables are used both for calibrating the uniform coverage probability and for computing the marginal quantiles; and those from Section 4.4 using the average CDF to approximate marginal quantiles but using simulation to calibrate the uniform coverage probability. For the latter two, looking across two-sided, lower one-sided, and upper one-sided bands alike, the simulated coverage probability is always in the interval $[0.896, 0.906]$, extremely close to the nominal 0.900 confidence level. This is expected because these methods have exact finite-sample coverage probability with the only error due to simulation, and this error is small because we use 10^5 draws to calibrate $\tilde{\alpha}$ and we run 10^4 simulation replications.

The uniform coverage probability of the first method (with pre-simulated $\tilde{\alpha}$) is slightly less precise but many orders of magnitude faster. That is, when using the pre-simulated values provided in the code from Goldman and Kaplan (2018a), the coverage probabilities are all in the interval $[0.902, 0.913]$, which is still very close to the 0.900 confidence level, and computation time is around 10,000 times faster by avoiding just-in-time simulation. For these particular DGPs, computation times using the values from Goldman and Kaplan (2018a) are on the order of 0.001 seconds, compared to on the order of 10 seconds to simulate the $\tilde{\alpha}$. Although waiting 10 to 20 seconds is not too bad, with larger J and/or multiple datasets the computational advantage becomes more important.

The second DGP has different numbers of bidders across auctions, but again the results act primarily as a sanity check because our methods have exact finite-sample coverage. Here, there are J_1 auctions with n_1 bidders, and J_2 auctions with n_2 bidders. Again the individual bids are $B_i \overset{iid}{\sim} \text{Unif}(0, 1)$ without loss of generality, and the observed transaction price in each auction is the second-highest.

Table 2: Coverage probability of 90% uniform confidence bands.

				Simulated marginals			Average CDF		
n_1	n_2	J_1	J_2	Two-sided	Lower	Upper	Two-sided	Lower	Upper
4	2	20	20	0.896	0.900	0.900	0.897	0.898	0.898
4	10	20	20	0.900	0.898	0.901	0.899	0.896	0.898
4	20	20	20	0.909	0.899	0.906	0.905	0.900	0.906
4	2	50	50	0.903	0.901	0.902	0.900	0.900	0.903
4	10	50	50	0.905	0.905	0.901	0.907	0.907	0.901
4	20	50	50	0.897	0.901	0.899	0.897	0.901	0.898

10,000 replications, 100,000 draws of $\{\beta_j\}_{j=1}^{J}$.

Table 2 shows the uniform coverage probability of our uniform confidence bands from Theorems 6 and 7 in Section 4.2 ("Simulated marginals") and from Section 4.4 ("Average CDF"). Nearly the same as in Table 1, looking across two-sided, lower one-sided, and upper one-sided bands alike, the simulated uniform coverage probability is always in the interval $[0.896, 0.909]$, extremely close to the nominal 0.900 confidence level, as expected. The average CDF alternative is faster, taking 45–70% the computation time of the method with simulated marginal quantiles (12–20 seconds vs. 20–41 seconds), while maintaining the same uniform coverage probability, but not exactly even pointwise coverage probability. Which method is preferred may depend on the empirical setting and the user's computing power (and patience).

8 Conclusion

We have provided new uniform confidence bands with exact finite-sample coverage probability for the underlying bid quantile function when only the transaction price (the highest or second-highest order statistic) is observed from each auction, assuming symmetric independent private values but allowing different numbers of bidders per auction. With unobserved heterogeneity, the lower bound confidence function remains valid for first-price auctions, and we provide other one-sided uniform confidence bands for first-price and second-price auctions, based on new bounds we derive for second-price auctions. We hope these new methods prove useful in practice, following our empirical illustration as a guide.

Our methodology may also be useful when multiple order statistics are observed. For example, related to Theorem 1(ii)(a) of Athey and Haile (2002), we can construct a uniform confidence band based only on the second-highest bid in each auction, and we can separately construct a band based only on the highest bid. Assuming symmetric independent private values, these bands both cover the same (true) bid distribution with the desired finite-sample probability, so there is a small probability that they do not overlap everywhere in $\tau \in [0, 1]$.

Thus, if the bands do not overlap, then we have evidence of a violation of the assumed symmetry or independence of private values. More generally, we can construct a separate band corresponding to each bid order statistic, and we can use a Bonferroni correction to make the multiple uniform confidence bands hold jointly with the desired finite-sample probability.

A Examples of counterfactuals and confidence intervals

In the following subsections, we describe some common counterfactuals of interest and show how to use our quantile function uniform confidence bands to derive finite-sample confidence intervals for the counterfactuals.

A.1 Optimal reserve price

We show how to use the results from Haile and Tamer (2003) to derive finite-sample confidence intervals for the optimal reserve price, as computed in our empirical application. Specifically, their Section IV shows how to derive bounds on the optimal reserve price given bounds on the value CDF. By their Theorem 4, these bounds are sharp. Although their paper focuses on (partial) identification, their footnote 12 anticipates our use of their results to derive a confidence interval for the optimal reserve price given a uniform confidence band for the value CDF ("Our solution could be applied in that case as well").

Notationally, let $F_V(\cdot)$ be the value CDF, v_0 the seller's value of the object being auctioned, and p a possible reserve price. More generally, $F_V(\cdot)$ can be the CDF of the truncated value distribution if the data come from auctions with a sometimes-binding reserve price, as defined in their (1) based on the untruncated value CDF $F_{V0}(\cdot)$: $F_V(v) = [F_{V0}(v) - F_{V0}(p_0)]/[1 - F_{V0}(p_0)]$ given historical reserve price p_0. As long as the optimal reserve price is higher than the historical reserve price, the solution can still be found. Haile and Tamer (2003) note that "no reserve prices or extremely low reserves" are common (page

15).

Haile and Tamer (2003) assume symmetric independent private values, as well as some shape restrictions on the value distribution. Specifically, they assume that $F_V(\cdot)$ is strictly increasing and continuously differentiable (page 13) and that $(p - v_0)[1 - F_{V0}(p)]$ is strictly pseudoconcave in p over the support of the value distribution (Assumption 3).

The confidence interval $[\hat{p}_L, \hat{p}_U]$ for the optimal reserve price is characterized as follows. We skip the proofs and ignore the "trivial special cases" (page 16). Let $[\hat{F}_L(\cdot), \hat{F}_U(\cdot)]$ be the bounds for $F_V(\cdot)$, specifically our uniform confidence band, assuming we are in a setting like an ascending auction where the transaction price can be interpreted as (approximately) the second-highest value. (The CDF band is simply our quantile function band with the axes reversed.) Following page 16 of Haile and Tamer (2003), but adding hats, define

$$\hat{\pi}_1(p) \equiv (p - v_0)[1 - \hat{F}_U(p)], \quad \hat{p}_1^* \equiv \arg\max_p \hat{\pi}_1(p), \quad \hat{\pi}_1^* \equiv \hat{\pi}_1(\hat{p}_1^*), \tag{22}$$

$$\hat{\pi}_2(p) \equiv (p - v_0)[1 - \hat{F}_L(p)], \tag{23}$$

$$\hat{p}_L \equiv \sup\{p < \hat{p}_1^* : \hat{\pi}_2(p) \leq \hat{\pi}_1^*\}, \quad \hat{p}_U \equiv \inf\{p > \hat{p}_1^* : \hat{\pi}_2(p) \leq \hat{\pi}_1^*\}. \tag{24}$$

As seen in our code, these are easy to implement and compute immediately.

The confidence interval's finite-sample validity follows from the results of Haile and Tamer (2003). They show that the bounds $\hat{F}_L(\cdot) \leq F_V(\cdot) \leq \hat{F}_U(\cdot)$ imply the optimal reserve price p^* satisfies $p^* \in [\hat{p}_L, \hat{p}_U]$, so the confidence interval's finite-sample coverage probability is

$$\mathrm{P}(p^* \in [\hat{p}_L, \hat{p}_U]) \geq \mathrm{P}(\hat{F}_L(\cdot) \leq F_V(\cdot) \leq \hat{F}_U(\cdot)) \geq 1 - \alpha, \tag{25}$$

assuming $[\hat{F}_L(\cdot), \hat{F}_U(\cdot)]$ is one of our uniform confidence bands with confidence level $1 - \alpha$.

A.2 Expectation of highest value

In an auction with n bidders with values V_i $(i = 1, \ldots, n)$ with quantile function $Q_V(\cdot)$, the distribution of the highest value $V_{n:n}$ or its mean are commonly of economic interest. If the seller's value of the auctioned object is zero, then $V_{n:n}$ is the total surplus: the winning bidder receives surplus $V_{n:n}$ minus the transaction price, and the seller receives the transaction price, so the sum is $V_{n:n}$. This is true for any mechanism that allocates the object to the bidder with the highest value. This measure of expected total surplus $\mathrm{E}(V_{n:n})$ is important for considering the welfare loss of potentially inefficient alternative allocation mechanisms, or for compare the welfare benefit with the costs of negative externalities (e.g., Armstrong, 2013, §2.2). In some cases below, we also construct confidence intervals for the expected total surplus under counterfactual reserve prices.

A.2.1 Ascending auctions

For ascending auctions, consider the case when our uniform confidence band can be interpreted as a band for the value quantile function, $Q_V(\cdot)$. Recall that we interpret the transaction price as (approximately) the second-highest value $V_{n:n-1}$, so we cannot directly learn about $V_{n:n}$, which is never observed. However, its distribution can be written in terms of $Q_V(\cdot)$. Specifically, by the probability integral transform, $V_{n:n} \stackrel{d}{=} Q_V(U_{n:n})$, where $U_{n:n}$ is the maximum of n random variables $U_i \stackrel{iid}{\sim} \mathrm{Unif}(0,1)$, $i = 1, \ldots, n$. (That is, we can write $V_i = Q_V(U_i)$ for $U_i \sim \mathrm{Unif}(0,1)$, and the increasingness of $Q_V(\cdot)$ implies the highest U_i generates the highest $V_i = Q_V(U_i)$.)

First, imagine there is no reserve price. Recall the distribution $U_{n:n} \sim \mathrm{Beta}(n,1)$ (e.g., Wilks, 1962, p. 236), which has PDF $f_{U_{n:n}}(u) = nu^{n-1}$. Then,

$$\mathrm{E}(V_{n:n}) = \mathrm{E}[Q_V(U_{n:n})] = \int_0^1 Q_V(u) f_{U_{n:n}}(u)\,\mathrm{d}u = n\int_0^1 Q_V(u) u^{n-1}\,\mathrm{d}u.$$

This is an increasing functional of $Q_V(\cdot)$, so $\hat{Q}_L(\cdot) \le Q_V(\cdot) \le \hat{Q}_U(\cdot)$ implies

$$n\int_0^1 \hat{Q}_L(u) u^{n-1}\,\mathrm{d}u \le \mathrm{E}(V_{n:n}) \le n\int_0^1 \hat{Q}_U(u) u^{n-1}\,\mathrm{d}u.$$

Consequently, the left-hand and right-hand sides above form a finite-sample confidence interval:

$$\mathrm{P}\biggl(n\int_0^1 \hat{Q}_L(u)u^{n-1}\,\mathrm{d}u \le \mathrm{E}(V_{n:n}) \le n\int_0^1 \hat{Q}_U(u)u^{n-1}\,\mathrm{d}u\biggr)$$
$$\ge \mathrm{P}\bigl(\hat{Q}_L(\cdot) \le Q_V(\cdot) \le \hat{Q}_U(\cdot)\bigr) \ge 1-\alpha$$

given $[\hat{Q}_L(\cdot), \hat{Q}_U(\cdot)]$ is a $1-\alpha$ uniform confidence band for $Q_V(\cdot)$. Similarly, one-sided confidence intervals for $\mathrm{E}(V_{n:n})$ can be constructed from one-sided uniform confidence bands for $Q_V(\cdot)$.

Second, imagine the reserve price is p. The object is not sold (zero surplus) if $V_{n:n} < p$, so the integral does not start at $u = 0$, but rather at $u = F_V(p)$:

$$\mathrm{E}(V_{n:n} \mid V_{n:n} \ge p) = \mathrm{E}[Q_V(U_{n:n}) \mid U_{n:n} \ge F_V(p)] = \int_{F_V(p)}^1 Q_V(u) f_{U_{n:n}}(u)\,\mathrm{d}u$$
$$= n\int_{F_V(p)}^1 Q_V(u)u^{n-1}\,\mathrm{d}u.$$

This also matches the expression in Table 1 of Andreyanov and Franguridi (2024). This is still increasing in $Q_V(\cdot)$: higher $Q_V(\cdot)$ corresponds to lower $F_V(p)$, which also increases the result. Thus, we can construct a finite-sample confidence interval similar to above:

$$\mathrm{P}\biggl(n\int_{\hat{F}_U(p)}^1 \hat{Q}_L(u)u^{n-1}\,\mathrm{d}u \le \mathrm{E}(V_{n:n}) \le n\int_{\hat{F}_L(p)}^1 \hat{Q}_U(u)u^{n-1}\,\mathrm{d}u\biggr)$$
$$\ge \mathrm{P}\bigl(\hat{Q}_L(\cdot) \le Q_V(\cdot) \le \hat{Q}_U(\cdot)\bigr) \ge 1-\alpha,$$

again given $[\hat{Q}_L(\cdot), \hat{Q}_U(\cdot)]$ is a $1-\alpha$ uniform confidence band for $Q_V(\cdot)$. Here $\hat{F}_L(\cdot)$ is just $\hat{Q}_U(\cdot)$ with the axes reversed, and similarly for $\hat{F}_U(\cdot)$ and $\hat{Q}_L(\cdot)$, so $\hat{Q}_L(\cdot) \le Q_V(\cdot) \le \hat{Q}_U(\cdot)$ implies $\hat{F}_L(p) \le F_V(p) \le \hat{F}_U(p)$.

Note that the upper bound requires an assumed upper bound of the support of the value distribution so that $\lim_{\tau\to 1} \hat{Q}_U(\tau) < \infty$. Even if the value distribution's support is unbounded above, we can still construct one-sided confidence intervals with a lower bound because (presumably) $V \ge 0$, so $\hat{Q}_L(\tau) = 0$ in the extreme lower tail. Additionally, we can still compute confidence intervals for alternatives like a trimmed mean of $V_{n:n}$ that removes the dependence on the extreme upper tail.

A.2.2 First-price auctions

For first-price sealed-bid auctions with symmetric independent private values and unobserved heterogeneity (of the type we consider), when only the winning bid $B_{n:n}$ is observed, Armstrong (2013) provides bounds on $\mathrm{E}(V_{n:n})$ in terms of $\mathrm{E}(B_{n:n})$, the expectation of the highest bid. Specifically, simplifying the expression in Theorem 2, the bounds on page 386 are

$$\mathrm{E}(B_{n:n}) \le \mathrm{E}(V_{n:n}) \le \frac{n}{n-1}\bar{b} - \frac{1}{n-1}\,\mathrm{E}(B_{n:n}),$$

where $\bar{b}$ is the upper bound of the bid distribution's support. A finite-sample one-sided confidence interval for $\mathrm{E}(V_{n:n})$ can be derived from a one-sided uniform confidence band satisfying $\mathrm{P}(\hat{Q}_L(\cdot) \le Q_B(\cdot)) \ge 1-\alpha$:

$$\begin{aligned}\mathrm{P}\Big(n\int_0^1 \hat{Q}_L(u)u^{n-1}\,\mathrm{d}u \le \mathrm{E}(V_{n:n})\Big)\\ \ge \mathrm{P}\big(\hat{Q}_L(\cdot) \le Q_B(\cdot)\big) \ge 1-\alpha\end{aligned}$$

The upper bound of a confidence interval depends heavily on $\bar{b}$, but could be derived similarly, plugging $\hat{Q}_L(\cdot)$ into the integral defining $\mathrm{E}(B_{n:n})$.

A.3 Mean of value distribution

In the same setting as Appendix A.2, the mean of the value distribution $\mathrm{E}(V)$ is also of economic and counterfactual interest. The following motivation is from Armstrong (2013, §2.1). First, the mean $\mathrm{E}(V)$ provides the expected surplus if instead of being auctioned, the object is assigned by a random lottery, a common mechanism used by the government (Armstrong, 2013, §2.4). Second, $\mathrm{E}(V)$ can be used to compute the value of a nonrival technology that could replace the auctioned object and be used by all participants. For example, if n participants have been bidding on emissions permits, $n\,\mathrm{E}(V)$ is the expected aggregate value of a new technology that reduces pollution (without affecting production efficiency) by the amount of the permit.

A.3.1 Ascending auctions

For ascending auctions where our uniform confidence band can be interpreted as a band for the value quantile function $Q_V(\cdot)$, a confidence interval for $\mathrm{E}(V)$ follows from arguments similar to those in Appendix A.2.1 for the $\mathrm{E}(V_{n:n})$ confidence interval. That is, we can write $\mathrm{E}(V)$ as an integral of the quantile function and plug in bounds, so

$$\begin{aligned}
&\mathrm{P}\Bigl(\int_0^1 \hat{Q}_L(u)\,\mathrm{d}u \le \mathrm{E}(V) = \int_0^1 Q_V(u)\,\mathrm{d}u \le \int_0^1 \hat{Q}_U(u)\,\mathrm{d}u\Bigr) \\
&\quad \ge \mathrm{P}\bigl(\hat{Q}_L(\cdot) \le Q_V(\cdot) \le \hat{Q}_U(\cdot)\bigr) \ge 1-\alpha,
\end{aligned}$$

and the corresponding one-sided confidence intervals (based on one-sided uniform confidence bands) also have finite-sample coverage probability. As before, the same caveat about $\hat{Q}_U(\cdot)$ in the upper tail applies here, where an assumed upper bound of the value distribution's support is needed to get a finite upper bound.

A.3.2 First-price auctions

For first-price sealed-bid auctions with symmetric independent private values and unobserved heterogeneity (of the type we consider), when only the winning bid $B_{n:n}$ is observed, Armstrong (2013) provides bounds on $\mathrm{E}(V)$ in terms of $\mathrm{E}(B)$, the mean of the bid distribution. Specifically, simplifying the expression in Theorem 1, the bounds on page 384 are

$$\mathrm{E}(B) \le \mathrm{E}(V) \le \frac{n-2}{n-1}\,\mathrm{E}(B) + \frac{1}{n-1}\bar{b},$$

where $\bar{b}$ is the upper bound of the bid distribution's support. A finite-sample two-sided confidence interval for $\mathrm{E}(V)$ can be derived from our two-sided uniform confidence band for $Q_B(\cdot)$:

$$\mathrm{P}\Bigl(\int_0^1 \hat{Q}_L(u)\,\mathrm{d}u \le \mathrm{E}(V) \le \frac{1}{n-1}\bar{b} + \frac{n-2}{n-1}\int_0^1 \hat{Q}_U(u)\,\mathrm{d}u\Bigr)$$

$$\geq \mathrm{P}\big(\hat{Q}_L(\cdot) \leq Q_B(\cdot) \leq \hat{Q}_U(\cdot)\big) \geq 1-\alpha,$$

and the corresponding one-sided confidence intervals (based on one-sided uniform confidence bands) also have finite-sample coverage probability. Unlike for $\mathrm{E}(V_{n:n})$, here the influence of $\bar{b}$ is much smaller, especially with a larger number of bidders n. As before, $\bar{b}$ is not required for the lower bound one-sided confidence interval.

A.4 Expected revenue

We derive confidence intervals for the expected revenue, that is, the expected transaction price. We focus on ascending auctions where our uniform confidence bands can be interpreted in terms of the value quantile function $Q_V(\cdot)$, and we are interested in expected revenue for an auction with counterfactual number of bidders n that differs from some or all of the auctions in the data. In this context, the transaction price is (approximately) $V_{n:n-1}$, so expected revenue is $\mathrm{E}(V_{n:n-1})$. This is closely related to our results in Appendix A.2 for $\mathrm{E}(V_{n:n})$, so we keep the explanations here brief.

Like for $\mathrm{E}(V_{n:n})$, write $V_i = Q_V(U_i)$ for $U_i \overset{iid}{\sim} \mathrm{Unif}(0,1)$, so $V_{n:n-1} = Q_V(U_{n:n-1})$, where standard uniform order statistic $U_{n:n-1} \sim \mathrm{Beta}(n-1,2)$ has PDF $f_{\beta(n-1,2)}(u) = n(n-1)u^{n-2}(1-u)$. Thus,

$$\mathrm{E}(V_{n:n-1}) = \mathrm{E}[Q_V(U_{n:n-1})] = \int_0^1 Q_V(u) f_{\beta(n-1,2)}\,\mathrm{d}u.$$

Because the PDF is non-negative, the lower bound is attained by plugging in a lower bound for $Q_V(\cdot)$, and the upper bound by the upper bound for $Q_V(\cdot)$. Thus, given a finite-sample uniform confidence band with $\mathrm{P}(\hat{Q}_L(\cdot) \leq Q_V(\cdot) \leq \hat{Q}_U(\cdot)) \geq 1-\alpha$, a two-sided finite-sample confidence interval can be derived:

$$\begin{aligned}
&\mathrm{P}\bigg(\int_0^1 \hat{Q}_L(u) f_{\beta(n-1,2)}(u)\,\mathrm{d}u \leq \mathrm{E}(V_{n:n-1}) \leq \int_0^1 \hat{Q}_U(u) f_{\beta(n-1,2)}(u)\,\mathrm{d}u\bigg) \\
&\quad \geq \mathrm{P}\big(\hat{Q}_L(\cdot) \leq Q_V(\cdot) \leq \hat{Q}_U(\cdot)\big) \geq 1-\alpha.
\end{aligned}$$

Similarly, one-sided confidence intervals for $\mathrm{E}(V_{n:n-1})$ can be constructed from one-sided uniform confidence bands for $Q_V(\cdot)$. The same caveat as before applies about the extreme upper tail of $\hat{Q}_U(\cdot)$ being bounded only by an assumed upper bound of the support of the value distribution, though the lower bound and corresponding one-sided confidence interval does not require such an assumption.

B Proofs

Proof of Proposition 1. Under A1 and A2,

$$
\begin{aligned}
&\mathrm{P}(W \le Q_B(F_\beta^{-1}(\tau))) \\
&= \mathrm{P}(F_B(W) \le F_\beta^{-1}(\tau)) && \text{by applying } F_B(\cdot), \text{ and } F_B(Q_B(x)) = x \text{ by A1} \\
&= \mathrm{P}(F_B(B_{n:k}) \le F_\beta^{-1}(\tau)) && \text{by definition of } W \equiv B_{n:k} \\
&= \tau.
\end{aligned}
$$

The last equality is because $F_B(B_{n:k}) \sim \mathrm{Beta}(k, n+1-k)$ by A1 and A2, given results from Wilks (1962, pp. 236–238). Thus, overall, $\mathrm{P}(W \le Q_B(F_\beta^{-1}(\tau))) = \tau$, which means $Q_B(F_\beta^{-1}(\tau))$ is the τ-quantile of W, which is equivalent to $Q_B(\tau) = Q_W(F_\beta(\tau))$ as stated in Proposition 1.

For the CDF identity, let $q = Q_B(\tau)$, so $F_B(q) = \tau$. Given the quantile identity, $q = Q_B(\tau) = Q_W(F_\beta(\tau))$, and applying $F_W(\cdot)$ yields $F_W(q) = F_\beta(\tau)$, given that $F_W(Q_W(x)) = x$ by A1. Finally, $F_\beta^{-1}(F_W(q)) = \tau = F_B(q)$. □

Proof of Theorem 2. With explanations for the equalities given below, under A1–A3, the coverage probability of the one-sided lower-bound confidence function $\hat{g}_L(\cdot; \tilde{\alpha})$ is

$$
\begin{aligned}
&\mathrm{P}\big(\hat{g}_L(\tau; \tilde{\alpha}) \le Q_B(\tau) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(W_{J:1} \le Q_B(\tau_1(1-\tilde{\alpha})) \text{ and } \ldots \text{ and } W_{J:J} \le Q_B(\tau_J(1-\tilde{\alpha}))\big) \\
&= \mathrm{P}\big(F_B(W_{J:1}) \le \tau_1(1-\tilde{\alpha}) \text{ and } \ldots \text{ and } F_B(W_{J:J}) \le \tau_J(1-\tilde{\alpha})\big)
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{P}(\beta_{J:1} \le \tau_1(1-\tilde{\alpha}) \text{ and } \ldots \text{ and } \beta_{J:J} \le \tau_J(1-\tilde{\alpha})) \\
&= \mathrm{P}\big(F_\beta(\beta_{J:1}) \le F_\beta(\tau_1(1-\tilde{\alpha})) \text{ and } \ldots \text{ and } F_\beta(\beta_{J:J}) \le F_\beta(\tau_J(1-\tilde{\alpha}))\big) \\
&= \mathrm{P}\big(F_\beta(\beta_{J:1}) \le \xi_{1-\tilde{\alpha},J,1} \text{ and } \ldots \text{ and } F_\beta(\beta_{J:J}) \le \xi_{1-\tilde{\alpha},J,J}\big) \\
&= 1-\alpha.
\end{aligned}
$$

The first equality is from the definition of $\hat{g}_L(\cdot;\tilde{\alpha})$ in (3) and the stair-step shape of the function combined with the non-decreasing property of $Q_B(\cdot)$. For example, temporarily writing $\tau_r(1-\tilde{\alpha})$ as τ_r for simplicity, because $\hat{g}_L(\tau;\tilde{\alpha}) = W_{J:1}$ for all $\tau \in [\tau_1, \tau_2)$ and $Q_B(\cdot)$ is non-decreasing (so $Q_B(\tau) \ge Q_B(\tau_1)$ over that same range), it follows that $\hat{g}_L(\tau;\tilde{\alpha}) \le Q_B(\tau)$ for all $\tau \in [\tau_1, \tau_2)$ if and only if $W_{J:1} \le Q_B(\tau_1)$. The second equality follows from A1. The third equality simply switches to the notation $\beta_{J:r} \equiv F_B(W_{J:r})$ from (1). The fourth equality applies $F_\beta(\cdot)$ to each term, where $F_\beta(\cdot)$ is the CDF of the marginal $\mathrm{Beta}(k, n+1-k)$ distribution of each β_j random variable. The fifth equality plugs in the definitions of the $\tau_r(1-\tilde{\alpha})$ from (4). The final equality comes from the fact that by (26) the $F_\beta(\beta_j) \overset{iid}{\sim} \mathrm{Unif}(0,1)$ over $j = 1, \ldots, J$, so

$$
(F_\beta(\beta_{J:1}), F_\beta(\beta_{J:2}) - F_\beta(\beta_{J:1}), \ldots, F_\beta(\beta_{J:J}) - F_\beta(\beta_{J:J-1}), 1 - F_\beta(\beta_{J:J}))
$$

jointly follows the Dirichlet distribution $\mathrm{Dir}(\overbrace{1, 1, \ldots, 1}^{J+1})$ as given by Wilks (1962, pp. 236–238) or Goldman and Kaplan (2018a, Thm. 4). Thus, by Method 1 of Goldman and Kaplan (2018a), $\tilde{\alpha}$ is defined to set the final expression equal to $1-\alpha$ exactly.

Similarly, under A1–A3, the coverage probability of $\hat{g}_U(\cdot;\tilde{\alpha})$ equals $1-\alpha$ exactly, with the same reason for each equality below including Method 1 of Goldman and Kaplan (2018a) for the final step:

$$
\begin{aligned}
&\mathrm{P}\big(\hat{g}_U(\tau;\tilde{\alpha}) \ge Q_B(\tau) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(W_{J:1} \ge Q_B(\tau_1(\tilde{\alpha})) \text{ and } \ldots \text{ and } W_{J:J} \ge Q_B(\tau_J(\tilde{\alpha}))\big) \\
&= \mathrm{P}\big(F_B(W_{J:1}) \ge \tau_1(\tilde{\alpha}) \text{ and } \ldots \text{ and } F_B(W_{J:J}) \ge \tau_J(\tilde{\alpha})\big)
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{P}(\beta_{J:1} \geq \tau_1(\tilde{\alpha}) \text{ and } \ldots \text{ and } \beta_{J:J} \geq \tau_J(\tilde{\alpha})) \\
&= \mathrm{P}\big(F_\beta(\beta_{J:1}) \geq F_\beta(\tau_1(\tilde{\alpha})) \text{ and } \ldots \text{ and } F_\beta(\beta_{J:J}) \geq F_\beta(\tau_J(\tilde{\alpha}))\big) \\
&= \mathrm{P}\big(F_\beta(\beta_{J:1}) \geq \xi_{\tilde{\alpha},J,1} \text{ and } \ldots \text{ and } F_\beta(\beta_{J:J}) \geq \xi_{\tilde{\alpha},J,J}\big) \\
&= 1 - \alpha. \qquad \square
\end{aligned}
$$

Proof of Theorem 3. Under Assumptions A1–A3,

$$
\begin{aligned}
&\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}) \leq Q_B(\tau) \leq \hat{g}_U(\tau;\tilde{\alpha}) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(Q_B(\tau_1(\tilde{\alpha})) \leq W_{J:1} \leq Q_B(\tau_1(1-\tilde{\alpha})) \text{ and } \ldots \\
&\qquad \ldots \text{ and } Q_B(\tau_J(\tilde{\alpha})) \leq W_{J:J} \leq Q_B(\tau_J(1-\tilde{\alpha}))\big) \\
&= \mathrm{P}\big(\tau_1(\tilde{\alpha}) \leq F_B(W_{J:1}) \leq \tau_1(1-\tilde{\alpha}) \text{ and } \ldots \text{ and } \tau_J(\tilde{\alpha}) \leq F_B(W_{J:J}) \leq \tau_J(1-\tilde{\alpha})\big) \\
&= \mathrm{P}\big(\tau_1(\tilde{\alpha}) \leq \beta_{J:1} \leq \tau_1(1-\tilde{\alpha}) \text{ and } \ldots \text{ and } \tau_J(\tilde{\alpha}) \leq \beta_{J:J} \leq \tau_J(1-\tilde{\alpha})\big) \\
&= \mathrm{P}\left(\bigcap_{r=1}^{J} \{F_\beta(\tau_r(\tilde{\alpha})) \leq F_\beta(\beta_{J:r}) \leq F_\beta(\tau_r(1-\tilde{\alpha}))\}\right) \\
&= \mathrm{P}\left(\bigcap_{r=1}^{J} \{\xi_{\tilde{\alpha},J,r} \leq F_\beta(\beta_{J:r}) \leq \xi_{1-\tilde{\alpha},J,r}\}\right) \\
&= 1 - \alpha.
\end{aligned}
$$

The equalities have the same reasons given in Theorem 2, except the final equality follows from Method 2 (instead of Method 1) of Goldman and Kaplan (2018a). □

Proof of Theorem 4. See (8) for the quantile function band. For the CDF,

$$
\begin{aligned}
&\mathrm{P}\big(F_\beta^{-1}(\hat{L}(q)) \leq F_B(q) \leq F_\beta^{-1}(\hat{U}(q)) \text{ for all } q \in \mathbb{R}\big) \\
&= \mathrm{P}\big(F_\beta^{-1}(\hat{L}(q)) \leq \overbrace{F_\beta^{-1}(F_W(q))}^{\text{by Proposition 1}} \leq F_\beta^{-1}(\hat{U}(q)) \text{ for all } q \in \mathbb{R}\big) \\
&= \mathrm{P}\big(\hat{L}(q) \leq F_W(q) \leq \hat{U}(q) \text{ for all } q \in \mathbb{R}\big). \qquad \square
\end{aligned}
$$

Proof of Theorem 5. Similar to the proof of Theorem 2,

$$\begin{aligned}
&\mathrm{P}\big(\hat{w}_L(\tau;\tilde{\alpha}) \le Q_W(\tau) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(W_{J:1} \le Q_W(\xi_{1-\tilde{\alpha},J,1}) \text{ and } \ldots \text{ and } W_{J:J} \le Q_W(\xi_{1-\tilde{\alpha},J,J})\big) \\
&= \mathrm{P}\big(F_W(W_{J:1}) \le \xi_{1-\tilde{\alpha},J,1} \text{ and } \ldots \text{ and } F_W(W_{J:J}) \le \xi_{1-\tilde{\alpha},J,J}\big) \\
&= 1-\alpha,
\end{aligned}$$

where now the last equality follows from the fact that the continuity of $F_W(\cdot)$ implies that $(F_W(W_{J:1}), \ldots, F_W(W_{J:J}))$ follows a Dirichlet distribution with every parameter equal to one (Wilks, 1962, pp. 236–238) combined with Method 1 of Goldman and Kaplan (2018a), noting that our $\xi_{1-\tilde{\alpha},J,r}$ in their notation would be $B_{r,n}^{1-\tilde{\alpha}}$. That is, their Method 1 defines $\tilde{\alpha}$ to set the last equality to $1-\alpha$. (For the equality before that, the assumption that $F_W(\cdot)$ is continuous and strictly increasing implies $F_W(Q_W(x)) = x$.) For the same reasons,

$$\begin{aligned}
&\mathrm{P}\big(\hat{w}_U(\tau;\tilde{\alpha}) \ge Q_W(\tau) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(W_{J:1} \ge Q_W(\xi_{\tilde{\alpha},J,1}) \text{ and } \ldots \text{ and } W_{J:J} \ge Q_W(\xi_{\tilde{\alpha},J,J})\big) \\
&= \mathrm{P}\big(F_W(W_{J:1}) \ge \xi_{\tilde{\alpha},J,1} \text{ and } \ldots \text{ and } F_W(W_{J:J}) \ge \xi_{\tilde{\alpha},J,J}\big) \\
&= 1-\alpha.
\end{aligned}$$

For the two-sided uniform confidence band,

$$\begin{aligned}
&\mathrm{P}\big(\hat{w}_L(\tau;\tilde{\alpha}) \le Q_W(\tau) \le \hat{w}_U(\tau;\tilde{\alpha}) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(Q_W(\xi_{\tilde{\alpha},J,1}) \le W_{J:1} \le Q_W(\xi_{1-\tilde{\alpha},J,1}) \text{ and } \ldots \\
&\qquad \ldots \text{ and } Q_W(\xi_{\tilde{\alpha},J,J}) \le W_{J:J} \le Q_W(\xi_{1-\tilde{\alpha},J,J})\big) \\
&= \mathrm{P}\big(\xi_{\tilde{\alpha},J,1} \le F_W(W_{J:1}) \le \xi_{1-\tilde{\alpha},J,1} \text{ and } \ldots \text{ and } \xi_{\tilde{\alpha},J,J} \le F_W(W_{J:J}) \le \xi_{1-\tilde{\alpha},J,J}\big) \\
&= 1-\alpha,
\end{aligned}$$

where the final equality now comes from Method 2 (instead of Method 1) of Goldman and Kaplan (2018a). □

Proof of Theorem 6. For the uniform coverage probability of $\hat{g}_L(\cdot;\tilde{\alpha}_L)$, as derived in the proof of Theorem 2,

$$\begin{aligned}&\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}_L) \le Q_B(\tau) \text{ for all } \tau \in [0,1]\big)\\&= \mathrm{P}\big(\beta_{J:1} \le \tau_1(1-\tilde{\alpha}_L) \text{ and } \dots \text{ and } \beta_{J:J} \le \tau_J(1-\tilde{\alpha}_L)\big),\end{aligned}$$

which equals $1-\alpha$ by (13). Similarly, by the proof of Theorem 2 and then (14),

$$\begin{aligned}\mathrm{P}\big(Q_B(\tau) \le \hat{g}_U(\tau;\tilde{\alpha}_U) \text{ for all } \tau \in [0,1]\big) &= \mathrm{P}\big(\beta_{J:1} \ge \tau_1(\tilde{\alpha}_U) \text{ and } \dots \text{ and } \beta_{J:J} \ge \tau_J(\tilde{\alpha}_U)\big)\\&= 1-\alpha.\end{aligned}$$

□

Proof of Theorem 7. As derived in the proof of Theorem 3,

$$\begin{aligned}&\mathrm{P}\big(\hat{g}_L(\tau;\tilde{\alpha}_T) \le Q_B(\tau) \le \hat{g}_U(\tau;\tilde{\alpha}_T) \text{ for all } \tau \in [0,1]\big)\\&= \mathrm{P}\big(\tau_1(\tilde{\alpha}_T) \le \beta_{J:1} \le \tau_1(1-\tilde{\alpha}_T) \text{ and } \dots \text{ and } \tau_J(\tilde{\alpha}_T) \le \beta_{J:J} \le \tau_J(1-\tilde{\alpha}_T)\big),\end{aligned}$$

which then equals $1-\alpha$ given that $\tilde{\alpha}_T$ satisfies (15). □

Proof of Theorem 8. Median-unbiased means the probability that the estimator ($W_{J:r}$) is below the true value ($\tau_r(0.5)$-quantile of B) is 0.5. Using (11) and (12),

$$\mathrm{P}\big(W_{J:r} \le F_B^{-1}(\tau_r(0.5))\big) = \mathrm{P}\big(F_B(W_{J:r}) \le \tau_r(0.5)\big) = \mathrm{P}\big(\beta_{J:r} \le 0.5\text{-quantile of } \beta_{J:r}\big) = 0.5$$

because the distribution of $\beta_{J:r}$ is continuous. □

Proof of Theorem 9. For the CDF bounds, starting from the inequalities derived in (16) and (17), apply $g^{-1}(\cdot)$ and $h^{-1}(\cdot)$ to get the form stated. Note $F_\beta(\cdot)$ is a strictly increasing function for any (k,n), so the greatest convex minorant and least concave majorant are both strictly increasing and thus invertible.

For the quantile function bounds, let $q = Q_B(\tau)$. Then, $F_B(q) \le g^{-1}(F_W(q))$ becomes $F_B(Q_B(\tau)) \le g^{-1}(F_W(Q_B(\tau)))$, which becomes

$$Q_W(g(\tau)) \le Q_B(\tau)$$

after applying $Q_W(g(\cdot))$ to both sides and using $Q_B(\cdot) = F_B^{-1}(\cdot)$ and $Q_W(\cdot) = F_W^{-1}(\cdot)$ given Assumption A1. Similarly, $F_B(q) \geq h^{-1}(F_W(q))$ becomes $Q_W(h(\tau)) \geq Q_B(\tau)$. $\square$

Proof of Corollary 10. Part (i): if $k = n$, then $\text{Beta}(k, n+1-k)$ is $\text{Beta}(n, 1)$, whose CDF is $F_\beta(x) = x^n$; the results follow by plugging this into Theorem 9.

Part (ii): the key is that for $k < n$ (and $k > 1$), the corresponding $\text{Beta}(k, n+1-k)$ distribution is unimodal with mode $(k-1)/(n-1)$ strictly between zero and one, so its CDF $F_\beta(\cdot)$ is convex to the left of the mode and concave to the right of the mode, like the example in Figure 1. Thus, the greatest convex minorant equals $F_\beta(\cdot)$ up to the tangency point t_{1n}, and then connects to $(1,1)$ with a straight line; and the least concave majorant is the opposite, connecting $(0,0)$ with a straight line to tangency point t_{2n}, then following $F_\beta(\cdot)$ along the rest of its concave portion. $\square$

Proof of Corollary 11. For the lower confidence bound function,

$$\begin{aligned}
\mathrm{P}\big(\hat{w}_L(g(\tau); \tilde{\alpha}) \leq Q_B(\tau) \text{ for all } \tau \in [0,1]\big) &\geq \mathrm{P}\big(\hat{w}_L(g(\tau); \tilde{\alpha}) \leq Q_W(g(\tau)) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(\hat{w}_L(\tau; \tilde{\alpha}) \leq Q_W(\tau) \text{ for all } \tau \in [0,1]\big) \\
&= 1 - \alpha,
\end{aligned}$$

where the first inequality is from Theorem 9, the first equality is from the fact that $g(\cdot)$ is continuous and increasing from $g(0) = 0$ to $g(1) = 1$, and the last equality is from Theorem 5.

For the upper confidence bound function,

$$\begin{aligned}
\mathrm{P}\big(\hat{w}_U(h(\tau); \tilde{\alpha}) \geq Q_B(\tau) \text{ for all } \tau \in [0,1]\big) &\geq \mathrm{P}\big(\hat{w}_U(h(\tau); \tilde{\alpha}) \geq Q_W(h(\tau)) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(\hat{w}_U(\tau; \tilde{\alpha}) \geq Q_W(\tau) \text{ for all } \tau \in [0,1]\big) \\
&= 1 - \alpha,
\end{aligned}$$

where similarly the (in)equalities are respectively from Theorem 9, the fact that $h(\cdot)$ is continuous and increasing from $h(0) = 0$ to $h(1) = 1$, and Theorem 5.

For the two-sided band,

$$
\begin{aligned}
&\mathrm{P}\big(\hat{w}_L(g(\tau);\tilde{\alpha}) \le Q_B(\tau) \le \hat{w}_U(h(\tau);\tilde{\alpha}) \text{ for all } \tau \in [0,1]\big) \\
&\ge \mathrm{P}\big(\hat{w}_L(g(\tau);\tilde{\alpha}) \le Q_W(g(\tau)) \text{ and } Q_W(h(\tau)) \le \hat{w}_U(h(\tau);\tilde{\alpha}) \text{ for all } \tau \in [0,1]\big) \\
&= \mathrm{P}\big(\hat{w}_L(\tau;\tilde{\alpha}) \le Q_W(\tau) \text{ and } Q_W(\tau) \le \hat{w}_U(\tau;\tilde{\alpha}) \text{ for all } \tau \in [0,1]\big) \\
&= 1-\alpha,
\end{aligned}
$$

for the same reasons given above for the one-sided bands. □

C Pointwise inference

Here we provide methods and results for (pointwise) confidence intervals for individual quantiles and interquantile ranges of the bid distribution, in the setting with a fixed number of bidders. We also provide quantile confidence intervals with a varying number of bidders.

Assumption A5 is a regularity condition that guarantees we can apply other prior results to obtain high-order pointwise coverage accuracy in Corollaries 13 and 14. Assumption A5 combined with the strictly increasing CDF in A1 implies $f_W(Q_W(\tau)) > 0$.

Assumption A5. At the quantile index of interest τ, the transaction price PDF $f_W(\cdot)$ exists and its second derivative $f''_W(\cdot)$ is continuous in a neighborhood of $Q_W(\tau)$.

C.1 Quantiles

Using the identity in Proposition 1, Corollary 12 says generally that we can construct a confidence interval (CI) for $Q_B(\tau)$ using any quantile CI for W, after transforming τ per Proposition 1.

Corollary 12. *Under A1 and A2, any CI for $Q_W(p)$ also covers $Q_B(\tau)$ with the same coverage probability, where $p = F_\beta(\tau)$.*

Proof of Corollary 12. Immediate from Proposition 1. □

Corollary 12 is a general result that allows us to use any established pointwise CI for $Q_W(\cdot)$ to be the CI for $Q_B(\cdot)$ after a transformation of the quantile index. For example, without the independence of A3, we can use a time-series quantile CI for $Q_W(p)$ as the CI for $Q_B(\tau)$. If we further assume independent auctions as in A3, then we can get a more precise CI with high-order accuracy as in Corollary 13.

We provide some intuition for the accuracy in the case of iid B_i (and thus iid W_j and β_j). Under A3 and (1) by the probability integral transform

$$F_\beta(\beta_j) \overset{iid}{\sim} \text{Unif}(0,1) \implies F_\beta(\beta_{J:r}) \sim \text{Beta}(r, J+1-r). \tag{26}$$

Thus, the coverage probability (CP) of an upper one-sided confidence interval $[W_{J:r}, \infty)$ using integer order statistic is

$$\begin{aligned} \text{CP} \equiv \text{P}(W_{J:r} \leq F_B^{-1}(\tau)) &= \text{P}(\beta_{J:r} \leq \tau) \\ &= \text{P}(F_\beta(\beta_{J:r}) \leq F_\beta(\tau)) = \text{P}(\text{Beta}(r, J+1-r) \leq F_\beta(\tau)), \end{aligned} \tag{27}$$

which is a known value given the known r and τ (and J, k, and n). That is, given a particular r and τ, we can compute the exact finite-sample coverage probability. However, given a desired τ and confidence level $1-\alpha$, generally there is no integer r that sets (27) equal to $1-\alpha$ exactly. We can either choose integer r to be conservative (CP strictly above $1-\alpha$), or we can try to get closer to $1-\alpha$ at the cost of some approximation error. Specifically, Goldman and Kaplan (2017) show how using fractional order statistics like Hutson (1999) yields only $O(J^{-3/2}[\log J]^3)$ error. Note this is an asymptotic error with fixed n and $J \to \infty$.

Corollary 13. *Under Assumptions A1–A3 and A5, with $p = F_\beta(\tau)$ as in Proposition 1, the level $1-\alpha$ confidence intervals for $Q_W(p)$ in Goldman and Kaplan's (2017) Theorem 4 have coverage probability $1-\alpha+O(J^{-3/2}[\log(J)]^3)$, or exactly $1-\alpha$ in the special case where the fractional order statistic is an integer order statistic.*

Proof of Corollary 13. The result follows from our Proposition 1 and Goldman and Kaplan's (2017) Theorem 4. That is, under A2, A3, and A5, Goldman and Kaplan's (2017) Theorem

4 provides one-sided and two-sided CIs that cover $Q_W(p)$ with $1-\alpha+O(J^{-3/2}[\log(J)]^3)$ coverage probability, and under A1 and A2, Proposition 1 says $Q_W(p) = Q_B(\tau)$ with $p = F_\beta(\tau)$. Thus, these CIs cover $Q_B(\tau)$ with coverage probability $1-\alpha+O(J^{-3/2}[\log(J)]^3)$. Goldman and Kaplan (2017) also show that the CP of these CIs is exactly $1-\alpha$ when the interpolating weights are 0 or 1, as also seen in our (27). $\square$

C.2 Interquantile ranges

In addition, Proposition 1 facilitates CIs for any interquantile range (IQR) $Q_B(\tau_2)-Q_B(\tau_1)$ of the underlying bid distribution. Specifically, Proposition 1 translates any IQR of B to a particular IQR of W, and Goldman and Kaplan (2018b, §3.2) provide one-sided CIs for any IQR of W.

Corollary 14. *Under A1–A3 and A5, for confidence level $1-\alpha$ and any $0 < \tau_1 < \tau_2 < 1$, Theorem 3.2 of Goldman and Kaplan (2018b) provides confidence intervals for $Q_B(\tau_2) - Q_B(\tau_1)$ with $1-\alpha+O(J^{-1/2}\log(J))$ coverage probability, by taking the corresponding confidence interval for $Q_W(p_2)-Q_W(p_1)$ with $p_1 = F_\beta(\tau_1)$ and $p_2 = F_\beta(\tau_2)$.*

Proof of Corollary 14. Under A2, A3, and A5, for any $0 < p_1 < p_2 < 1$, Theorem 3.2 of Goldman and Kaplan (2018b) provides CIs for $Q_W(p_2)-Q_W(p_1)$ with $1-\alpha+O(J^{-1/2}\log(J))$ coverage probability. Under A1 and A2, Proposition 1 says $Q_B(\tau_2)-Q_B(\tau_1) = Q_W(p_2) - Q_W(p_1)$ with $p_1 = F_\beta(\tau_1)$ and $p_2 = F_\beta(\tau_2)$. Thus, the CIs for $Q_W(p_2)-Q_W(p_1)$ cover $Q_B(\tau_2)-Q_B(\tau_1)$ with $1-\alpha+O(J^{-1/2}\log(J))$ coverage probability, too. $\square$

C.3 Quantiles with varying number of bidders

At a given $1-\alpha$ confidence level, we choose r_{1L} that satisfies

$$\mathrm{P}(\beta_{J:r_{1L}} \le \tau) \ge 1-\alpha \text{ and } \mathrm{P}(\beta_{J:r_{1L}+1} \le \tau) < 1-\alpha, \tag{28}$$

choose r_{1U} that satisfies

$$\mathrm{P}(\beta_{J:r_{1U}} \geq \tau) \geq 1-\alpha \text{ and } \mathrm{P}(\beta_{J:r_{1U}-1} \geq \tau) < 1-\alpha, \tag{29}$$

and choose any (r_{2L}, r_{2U}) satisfying

$$\begin{aligned} &\mathrm{P}(\beta_{J:r_{2L}} \leq \tau) \geq 1-\alpha/2 \text{ and } \mathrm{P}(\beta_{J:r_{2L}+1} \leq \tau) < 1-\alpha/2; \\ &\mathrm{P}(\beta_{J:r_{2U}} \geq \tau) \geq 1-\alpha/2 \text{ and } \mathrm{P}(\beta_{J:r_{2U}-1} \geq \tau) < 1-\alpha/2. \end{aligned} \tag{30}$$

The upper and lower one-sided $[W_{J:r_{1L}}, \infty)$ and $(-\infty, W_{J:r_{1U}}]$ and two-sided $[W_{J:r_{2L}}, W_{J:r_{2U}}]$ are conservative CIs that cover $Q_B(\tau)$ with at least $1-\alpha$ CP. The (r_{2L}, r_{2U}) can be uniquely determined if we impose enough additional objectives (like minimizing $r_{2U} - r_{2L}$).

Corollary 15. *Under A1, A3, and A4, the upper one-sided CI* $[W_{J:r_{1L}}, \infty)$ *with* r_{1L} *satisfying* (28) *covers* $Q_B(\tau)$ *with exact known CP* $\mathrm{P}(\beta_{J:r_{1L}} \leq \tau)$*; the lower one-sided CI* $(-\infty, W_{J:r_{1U}}]$ *with* r_{1U} *satisfying* (29) *covers* $Q_B(\tau)$ *with exact known CP* $\mathrm{P}(\beta_{J:r_{1U}} \geq \tau)$*; the two-sided CI* $[W_{J:r_{2L}}, W_{J:r_{2U}}]$ *with* r_{2L} *and* r_{2U} *satisfying* (30) *covers* $Q_B(\tau)$ *with exact known CP* $\mathrm{P}(\beta_{J:r_{2L}} \leq \tau \leq \beta_{J:r_{2U}})$.

Proof of Corollary 15. By simulation, given (11) under A3 and A4, we know the values of $\mathrm{P}(\beta_{J:r} \leq \tau)$ for $r = 1, \ldots, J$, up to arbitrarily small simulation error. We can choose r_{1L}, r_{1U}, r_{2L}, and r_{2U} satisfying (28)–(30) accordingly. The CP of these CIs is at least $1-\alpha$ and comes from the general equalities

$$\mathrm{P}(W_{J:r} \leq F_B^{-1}(\tau)) = \mathrm{P}(F_B(W_{J:r}) \leq \tau) = \mathrm{P}(\beta_{J:r} \leq \tau).$$

For example, the CP of the two-sided CI is

$$\begin{aligned} \mathrm{P}(W_{J:r_{2L}} \leq F_B^{-1}(\tau) \leq W_{J:r_{2U}}) &= \mathrm{P}(F_B(W_{J:r_{2L}}) \leq \tau \leq F_B(W_{J:r_{2U}})) \\ &= \mathrm{P}(\beta_{J:r_{2L}} \leq \tau \leq \beta_{J:r_{2U}}) \text{ by A1} \\ &= \mathrm{P}(\beta_{J:r_{2L}} \leq \tau) - \mathrm{P}(\beta_{J:r_{2U}} \leq \tau) \\ &= \mathrm{P}(\beta_{J:r_{2L}} \leq \tau) + \mathrm{P}(\beta_{J:r_{2U}} \geq \tau) - 1 \end{aligned}$$

$$\geq 1 - \alpha/2 + 1 - \alpha/2 - 1 = 1 - \alpha. \qquad \square$$

References

Aldor-Noiman, Sivan, Lawrence D. Brown, Andreas Buja, Wolfgang Rolke, and Robert A. Stine. 2013. "The Power to See: A New Graphical Test of Normality." *The American Statistician* 67 (4):249–260. URL https://doi.org/10.1080/00031305.2013.847865.

Andreyanov, Pasha and Grigory Franguridi. 2024. "Nonparametric inference on counterfactuals in first-price auctions." Working paper available at https://arxiv.org/abs/2106.13856.

Aradillas-López, Andrés, Amit Gandhi, and Daniel Quint. 2013. "Identification and Inference in Ascending Auctions With Correlated Private Values." *Econometrica* 81 (2):489–534. URL https://doi.org/10.3982/ECTA9431.

Armstrong, Timothy B. 2013. "Bounds in auctions with unobserved heterogeneity." *Quantitative Economics* 4 (3):377–415. URL https://doi.org/10.3982/QE167.

Athey, Susan and Philip A. Haile. 2002. "Identification of Standard Auction Models." *Econometrica* 70 (6):2107–2140. URL https://www.jstor.org/stable/3081982.

———. 2007. "Nonparametric Approaches to Auctions." In *Handbook of Econometrics*, vol. 6A, edited by James J. Heckman and Edward E. Leamer, chap. 60. Elsevier, 3847–3965. URL https://doi.org/10.1016/S1573-4412(07)06060-6.

Bapat, R. B. and M. I. Beg. 1989. "Order Statistics for Nonidentically Distributed Variables and Permanents." *Sankhyā: The Indian Journal of Statistics, Series A* 51 (1):79–93. URL http://www.jstor.org/stable/25050725.

Buja, Andreas and Wolfgang Rolke. 2006. "Calibration for Simultaneity: (Re)Sampling Methods for Simultaneous Inference with Applications to Function Estimation and Functional Data." Working paper, available at http://stat.wharton.upenn.edu/~buja/PAPERS/paper-sim.pdf.

Chesher, Andrew and Adam M. Rosen. 2017. "Incomplete English auction models with heterogeneity." Working paper available at https://econpapers.repec.org/paper/ifscemmap/27_2f17.htm.

Compiani, Giovanni, Philip Haile, and Marcelo Sant'Anna. 2020. "Common Values, Unobserved Heterogeneity, and Endogenous Entry in US Offshore Oil Lease Auctions." *Journal of Political Economy* 128 (10):3872–3912. URL https://doi.org/10.1086/710026.

Fisher, Ronald Aylmer. 1932. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, 4th ed.

Freyberger, Joachim and Bradley J. Larsen. 2022. "Identification in ascending auctions, with an application to digital rights management." *Quantitative Economics* 13 (2):505–543. URL https://doi.org/10.3982/QE1151.

Gimenes, Nathalie. 2017. "Econometrics of Ascending Auctions by Quantile Regression." *Review of Economics and Statistics* 99 (5):944–953. URL https://www.jstor.org/stable/26616171.

Gimenes, Nathalie and Emmanuel Guerre. 2022. "Quantile regression methods for first-price auctions." *Journal of Econometrics* 226 (2):224–247. URL https://doi.org/10.1016/j.jeconom.2021.02.009.

Glueck, D. H., A. Karimpour-Fard, J. Mandel, L. Hunter, and K. E. Muller. 2008. "Fast Computation by Block Permanents of Cumulative Distribution Functions of Order Statistics from Several Populations." *Communications in Statistics - Theory and Methods* 37 (18):2815–2824. URL https://doi.org/10.1080/03610920802001896.

Goldman, Matt and David M. Kaplan. 2017. "Fractional order statistic approximation for nonparametric conditional quantile inference." *Journal of Econometrics* 196 (2):331–346. URL https://doi.org/10.1016/j.jeconom.2016.09.015.

———. 2018a. "Comparing distributions by multiple testing across quantiles or CDF values." *Journal of Econometrics* 206 (1):143–166. URL https://doi.org/10.1016/j.jeconom.2018.04.003.

———. 2018b. "Non-parametric inference on conditional quantile differences and linear combinations, using L-statistics." *The Econometrics Journal* 21 (2):136–169. URL https://doi.org/10.1111/ectj.12095.

Grundl, Serafin and Yu Zhu. 2023. "Robust inference in first-price auctions: Overbidding as an identifying restriction." *Journal of Econometrics* 235 (2):484–506. URL https://doi.org/10.1016/j.jeconom.2022.06.001.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong. 2000. "Optimal Nonparametric Estimation of First-price Auctions." *Econometrica* 68 (3):525–574. URL https://www.jstor.org/stable/2999600.

———. 2009. "Nonparametric Identification of Risk Aversion in First-Price Auctions under Exclusion Restrictions." *Econometrica* 77 (4):1193–1227. URL https://www.jstor.org/stable/40263858.

Haile, Philip A. and Yuichi Kitamura. 2019. "Unobserved heterogeneity in auctions." *The Econometrics Journal* 22 (1):C1–C19. URL https://doi.org/10.1111/ectj.12121.

Haile, Philip A. and Elie Tamer. 2003. "Inference with an Incomplete Model of English Auctions." *Journal of Political Economy* 111 (1):1–51. URL https://www.jstor.org/stable/10.1086/344801.

Hutson, Alan D. 1999. "Calculating nonparametric confidence intervals for quantiles using fractional order statistics." *Journal of Applied Statistics* 26 (3):343–353. URL https://doi.org/10.1080/02664769922458.

Kaplan, David M. 2019. "distcomp: Comparing distributions." *Stata Journal* 19 (4):832–848. URL https://doi.org/10.1177/1536867x19893626.

Luo, Yao. 2020. "Unobserved heterogeneity in auctions under restricted stochastic dominance." *Journal of Econometrics* 216 (2):354–374. URL https://doi.org/10.1016/j.jeconom.2019.07.009.

Menzel, Konrad and Paolo Morganti. 2013. "Large sample properties for estimators based on the order statistics approach in auctions." *Quantitative Economics* 4 (2):329–375. URL https://doi.org/10.3982/QE177.

Neyman, Jerzy. 1937. "»Smooth test» for goodness of fit." *Skandinavisk Aktuarietidskrift* 20 (3–4):149–199. URL https://doi.org/10.1080/03461238.1937.10404821.

Pearson, Karl. 1933. "On a Method of Determining Whether a Sample of Size n Supposed to Have Been Drawn From a Parent Population Having a Known Probability Integral has Probably Been Drawn at Random." *Biometrika* 25:379–410. URL https://doi.org/10.

1093/biomet/25.3-4.379.
R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.
Syrgkanis, Vasilis, Elie Tamer, and Juba Ziani. 2021. "Inference on Auctions with Weak Assumptions on Information." Working paper available at https://scholar.harvard.edu/files/tamer/files/bce_econometrics.pdf.
Wilks, Samuel Stanley. 1962. *Mathematical Statistics*. New York: Wiley. URL https://www.worldcat.org/oclc/602895854.