# Nonparametric inference on counterfactuals in first-price auctions*

Pasha Andreyanov† Grigory Franguridi‡

June 14, 2022

**Abstract**

In a classical model of the first-price sealed-bid auction with independent private values, we develop nonparametric estimation and inference procedures for a class of policy-relevant metrics, such as total expected surplus and expected revenue under counterfactual reserve prices. Motivated by the linearity of these metrics in the quantile function of bidders' values, we propose a bid spacings-based estimator of the latter and derive its Bahadur-Kiefer expansion. This makes it possible to construct exact uniform confidence bands and assess the optimality of a given auction rule. Using the data on U.S. Forest Service timber auctions, we test whether setting zero reserve prices in these auctions was revenue maximizing.

**JEL Classfication:** C57

**Keywords:** first-price auction, uniform inference, quantile density, spacings, counterfactual reserve price, Bahadur-Kiefer expansion, USFS auctions

*This paper supersedes the earlier version entitled "Simple nonparametric inference for first-price auctions via bid spacings" available at https://arxiv.org/pdf/2106.13856v1.pdf. A python package implementing the methodology in the paper is available at https://pypi.org/project/simple-fpa

†Faculty of Economic Sciences, HSE University. Email: pandreyanov@gmail.com

‡Department of Economics, University of Southern California. Email: franguri@usc.edu

# 1 Introduction

In the empirical studies of first-price auctions, a structural approach to estimation and inference is often used. This approach exploits restrictions derived from economic theory to recover bidders' latent valuations from the observed bids. With these valuations at hand, the researcher can make predictions about the effects of changes in auction rules or composition of bidders. Various methods, in both parametric and nonparametric frameworks, have been developed, see, e.g., Paarsch et al. (2006), Athey and Haile (2007), and Perrigne and Vuong (2019) for an overview.

Since the seminal papers by Elyakime et al. (1994), Guerre et al. (2000) and Li et al. (2000), it is the probability density function (PDF) of bidders' values that has been considered a default medium containing the model primitives. This choice is natural since it allows for constructive identification (Matzkin, 2013) when valuations are independent,[1] and since various counterfactual quantities can be expressed as functionals of the value density. However, the standard estimator of the value density (Guerre et al., 2000) is a two-step nonparametric procedure not amenable to simple theoretical analysis.[2] More importantly, most counterfactuals are *nonlinear* functionals of the value density, rendering rigorous counterfactual inference prohibitively hard. This leads to researchers reporting confidence intervals based on simulation from the estimated PDF (e.g., Li and Perrigne, 2003) or none at all.

Our main contribution is developing the methodology for nonparametric estimation and uniform inference for a class of important counterfactual quantities, such as bidder's expected surplus, total expected surplus, and expected revenue under counterfactual reserve prices. This methodology relies on the *quantile function of valuations* — an alternative candidate for constructive identification — instead of the PDF, and on recognizing that many counterfactuals are *continuous linear functionals* of this quantile function.

Since the value quantile function is the key ingredient of our counterfactual evaluation, we provide its complete first-order asymptotic analysis. Namely, we derive the uniform, asymptotically linear (Bahadur-Kiefer, or BK) expansion for the kernel estimator $\hat{v}_h$ of the value quantile function $v$, where $h$ is a smoothing bandwidth. This expansion implies that, despite converging to a Gaussian distribution pointwise, the estimator does not admit a functional central limit theorem, which calls for alternative ways of conducting uniform inference. Luckily, the linear term of the studentized estimator is *known* and *pivotal*, allowing us to suggest simple simulation-based confidence bands and establish their validity using the anti-concentration theory of Chernozhukov et al. (2014).

[1] With correlated valuations, nonparametric identification is partial, see Aradillas-López et al. (2013).

[2] For example, uniform inference for the value density was only developed recently in Ma et al. (2019).

With the theoretical properties of the value quantile function at hand, we move towards the analysis of the estimators of the aforementioned counterfactuals. We show that these can be divided into two broad classes. One class contains the "smoother" (w.r.t. the value quantile function) counterfactuals that are estimable at the parametric rate $n^{-1/2}$ and converge weakly to a Gaussian process in $\ell^\infty[0,1]$. The other class contains the "less smooth" counterfactuals that are only estimable at the slower rate $(nh)^{-1/2}$ and do not converge weakly in $\ell^\infty[0,1]$. For each class, we develop a distinct protocol for construction of confidence intervals and bands, and establish their validity.

To demonstrate how our methodology can be used to answer a concrete economic question, we use Phil Haile's data on U.S. Forest Service timber auctions, where a reserve price was never employed, and assess the optimality of this auction design. Namely, we test whether the seller's expected revenue could have been increased, had the auction designer chosen a nonzero reserve price.

Our work contributes to the expanding literature on quantile methods in first-price auctions, see Marmer and Shneyerov (2012) and Enache and Florens (2017) for kernel-based estimators, Luo and Wan (2018) for isotone regression-based estimators, and Guerre and Sabbah (2012) and Gimenes and Guerre (2021) for local polynomial estimators. Interestingly, our estimator of valuation quantiles is a weighted sum of the differences of ordered bids, often referred to as *bid spacings*. The latter have been used for collusion detection in Ingraham (2005), for set identification of bidders' rents in Paul and Gutierrez (2004) and Marra (2020), and in the prior-free clock auction design in Loertscher and Marx (2020).

Finally, Zincenko (2021) has recently shown that it is possible to perform uniform inference on the seller's expected revenue with a pseudo-value-based estimator. The validity of his confidence bands relies on either the bootstrap delta method or extreme value theory for kernel density estimators. Our quantile-based approach, on the other hand, does not require bootstrapping or extreme value approximations, is applicable to a larger class of counterfactuals and may be computationally advantageous due to a natural choice of the grid and the possibility of using the Fast Fourier Transform for discrete convolution.

The rest of the paper is organized as follows. In Section 2, we set up the theoretical and econometric framework for our analysis. In Section 3, we develop estimation and inference on the value quantile function. In Section 4, we develop estimation and inference on the counterfactual quantities of interest. In Section 5, we provide the Monte Carlo simulations of the finite-sample coverage of our confidence bands. In Section 6, we use the timber auction data to test whether counterfactual reserve prices are revenue-enhancing. Section 7 contains a discussion of some practical aspects of our methodology. Section 8 concludes the paper. Proofs of theoretical results are provided in the Appendix.

# 2 Framework

## 2.1 Auction model

Our setting is a sealed-bid first-price auction with independent private values and a (potentially binding) reserve price.

There are $M \geqslant 2$ *potential bidders* in the auction, who are ex ante identical and risk-neutral.[3] A potential bidder becomes *active* if two conditions are met: her valuation exceeds the publicly announced reserve price $\underline{r} \geqslant 0$ (i.e. the lowest price at which the auctioneer is willing to sell the auctioned object), and she passes an exogenous (independent of her valuation, identity and the reserve price) selection procedure. Every active bidder submits a bid $b$ without observing the number of active bidders or their bids. Naturally, each bidder assigns the same belief $p_m$ to the event that there are exactly $m$ active bidders in the auction. We assume away the degenerate case $p_0 + p_1 = 1$.[4] The object is won by the highest bidder who pays the face value of her bid to the auctioneer.

Denote the valuation of a potential bidder by $v$. We impose the following assumption on the value distribution (Guerre et al., 2009, Definition 2).

**Assumption 1** (Distribution of values)**.** *The values $v_1, \ldots, v_M$ of potential bidders are drawn independently from a common CDF $\tilde{G}$ such that:*

1. *The support of $\tilde{G}$ has the form $[\underline{v}, \bar{v}]$, where $0 \leqslant \underline{v} < \bar{v} < +\infty$.*
2. *$\tilde{G}$ is twice continuously differentiable on $[\underline{v}, \bar{v}]$.*
3. *$\tilde{g}(v) = \tilde{G}'(v) > 0$ for all $v \in [\underline{v}, \bar{v}]$.*

We assume that the primitives $\tilde{G}, \underline{r}, (p_m)_{m=1}^{M}$ of the model are common knowledge. Therefore, the equilibrium behavior depends only on the distribution of valuations of active bidders; this distribution has the CDF

$$G(v) = \begin{cases} 0, & v < \underline{r}, \\ \frac{\tilde{G}(v) - \tilde{G}(\underline{r})}{1 - \tilde{G}(\underline{r})}, & v \geqslant \underline{r}. \end{cases} \tag{1}$$

We denote the associated PDF by $g(v) = G'(v)$. If the reserve price is not binding, i.e. $\tilde{G}(\underline{r}) = 0$, then $G = \tilde{G}$.

[3] For risk-averse bidders with known CRRA utilities, the analysis is similar, see, e.g., Zincenko (2021).

[4] We emphasize that, although the equilibrium beliefs $(p_m)_{m=1}^{M}$ depend on the reserve price $\underline{r}$ as well as the unspecified selection procedure, its nature is irrelevant as long as the beliefs are identical, see, e.g., Krishna (2009, Section 3.2.2).

Define the auxiliary functions

$$A_1(u) := \sum_{m=1}^{M} p_m u^{m-1}, \quad A_2(u) := uA_1(u), \tag{2}$$

$$A_3(u) := (1-u)A_1(u), \quad A(u) := \frac{A_1(u)}{A_1'(u)}, \tag{3}$$

where $A_1'(u) \neq 0$ since $\sum_{m=2}^{M} p_m > 0$ by non-degeneracy of the beliefs.

The bidding strategy $\beta(v)$ of the symmetric Bayes-Nash equilibrium can be characterized either via the first order conditions

$$\beta'(v) = \frac{(v-\beta(v))g(v)}{A(G(v))}, \quad \beta(\underline{v}) = \underline{v}, \tag{4}$$

or the envelope conditions

$$\beta(v) = v - \frac{\int_{\underline{r}}^{v} A_1(G(x))\,dx}{A_1(G(v))}, \tag{5}$$

see, e.g., Riley and Samuelson (1981) or Krishna (2009). Clearly, this strategy is weakly increasing and twice continuously differentiable for all $v \in [\underline{v}, \bar{v}]$. Equations (4) and (5) together imply that $\beta(v) < v$ and $\beta'(v) > 0$ for all $v > \underline{v}$ and, moreover, $\beta'(\underline{v}) = 1/(1 + A'(0)) > 0$ by L'Hôpital's rule. Consequently, the strategy is strictly increasing with the slope $\beta'(v) > 0$, for all $v \in [\underline{v}, \bar{v}]$.

Denoting by $F$ the CDF of the equilibrium bid, and $f = F'$, the inverse bidding strategy can be written as

$$v = \beta^{-1}(b) = b + \frac{A(F(b))}{f(b)}, \tag{6}$$

allowing to recover the latent values from the observed bids. This suggests a nonparametric estimation approach that was popularized by Guerre et al. (2000) and Li et al. (2000).

Alternatively, we can rewrite the equation (6) in terms of the quantiles of the participating values. Denote by $Q(u) := F^{-1}(u)$ the *bid quantile function* and by $q(u) := Q'(u)$ the associated *bid quantile density*. Let $v(u) := G^{-1}(u)$ be the $u$-th quantile of the participating value distribution $G$. Then equation (6) can be rewritten as

$$v(u) = \beta^{-1}(Q(u)) = Q(u) + A(u)q(u), \tag{7}$$

where we use the change of variables $b = Q(u)$ along with the identities $F(Q(u)) = u$ and $f(Q(u))q(u) = 1$. Since, by definition, $Q(u) = \beta(G^{-1}(u))$, and both $(G^{-1})'(v)$ and $\beta'(v)$ are strictly positive for all $v \in [\underline{v}, \bar{v}]$, we arrive at the following property of the equilibrium distribution of bids.

**Proposition 1** (Distribution of bids)**.** *Under Assumption 1, the equilibrium bids are drawn independently from a distribution with a quantile function $Q$ such that:*

1. *$Q$ is twice continuously differentiable on $[0,1]$,*
2. *$q(u) = Q'(u) > 0$ for all $u \in [0,1]$.*

## 2.2 Counterfactuals

The counterfactual experiment of interest is the increase in the reserve price from $\underline{r}$ to $r^* > \underline{r}$.[5] We show that a variety of counterfactual metrics can be written in terms of the counterfactual reserve price $r^*$ and the distribution $G$ of bids submitted under the *original* reserve price $\underline{r}$, despite the fact that the counterfactual experiment is accompanied by the change in the beliefs about the number of participants due to endogenous entry (see the last row of Table 1). We then show that, in our model, these counterfactual metrics can be rewritten as *linear* functionals of the value quantile function $v(\cdot)$, the key observation enabling simple inference procedures in Section 4.

One such counterfactual is the total expected (ex ante) surplus. In a symmetric equilibrium, it is ex post equal to the highest valuation if it exceeds $r^*$, and zero otherwise, which is a random variable with CDF $A_2(G(\cdot))$. Hence the total surplus is its expectation

$$TS(r^*) \coloneqq \int_{r^*}^{\bar{v}} x d\left(A_2(G(x))\right). \tag{8}$$

Another counterfactual is bidder's expected surplus. We will distinguish the interim surplus of a potential bidder $\pi^p(v)$ from that of an active bidder $\pi^a(v)$. By the revenue equivalence principle (see Krishna, 2009), the interim expected utility of an active bidder is related to her equilibrium probability of winning, equal to $A_1(G(v))$, via the envelope conditions

$$\pi^a(v) \coloneqq \int_{r^*}^{v} A_1(G(x))\, dx. \tag{9}$$

The interim expected utility of a potential bidder only differs from that of an active bidder by a factor of $a$, the expected participation rate (the ratio of the number of active bidders

[5] We do not consider the counterfactual decrease in the reserve price since no identification is possible in this case, unless, of course, $\underline{r}$ is non-binding (which is itself a non-testable assumption in our framework). This is due to the fact that we never observe the bidders whose valuations are smaller than $\underline{r}$.

to the number of potential bidders) under the original reserve price $\underline{r}$,[6]

$$\pi^p(v) := a \cdot \pi^a(v), \quad a := \frac{1}{M}\sum_{m=1}^{M} m p_m. \tag{10}$$

To derive the potential bidder's expected (ex ante) surplus $BS$, we need to take the expectation of $\pi^p(v)$ w.r.t. the distribution of $v$. Integration by parts yields the formula

$$BS(r^*) := \int_{r^*}^{\bar{v}} \pi^p(x) dG(x) = a \int_{r^*}^{\bar{v}} \int_{r^*}^{v} A_1(G(x)) dx d(G(v) - 1) = a \int_{r^*}^{\bar{v}} A_3(G(x)) dx.$$

Finally, we consider the seller's expected revenue under the counterfactual reserve price $r^*$, which is equal to the difference between the total expected surplus and $M$ times the potential bidder's expected surplus,

$$\begin{aligned} \mathrm{Rev}(r^*) &:= TS - M \cdot BS = \int_{r^*}^{\bar{v}} x d\left(A_2(G(x))\right) - Ma \int_{r^*}^{\bar{v}} A_3(G(x)) dx = \\ &= MaA_3(G(r^*))r^* + \int_{r^*}^{\bar{v}} x d\left(A_2(G(x)) + MaA_3(G(x))\right), \end{aligned}$$

and the bidder's strategy under $r^*$,

$$\beta_{r^*}(v) = v - \frac{\int_{r^*}^{v} A_1(G(x)) dx}{A_1(G(v))}, \quad v \geqslant r^*. \tag{11}$$

It can be seen that all the aforementioned counterfactuals are complicated, nonlinear functionals of the primitives $(r^*, G)$. However, using change of variables $z = G(x)$ (i.e. passing to the *ranks* of valuations from their levels) and denoting $u^* = G(r^*)$ yields expressions that are *linear* in the quantile function $v(\cdot)$, see Table 1. This makes $v(\cdot)$ a key object needed for the counterfactual analysis.

### 2.3 Data generating process

The observed data is a random sample of bids $\{b_{il},\ i = 1, \ldots, m_l,\ l = 1, \ldots, L\}$, where $b_{il}$ denotes the bid submitted by the $i$-th participant in the $l$-th auction. All the auctions are ex ante symmetric and independent, and $m_l$ is the number of participants in the $l$-th auction.

[6] We emphasize that endogenous participation is fully captured by the $A_1, A_2, A_3, A$ functions and the $a$ constant, and so we do not need the counterfactual participation rate in our calculations.

| | Classical (nonlinear) form | Quantile (linear in $v$) form |
|---|---|---|
| total expected surplus | $\int_{r^*}^{\bar{v}} x\, d\left(A_2(G(x))\right)$ | $\int_{u^*}^{1} A_2'(z) v(z)\, dz$ |
| potential bidder's expected surplus | $a \int_{r^*}^{\bar{v}} A_3(G(x))\, dx$ | $-aA_3(u^*)v(u^*) - a\int_{u^*}^{1} A_3'(z)v(z)\, dz$ |
| expected revenue | $MaA_3(G(r^*))r^* + \int_{r^*}^{\bar{v}} x\, d\left(A_2(G(x)) + MaA_3(G(x))\right)$ | $MaA_3(u^*)v(u^*) + \int_{u^*}^{1} \left(A_2'(z) + MaA_3'(z)\right) v(z)\, dz$ |
| optimal bid given value $v = v(u)$ | $v - \dfrac{\int_{r^*}^{v} A_1(G(x))\, dx}{A_1(G(v))}$ | $\dfrac{A_1(u^*)}{A_1(u)} v(u^*) + \int_{u^*}^{u} \dfrac{A_1'(z)}{A_1(u)} v(z)\, dz$ |
| probability of $m$ active bidders given $r^*$ | $\sum_{i=m}^{M} p_i \binom{i}{m} (1-G(r^*))^m G(r^*)^{i-m}$ | $\sum_{i=m}^{M} p_i \binom{i}{m} (1-u^*)^m (u^*)^{i-m}$ |

Table 1: Typical counterfactuals in first-price auctions as linear functionals of $v(\cdot)$.

For brevity, we denote the (random) sample size by $n = n(L) = \sum_{l=1}^{L} m_l$, and define

$$b_1 := b_{11},\ b_2 := b_{21},\ \ldots\ , b_n := b_{m_L L}. \tag{12}$$

Note that, since the bidders do not know the realizations of the number of active bidders, the samples $\{b_1, \ldots, b_n\}$ and $\{m_1, \ldots, m_L\}$ are independent. Besides, as $L \to \infty$, the sample size $n(L) \to \infty$ with probability one. Therefore, without loss of generality, we condition our subsequent exposition on a realization $\{m_l\}_{l=1}^{\infty}$ such that $n(L) \to \infty$. This has the following important implication: although the auxiliary functions $A_1, A_2, A_3, A$ and the constant $a$ need to be estimated from the data, we can assume that they are known, since their estimators only depends on the conditioning variables $m_1, \ldots, m_L$, see equations (16)-(18) below.

# 3 Estimation and inference for value quantiles

As explained in Section 2.2, the value quantile function $v(\cdot)$ is the key object needed for the counterfactual analysis. In this section we develop the asymptotic theory for its natural (plug-in) estimator. To define the estimator, we need to introduce two auxiliary objects.

The first object is the kernel estimator of the bid quantile density $q(u)$, defined by

$$\hat{q}_h(u) := \int_0^1 K_h(u - z)\, d\hat{Q}(z), \quad u \in [0, 1]. \tag{13}$$

Here $K$ is a compactly supported kernel, $K_h(z) := h^{-1}K\left(h^{-1}z\right)$, $h > 0$ is a bandwidth, and $\hat{Q}(u)$ is the empirical bid quantile function,

$$\hat{Q}(u) = \begin{cases} b_{(\lfloor nu \rfloor + 1)}, & u \in [0, 1), \\ b_{(n)}, & u = 1, \end{cases} \tag{14}$$

where $b_{(1)} \leqslant \ldots \leqslant b_{(n)}$ are the order statistics of the observed bids $b_1, \ldots, b_n$. We note that $\hat{q}_h$ takes the form of a weighted sum of *bid spacings* $b_{(i+1)} - b_{(i)}$,

$$\hat{q}_h(u) = \sum_{i=1}^{n-1} K_h(u - i/n) \left(b_{(i+1)} - b_{(i)}\right). \tag{15}$$

This estimator was previously studied by Siddiqui (1960) and Bloch and Gastwirth (1968) for the case of rectangular kernel, and by Falk (1986), Welsh (1988), Csörgő et al. (1991) and Jones (1992) for general kernels.

The second auxiliary object is the plug-in estimators of $A_1, A_2, A_3, A$ and $a$ defined by

$$\check{A}_1(u) := \sum_{m=1}^{M} \check{p}_m u^{m-1}, \quad \check{A}_2(u) := u\check{A}_1(u), \quad \check{A}_3(u) := (1-u)\check{A}_1(u), \tag{16}$$

$$\check{A}(u) := \frac{\check{A}_1(u)}{\check{A}_1'(u)}, \quad \check{a} := \frac{1}{M}\sum_{m=1}^{M} m\check{p}_m, \tag{17}$$

where $\check{p}_m$ is the empirical frequency of auctions with $m$ bidders,

$$\check{p}_m := \frac{1}{L}\sum_{l=1}^{L} 1(m_l = m), \quad m = 1, \ldots, M. \tag{18}$$

We use the "check" (as opposed to "hat") notation here to highlight that $\check{A}_1, \check{A}_2, \check{A}_3, \check{A}, \check{a}$ are treated as known since, as explained in Section 2.3, our analysis is conditional on $m_1, \ldots, m_L$.

Given $\check{A}$ and $\hat{q}_h$, we define our estimator of the value quantile $v(u)$ by

$$\hat{v}_h(u) := \hat{Q}(u) + \check{A}(u)\hat{q}_h(u), \quad u \in [0, 1]. \tag{19}$$

We note that $\hat{v}_h$ consists of two parts: (i) the empirical quantile function $\hat{Q}$ that is uniformly consistent and converges to a Gaussian process in $\ell^\infty[0,1]$ at the parametric rate $n^{-1/2}$, and (ii) the kernel quantile density $\hat{q}_h$ that is uniformly consistent only away from the boundary $\{0, 1\}$ and does *not* converge to a (tight) limit in $\ell^\infty[\varepsilon, 1-\varepsilon]$ even if $\varepsilon > 0$, but converges pointwise to a Gaussian limit at the nonparametric rate $(nh)^{-1/2}$, see the proof of Theorem 1. Therefore, the first-order asymptotic properties of $\hat{v}_h$ are determined by the kernel quantile density $\hat{q}_h$.

We impose the following assumptions.

**Assumption 2** (Kernel function)**.**

1. *$K : \mathbb{R} \to \mathbb{R}$ is a nonnegative function such that*

$$\int_{\mathbb{R}} K(z)\, dz = 1 \text{ and } R_K \coloneqq \int_{\mathbb{R}} K(z)^2\, dz < \infty. \tag{20}$$

2. *$K$ is a Lipschitz function supported on the interval $[-1, 1]$.*

**Assumption 3** (Bandwidth)**.** *The bandwidth $h = h_n$ is such that $h \to 0$ and*

1. *there exist $c > 0$ and $\alpha > 0$ such that $h_n \geqslant cn^{-1/2+\alpha}$ for all $n$.*

2. *there exist $C > 0$ and $\beta > 0$ such that $h_n \leqslant Cn^{-1/3-\beta}$ for all $n$.*

Assumption 2.1 states that $K$ is a valid, square-integrable PDF. Assumption 2.2 is standard in the literature on strong approximations of local empirical processes (see, e.g., Rio, 1994). In particular, it implies that $K$ is a function of bounded variation, which is crucial in our derivation of the BK expansion. Assumption 3.1 states that $h_n$ decays at a rate that is slower than $n^{-1/2}$. This assumption is mild and guarantees that the Gaussian approximation to the supremum of our (studentized) estimator has at least the rate $o_p(\log^{-1/2} n)$. This rate is needed to establish validity of the confidence bands in Section 3.2. Finally, Assumption 3.2 imposes undersmoothing that eliminates the smoothing bias in $\hat{v}_h$ and nonsmooth counterfactuals, see Section 4.2.

## 3.1 The Bahadur–Kiefer expansion

In this section, we derive the Bahadur–Kiefer (i.e. almost sure, uniform, asymptotically linear) representation of the form

$$\frac{\sqrt{nh}(\hat{v}_h(u) - v(u))}{\hat{q}_h(u)} = -\check{A}(u)\mathbb{G}_{n,h}(u) + R_n(u), \quad u \in [h, 1-h], \tag{21}$$

where

$$\mathbb{G}_{n,h}(u) \coloneqq \sqrt{nh} \cdot \frac{1}{n} \sum_{i=1}^{n} \left[K_h\left(u - F(b_i)\right) - \mathbb{E}K_h\left(u - F(b_i)\right)\right] \tag{22}$$

and the remainder $R_n(u)$ converges to zero a.s. uniformly in $u \in [h, 1-h]$ with an explicit rate.

The key feature of this representation is that the main term is fully *known* and *pivotal*: its distribution does not depend on the data generating process since $U_i := F(b_i) \sim$ iid Uniform$[0,1]$. Heuristically, this suggests that the distribution of the left-hand side under *any* DGP is a valid approximation for its true distribution. Indeed, in Section 3.2, we show the validity of such approximation by combining pivotality with the anti-concentration theory of Chernozhukov et al. (2014). This leads to a simple algorithm for the confidence bands on the quantile function $v(\cdot)$.

To derive this representation, we rely on the classical BK expansion of the quantile function (Bahadur, 1966; Kiefer, 1967),

$$\hat{Q}(u) - Q(u) = -q(u)\left(\hat{F}(Q(u)) - u\right) + r_n(u), \tag{23}$$

$$\text{where } r_n(u) = O_{a.s.}\left(n^{-3/4}\ell(n)\right) \text{ uniformly in } u \in [0,1]. \tag{24}$$

Here $\ell(n) = (\log n)^{1/2}(\log\log n)^{1/4}$ is a logarithmic offset factor that arises due to the uniform nature of the approximation and may often be disregarded in practice. Note that the BK expansion represents a *nonlinear* estimator $\hat{Q}(u)$ as a sum of the *linear* estimator — the empirical distribution function $\hat{F}(Q(u))$ — and the remainder $r_n(u)$ that converges to zero a.s. uniformly at a nonparametric (slow) rate $n^{-3/4}\ell(n)$.

**Theorem 1** (Bahadur-Kiefer expansion for value quantiles)**.** *Under the Assumptions 1 and 2, the estimator $\hat{v}_h(u)$ has the representation*

$$Z_n(u) = Z_n^*(u) + R_n(u), \quad u \in [h, 1-h], \tag{25}$$

*where*

$$Z_n(u) := \frac{\sqrt{nh}\left(\hat{v}_h(u) - v(u)\right)}{\hat{q}_h(u)}, \quad Z_n^*(u) := -\breve{A}(u)\mathbb{G}_{n,h}(u), \tag{26}$$

$$R_n(u) = O_{a.s.}\left(n^{1/2}h^{3/2} + h^{1/2} + h^{-1/2}n^{-1/4}\ell(n)\right) \text{ uniformly in } u \in [h, 1-h]. \tag{27}$$

**Remark 1** (BK expansion for quantile density)**.** *The proof of the preceding theorem also implies the BK expansion for the normalized quantile density $\sqrt{nh}\left(\hat{q}_h(u) - q_h(u)\right)$, which may be of independent interest. In this case, the right-hand side does not have the factor $\breve{A}(u)$, while the term $h^{1/2}$ in the remainder rate can be replaced by the faster term $h\log h$.*

We note that two types of biases arise in the estimation of $v(\cdot)$. The first type of bias is the *smoothing bias* $\mathbb{E}\hat{v}_h(u) - v(u)$ which manifests in the term $n^{1/2}h^{3/2}$ in the remainder rate.

This bias can be eliminated by undersmoothing $h = o(n^{-1/3})$, i.e. choosing a (suboptimally) small bandwidth such that $\sqrt{nh}(\mathbb{E}\hat{v}_h(u) - v(u)) \to 0$. Conversely, if the rate of $h$ is larger than $n^{-1/3}$, as in the case of Silverman's rule-of-thumb bandwidth $h^{rot} = O\left(n^{-1/5}\right)$, the confidence bands will be centered at $\mathbb{E}\hat{v}_h(u)$ rather than $v$. This conflict between MSE-optimal estimation and correct inference is a feature of most nonparametric estimators, see Horowitz (2001) and Hall (2013).

The other type of bias is the *boundary bias*, stemming from the estimator $\hat{v}_h(u)$ being inconsistent when $u$ is close to the boundary $\{0, 1\}$ of its domain $[0, 1]$. Because our interest is in valid hypothesis testing, and not the confidence bands *per se*, we can eliminate this bias by introducing the trimming $u \in [h, 1-h]$ while maintaining the validity of inference procedures based on the representation (25).

## 3.2 Inference on value quantiles

Theorem 1 allows us to construct pointwise confidence intervals and uniform confidence bands for the value quantile function. In particular, the following corollary provides the asymptotic distribution of the estimator of a fixed valuation quantile.

**Corollary 1.** *Under the Assumptions 1, 2 and 3, we have, for every* $u \in (0, 1)$*,*

$$\sqrt{nh}\left(\hat{v}_h(u) - v(u)\right) \rightsquigarrow N(0, V(u)), \tag{28}$$

$$V(u) := A^2(u) q^2(u) R_K. \tag{29}$$

Special cases of this result (for the quantile density estimator $\hat{q}_h$) were derived by Siddiqui (1960) and Bloch and Gastwirth (1968). It implies that a confidence interval of nominal confidence level $(1-\alpha)$ for $v_h(u)$ can be constructed as

$$\left[\hat{v}_h(u) - \frac{\check{A}(u)\hat{q}_h(u)\sqrt{R_K}}{\sqrt{nh}} z_{1-\frac{\alpha}{2}}, \quad \hat{v}_h(u) + \frac{\check{A}(u)\hat{q}_h(u)\sqrt{R_K}}{\sqrt{nh}} z_{1-\frac{\alpha}{2}}\right], \tag{30}$$

where $z_{1-\frac{\alpha}{2}}$ is the standard normal quantile of level $1 - \frac{\alpha}{2}$.

We now turn to the problem of uniform inference on $v(\cdot)$.

Note that if the process $Z_n$ converged weakly in $\ell^\infty[h, 1-h]$ to a known (or estimable) process $Z$, this would have enabled construction of asymptotically valid confidence bands by using the quantiles of $\sup_u |Z(u)|$ as critical values.[7] Unfortunately, although $Z_n(u)$ is asymptotically Gaussian at each point $u \in (0, 1)$, it does *not* converge in $\ell^\infty[h, 1-h]$. This follows from the fact that the main term $Z_n^*$ in the BK expansion is the scaled kernel density

[7]For one-sided confidence bands, one would use the quantiles of $\sup_u Z_n(u)$ instead.

process, which is known to lack functional convergence (see, e.g., Rio, 1994).[8] In such a case, there are two common ways to circumvent the problem and derive valid confidence intervals.

One approach is to derive the asymptotic distribution of a normalized version of $\sup_u Z_n(u)$ using extreme value theory, and then rely on the knowledge of the normalizing constants to construct the confidence band. In the case of kernel and histogram density estimation, this approach was pioneered by Smirnov (1950) and Bickel and Rosenblatt (1973).[9] However, convergence to the asymptotic distribution turns out to be very slow, leading to the coverage error of the resulting confidence band to be $O(1/\log n))$, as shown by Hall (1991).

The other approach is to rely on finite-sample approximations for (the distribution of) the supremum

$$W_n \coloneqq \sup_{u\in[h,1-h]} |Z_n(u)|\,. \tag{31}$$

If such an approximation admits simulation, it can be used for construction of confidence bands. This is the approach we take in this paper.

We consider two types of approximations, both of which are *pivotal*, and hence allow for simulation. One is simply the supremum of the linear term $Z_n^*$, viz.

$$W_n^* \coloneqq \sup_{u\in[h,1-h]} |Z_n^*(u)|\,. \tag{32}$$

The other is the supremum of $Z_n$ under an alternative, uniform[0,1] distribution of bids, viz.

$$W_n^{U[0,1]} \coloneqq \sup_{u\in[h,1-h]} \left|Z_n^{U[0,1]}(u)\right|, \tag{33}$$

where $Z_n^{U[0,1]}(u)$ is the process $Z_n(u)$ calculated using the pseudo-sample

$$\{\tilde{b}_i\}_{i=1}^n \sim \text{iid Uniform}[0,1]. \tag{34}$$

This approximation is nonstandard and makes use of the asymptotic pivotality of $W_n$. In principle, any distribution of the pseudo-bids rationalized by a value distribution satisfying Assumption 1 can be chosen; however, the uniform distribution is convenient since, in this

[8] For an example of a sequence of stochastic processes on $[0,1]$ that weakly converges pointwise, but not in $\ell^\infty[0,1]$, consider $X_n(u) = h_n^{-1/2}\left(B\left(u+h_n\right) - B\left(u\right)\right)$, where $B$ is the Brownian motion, $h_n \to 0$ and $u \in (0,1)$. Clearly, $X_n(u) \rightsquigarrow N(0,1)$ for all $u$, but, by Lévy's modulus of continuity theorem, $\sup_u |X_n(u)| \to \infty$ a.s., and so there is no convergence in $\ell^\infty[0,1]$.

[9] For a nonasymptotic version of Smirnov-Bickel-Rosenblatt extreme value theorem, see Rio (1994, Theorem 1.2).

case, we have, for all $u \in [0,1]$,

$$Q(u) \coloneqq u, \quad q(u) \coloneqq 1, \quad v(u) \coloneqq u + A(u), \tag{35}$$

and hence

$$Z_n^{U[0,1]}(u) \coloneqq \sqrt{nh}\left(\hat{v}_h(u;\{\tilde{b}_i\}_{i=1}^n) - u - \check{A}(u)\right). \tag{36}$$

We emphasize that it is not immediate that the distributions of $W_n^*$ and $W_n^{U[0,1]}$ approximate the distribution of $W_n$ in a way that guarantees the validity of the associated confidence bands

$$\left[\hat{v}_h(u) - \frac{\hat{q}_h(u)c_{n,1-\alpha/2}}{\sqrt{nh}},\ \hat{v}_h(u) + \frac{\hat{q}_h(u)c_{n,1-\alpha/2}}{\sqrt{nh}}\right], \quad u \in [h, 1-h], \tag{37}$$

where $c_{n,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of either $W_n^*$ or $W_n^{U[0,1]}$. Indeed, note that Theorem 1 implies the inequality

$$|W_n - W_n^*| = \left| \sup_{u\in[h,1-h]} |Z_n(u)| - \sup_{u\in[h,1-h]} |Z_n^*(u)| \right| \leqslant \sup_{u\in[h,1-h]} |R_n(u)|, \tag{38}$$

and hence the coupling

$$W_n = W_n^* + r_n, \tag{39}$$

where $r_n$ tends to zero a.s. at a known rate. If one could show that this implies Kolmogorov convergence

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(W_n \leqslant t) - \mathbb{P}(W_n^* \leqslant t)| \to 0, \tag{40}$$

then the confidence bands based on $W_n^*$ would be valid. However, (40) need not follow from (39) even if the a.s. convergence rate of $r_n$ is very fast, unless further conditions are imposed on $W_n^*$.

As an illustration of this phenomenon, consider an abstract example $W_n = n^{-1}U$, $W_n^* = n^{-1}(U-1)$, where $U \sim \text{Uniform}[0,1]$. Then $r_n = W_n - W_n^* = n^{-1}$, but

$$\mathbb{P}(W_n \leqslant 0) - \mathbb{P}(W_n^* \leqslant 0) = 0 - 1 = -1 \not\to 0,$$

and so (40) does not hold. On the other hand, if $W_n^*$ had an absolutely continuous asymptotic distribution $\mathcal{D}$, then the CDF of $W_n$ would converge to the CDF of $\mathcal{D}$ pointwise, and hence the quantiles of $\mathcal{D}$ would serve as valid critical values.

Therefore, intuitively, a certain degree of *anti-concentration* of $W_n^*$ is needed to guarantee that the coupling (39) implies Kolmogorov convergence (40) and hence validity of simulated critical values. The anti-concentration literature mainly focuses on Gaussian processes, while the process $Z_n^*$ is non-Gaussian. Fortunately, $Z_n^*$ is the normalized kernel density estimator for uniform data, which is a well-studied process. In particular, we rely on the seminal work Chernozhukov et al. (2014) to establish a coupling of $W_n$ with the supremum of a Gaussian process and show that the latter exhibits sufficient anti-concentration. We then argue that an identical argument works for $W_n^*$. Finally, the pivotality of $W_n^*$ and the coupling (39) imply the Kolmogorov convergence for $W_n^{U[0,1]}$. Formally, we have the following result.

**Theorem 2.** *Under the Assumptions 1, 2 and 3,*

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(W_n \leqslant x) - \mathbb{P}(W_n^* \leqslant x)| \to 0, \tag{41}$$

$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}(W_n \leqslant x) - \mathbb{P}(W_n^{U[0,1]} \leqslant x)\right| \to 0, \tag{42}$$

*and hence the confidence bands (37) are asymptotically valid and exact.*

**Remark 2.** *For the purpose of constructing the* one-sided *confidence bands, we note that the same result holds with $W_n$, $W_n^*$, and $W_n^{U[0,1]}$ replaced by $\sup_{u\in[h,1-h]} Z_n(u)$, $\sup_{u\in[h,1-h]} Z_n^*(u)$, and $\sup_{u\in[h,1-h]} Z_n^{U[0,1]}(u)$, respectively.*

# 4 Estimation and inference for counterfactuals

In this section, we develop the asymptotic theory for the counterfactuals in Table 1 which heavily relies on the analysis of the estimator of value quantiles in the previous section.

Clearly, every such counterfactual has the general form

$$T(u^*) := \varphi(u^*)v(u^*) + \int_{u^*}^1 \psi(x)v(x)\,dx, \tag{43}$$

where $\varphi$ and $\psi$ are continuously differentiable functions on $[0,1]$ that only depend on the auxiliary objects $A_1, A_2, A_3, A$ and $a$ (or their derivatives). As an example, for the total expected surplus $\varphi(x) \equiv 0$ and $\psi(x) = A_2'(x)$, while for the expected revenue $\varphi(u^*) = MaA_3(u^*)$ and $\psi(x) = A_2'(x) + MaA_3'(x)$.

The representation (43) implies $T(u^*)$ is a (weighted) sum of two continuous linear functionals of $v$ of different smoothness: (i) evaluation at a point $v(u^*)$ and (ii) integration

$$S_\psi(u^*) := \int_{u^*}^1 \psi(x)v(x)\,dx. \tag{44}$$

The natural estimators of the two components have fundamentally different asymptotic properties. Namely, in Section 3.1 we showed that the less smooth functional $v(u^*)$ is only estimable at the nonparametric rate $(nh)^{-1/2}$ and does *not* converge in $\ell^\infty[\varepsilon, 1-\varepsilon]$ even for $\varepsilon > 0$. On the other hand, we will show in Section 4.1 that the smoother functional $S_\psi(u^*)$ is estimable at the parametric rate $n^{-1/2}$ and converges to a Gaussian process in $\ell^\infty[0,1]$. We will combine the two results in Section 4.2 to show that, whenever $\varphi \neq 0$, inference on $T$ can be performed similarly to that on $v$.

### 4.1 Smooth ($S$-type) counterfactuals

First, let us consider estimation and inference for functionals (44), where $\psi : [0,1] \to \mathbb{R}$ is a known, continuously differentiable function.

To motivate our estimator, use integration by parts to rewrite

$$S_\psi(u^*) = \int_{u^*}^1 \psi(u)Q(u)\,du + \int_{u^*}^1 \breve{A}(u)\psi(u)\,dQ(u) \tag{45}$$

$$= \int_{u^*}^1 \psi(u)Q(u)\,du + \breve{A}(u)\psi(u)Q(u)|_{u^*}^1 - \int_{u^*}^1 Q(u)\,d[\breve{A}(u)\psi(u)] = \tag{46}$$

$$= \int_{u^*}^1 \psi(u)Q(u)\,du + \breve{A}(u)\psi(u)Q(u)|_{u^*}^1 - \int_{u^*}^1 Q(u)(\breve{A}'(u)\psi(u) + \breve{A}(u)\psi'(u))\,du \tag{47}$$

$$= \int_{u^*}^1 \chi_\psi(u)Q(u)\,du - \breve{A}(u^*)\psi(u^*)Q(u^*) + \breve{A}(1)\psi(1)Q(1), \tag{48}$$

where we denote

$$\chi_\psi(u) := (1 - \breve{A}'(u))\psi(u) - \breve{A}(u)\psi'(u). \tag{49}$$

Note that the latter formula expresses $S_\psi(u^*)$ as a continuous linear functional of the quantile function $Q$, which is estimable at a parametric rate. This leads to a natural estimator of $S_\psi$ that does not contain tuning parameters. Namely, for any $u^* \in [0,1]$, we define the estimator

$$\hat{S}_\psi(u^*) \coloneqq \int_{u^*}^{1} \chi_\psi(u)\hat{Q}(u)\,du - \check{A}(u^*)\psi(u^*)\hat{Q}(u^*) + \check{A}(1)\psi(1)\hat{Q}(1) \tag{50}$$

$$= \sum_{i=\lfloor nu^* \rfloor}^{n-1} b_{(i+1)} \int_{\frac{\max(i,nu^*)}{n}}^{\frac{i+1}{n}} \chi_\psi(u)\,du - \check{A}(u^*)\psi(u^*)b_{(\lfloor nu^* \rfloor+1)} + \check{A}(1)\psi(1)b_{(n)}. \tag{51}$$

The following theorem establishes standard Gaussian process asymptotics for $\hat{S}_\psi$.

**Theorem 3.** *Under Assumptions 1, 2 and 3,*

$$\sqrt{n}(\hat{S}_\psi(\cdot) - S_\psi(\cdot)) \rightsquigarrow \mathcal{G}_{\psi,q}(\cdot) \text{ in } \ell^\infty[0,1], \tag{52}$$

*where $\mathcal{G}_{\psi,q}(\cdot)$ is a tight, centered Gaussian process on $[0,1]$ with the covariance function*

$$\mathbb{E}\mathcal{G}_{\psi,q}(u^*)\mathcal{G}_{\psi,q}(v^*) = Cov\left(f_{u^*}(U), f_{v^*}(U)\right), \quad U \sim uniform[0,1], \quad u^*, v^* \in [0,1], \tag{53}$$

*and*

$$f_{u^*}(U) \coloneqq -\int_{u^*}^{1} \chi_\psi(u)q(u)1(U \leqslant u)\,du + \check{A}(u^*)\psi(u^*)q(u^*)1(U \leqslant u^*). \tag{54}$$

**Remark 3.** *The integral in the expression for $\hat{S}_\psi(u^*)$ can be replaced by $\chi_\psi(\zeta(i))/n$ for any $\zeta(i) \in \left[\frac{i}{n}, \frac{i+1}{n}\right]$. This would have no impact on the statement of Theorem 3.*

## 4.2 Nonsmooth ($T$-type) counterfactuals

Given the asymptotic results for the two components of the generic counterfactual (43), we may now turn to estimation and inference on the latter.

To this end, define the estimator

$$\hat{T}_h(u^*) = \check{\varphi}(u^*)\hat{v}_h(u^*) + \hat{S}_{\check{\psi}}(u^*), \quad u^* \in [h, 1-h], \tag{55}$$

where $\hat{v}_h(u^*)$ is defined in (19) and $\hat{S}_{\check{\psi}}(u^*)$ is the estimator (51) with $\psi = \check{\psi}$. Since $\hat{S}_{\check{\psi}}(u^*)$ converges fast, while $\hat{v}_h(u^*)$ converges slowly, the asymptotics of $\hat{T}_h(u^*)$ is dominated by the latter, as illustrated by the following theorem.

**Theorem 4.** *Under the Assumptions 1 and 2, we have the representation*

$$Z_n^T(u^*) = Z_n^{T*}(u^*) + R_n^T(u^*), \quad u^* \in [h, 1-h], \tag{56}$$

*where*

$$Z_n^T(u^*) := \frac{\sqrt{nh}\left(\hat{T}_h(u^*) - T(u^*)\right)}{\hat{q}_h(u^*)}, \quad Z_n^{T*}(u^*) := -\check{\varphi}(u^*)\check{A}(u^*)\mathbb{G}_{n,h}(u^*), \tag{57}$$

$$R_n^T(u^*) := O_{a.s.}\left(n^{1/2}h^{3/2} + h^{1/2} + h^{-1/2}n^{-1/4}\ell(n)\right) \text{ uniformly in } u^* \in [h, 1-h]. \tag{58}$$

Theorem 4 immediately yields the following result on the asymptotic distribution of $\hat{T}_h(u^*)$ at a fixed point $u^* \in (0, 1)$.

**Corollary 2.** *Under the Assumptions 1, 2 and 3, we have, for every $u^* \in (0, 1)$,*

$$\sqrt{nh}(\hat{T}_h(u^*) - T(u^*)) \rightsquigarrow N(0, V(u^*)), \tag{59}$$

$$V(u^*) = (A(u^*)q(u^*)\varphi(u^*))^2 R_K. \tag{60}$$

We now consider uniform inference on $T(\cdot)$. Since the estimator $\hat{T}_h(\cdot)$ does not converge in $\ell^\infty[h, 1-h]$, but the approximating process $Z_n^{T*}$ is known and pivotal, the methodology will be similar to the case of the valuation quantile function, see Section 3.2. In particular, we show that valid confidence bands for $T(\cdot)$ can be based on simulation from either (i) the approximating process $Z_n^{T*}$, or (ii) the process $Z_n^T$ under an alternative distribution of bids. To this end, define

$$W_n^T := \sup_{u^* \in [h,1-h]} \left|Z_n^T(u^*)\right|, \tag{61}$$

$$W_n^{T*} := \sup_{u^* \in [h,1-h]} \left|Z_n^{T*}(u^*)\right|, \tag{62}$$

$$W_n^{T,U[0,1]} := \sup_{u^* \in [h,1-h]} \left|Z_n^{T,U[0,1]}(u^*)\right|, \tag{63}$$

where $Z_n^{T,U[0,1]}(u^*)$ is the process $Z_n^T(u^*)$ calculated using the pseudo-sample

$$\{\tilde{b}_i\}_{i=1}^n \sim \text{iid Uniform}[0, 1], \tag{64}$$

cf. equations (31)-(34). Define the confidence bands by

$$\left[\hat{T}_h(u^*) - \frac{\hat{q}_h(u^*)c_{n,1-\alpha/2}}{\sqrt{nh}}, \hat{T}_h(u^*) + \frac{\hat{q}_h(u^*)c_{n,1-\alpha/2}}{\sqrt{nh}}\right], \quad u \in [h, 1-h], \tag{65}$$

where $c_{n,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of either $W_n^{T*}$ or $W_n^{T,U[0,1]}$.

**Theorem 5.** *Under the Assumptions 1, 2 and 3, we have*

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(W_n^T \leqslant t) - \mathbb{P}(W_n^{T*} \leqslant t)\right| \to 0, \tag{66}$$

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(W_n^T \leqslant t) - \mathbb{P}(W_n^{T,U[0,1]} \leqslant t)\right| \to 0, \tag{67}$$

*and hence the confidence bands (65) are asymptotically valid and exact.*

**Remark 4.** *For the purpose of constructing the* one-sided *confidence bands, we note that the same result holds with $W_n^T$, $W_n^{T*}$, and $W_n^{T,U[0,1]}$ replaced by $\sup_{u\in[h,1-h]} Z_n^T(u)$, $\sup_{u\in[h,1-h]} Z_n^{T*}(u)$, and $\sup_{u\in[h,1-h]} Z_n^{T,U[0,1]}(u)$, respectively.*

**Remark 5** (Shape of confidence bands)**.** *Note that, for any function $\iota(\cdot)$ bounded away from zero on $[0,1]$, the representation (56) is equivalent to*

$$Z_n^T(u^*)/\iota(u^*) = Z_n^{T*}(u^*)/\iota(u^*) + \tilde{R}_n^T(u^*), \quad u^* \in [h, 1-h], \tag{68}$$

*where $\tilde{R}_n^T(u^*)$ has the same uniform convergence rate as $R_n^T(u^*)$. Similarly to (65), the two-sided confidence bands based on such representation are*

$$\left[\hat{T}_h(u^*) - \frac{\iota(u^*)\hat{q}_h(u^*)c_{n,1-\alpha/2}}{\sqrt{nh}}, \hat{T}_h(u^*) + \frac{\iota(u^*)\hat{q}_h(u^*)c_{n,1-\alpha/2}}{\sqrt{nh}}\right], \quad u \in [h, 1-h], \tag{69}$$

*where $c_{n,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of either of the random variables*

$$W_n^{T,\iota} := \sup_{u^*\in[h,1-h]} \left|Z_n^{T*}(u^*)/\iota(u^*)\right|, \tag{70}$$

$$W_n^{T,U[0,1],\iota} := \sup_{u^*\in[h,1-h]} \left|Z_n^{T,U[0,1]}(u^*)/\iota(u^*)\right|, \tag{71}$$

*These bands stay asymptotically exact for any $\iota$, but have the shape $\iota(u^*)\hat{q}_h(u^*)$, which may affect their finite-sample performance and asymptotic power.*[10]

# 5 Monte Carlo experiments

While our theoretical results establish the asymptotic validity of the confidence bands, they do not rule out a substantial finite-sample size distortion. In this section, we evaluate the extent of this distortion in a set of Monte Carlo experiments.

[10] Montiel Olea and Plagborg-Møller (2019) discuss a similar issue with simultaneous confidence bands for a *vector* (rather than functional) parameter.

| Estimand | (i) | (ii) | (iii) | (iv) | (v) |
|---|---|---|---|---|---|
| | Sample size = 1,000, trim = 3% | | | | |
| beta(1,1) | 0.95 | 0.952 | 0.912 | 0.91 | 0.974 |
| beta(2,2) | 0.954 | 0.954 | 0.912 | 0.904 | 0.97 |
| beta(5,2) | 0.952 | 0.954 | 0.924 | 0.916 | 0.966 |
| beta(2,5) | 0.956 | 0.962 | 0.902 | 0.898 | 0.968 |
| powerlaw(2) | 0.952 | 0.952 | 0.928 | 0.922 | 0.976 |
| powerlaw(3) | 0.948 | 0.948 | 0.93 | 0.926 | 0.978 |
| | Sample size = 10,000, trim = 1.5% | | | | |
| beta(1,1) | 0.95 | 0.948 | 0.932 | 0.936 | 0.96 |
| beta(2,2) | 0.954 | 0.954 | 0.932 | 0.934 | 0.96 |
| beta(5,2) | 0.952 | 0.954 | 0.93 | 0.932 | 0.962 |
| beta(2,5) | 0.952 | 0.952 | 0.918 | 0.93 | 0.958 |
| powerlaw(2) | 0.954 | 0.952 | 0.94 | 0.938 | 0.96 |
| powerlaw(3) | 0.948 | 0.952 | 0.934 | 0.938 | 0.96 |
| | Sample size = 100,000, trim = .7% | | | | |
| beta(1,1) | 0.95 | 0.948 | 0.938 | 0.942 | 0.954 |
| beta(2,2) | 0.952 | 0.948 | 0.944 | 0.946 | 0.956 |
| beta(5,2) | 0.954 | 0.952 | 0.944 | 0.948 | 0.956 |
| beta(2,5) | 0.956 | 0.952 | 0.932 | 0.948 | 0.954 |
| powerlaw(2) | 0.944 | 0.948 | 0.948 | 0.948 | 0.954 |
| powerlaw(3) | 0.946 | 0.948 | 0.952 | 0.95 | 0.952 |

Table 2: Simulated coverage of the 95% uniform confidence bands.

For simplicity, we simulate an auction with exactly two bidders ($M = 2$ and $p_2 = \check{p}_2 = 1$) and a non-binding (original) reserve price $\underline{r} = 0$. We consider three choices for the distributions of the observed bids: uniform, beta and power-law; all of which are supported on the interval $[0, 1]$. The simplest choice is the uniform distribution, since it has a strictly positive density with $q(u) = 1$. The beta$(\alpha, \beta)$ distribution features a bell-shaped density for $\alpha, \beta > 1$, with varying skewness. In contrast, the density $f(x) = \alpha x^{\alpha-1}$ of the power-law distribution increases on the support $[0, 1]$ for $\alpha > 1$. We censor these distributions at top 5% and bottom 5% quantile levels, so that the quantile density is strictly positive and satisfies the statement of Proposition 1.[11]

The estimation targets are (i) the bid quantile function $q$, (ii) the value quantile function $v$; and the following quantities as functions of the counterfactual reserve price: (iii) the potential bidder's expected surplus, (iv) the expected revenue, and (v) the total expected surplus, see Table 1.

[11]The censoring of the distribution in the Monte Carlo simulations is achieved by replacing their true quantile function $Q(u)$ with $(Q(0.05 + 0.9u) - Q(0.05))/(Q(0.95) - Q(0.05))$. We emphasize that every simulated bid distribution can be rationalized by some value distribution satisfying Assumption 1.

For the non-counterfactual targets (i), (ii) and for the $T$-type counterfactuals (iii), (iv), we calculate the confidence bands by simulation from the left-hand side of the respective BK expansions under the uniform[0,1] bid distribution, see Theorem 5. For the $S$-type functional (v), the confidence bands are constructed by simulation from the estimated process

$$\hat{\mathcal{G}}(u^*) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( \hat{f}_{u^*}(U_i) - \mathbb{E}\, \hat{f}_{u^*}(U_i) \right), \tag{72}$$

where $U_i \sim$ iid Uniform$[0,1]$ and $\hat{f}_{u^*}$ is equal to $f_{u^*}$ with the true values $q$ replaced by their estimates $\hat{q}_h$, see Theorem 3. We use the undersmoothing bandwidth $h = 1.06 s \cdot n^{-0.34}$, where $s$ is the standard deviation of bid spacings, and set both the number of DGP simulations and the number of simulations for the critical values to 500.

The results are shown in Table 2. The simulated coverage can be seen to be close to the nominal level of 0.95 for larger sample sizes, which validates our theoretical results in Sections 3 and 4.

# 6 Empirical application

In this section, we apply our methodology to test the hypothesis about the optimality of the auction design employed in timber sales held by the U.S. Forest Service in between 1974 and 1989, see, e.g., Haile (2001). These auctions did not feature a reserve price (i.e. $\underline{r} = 0$) which raises the question whether the collected revenue could have been higher, had the reserve price been set at a positive level. We use the data kindly provided by Phil Haile on his website.[12]

We select a subsample of auctions that are sealed-bid and have at least 2 bidders, see Figure 1. As is common in the literature, we residualize the log-bids using available auction-level characteristics: year and location dummies, the logarithms of the tract advertised value and the Herfindahl index.[13] The latter is a measure of homogeneity of the tract with respect to the timber species. This procedure is consistent with a multiplicative model of observed auction heterogeneity.[14] The distribution of bid residuals is truncated at the 5-th percentile on each end, leaving 60758 observations, see Figure 1.

[12] `http://www.econ.yale.edu/~pah29/`

[13] The exponentiated log-bid residuals (to which we will refer simply as *bid residuals*) are interpreted as estimates of the idiosyncratic component of bids, while the exponentiated fitted values are interpreted as estimates of the common component of bids, see Haile et al. (2003).

[14] The asymptotic distributions of the test statistics are not affected by the error in the residualization procedure, as long as the estimates of the common component of bids converge at a faster (in our case parametric) rate, see Haile et al. (2003) and Athey and Haile (2007).

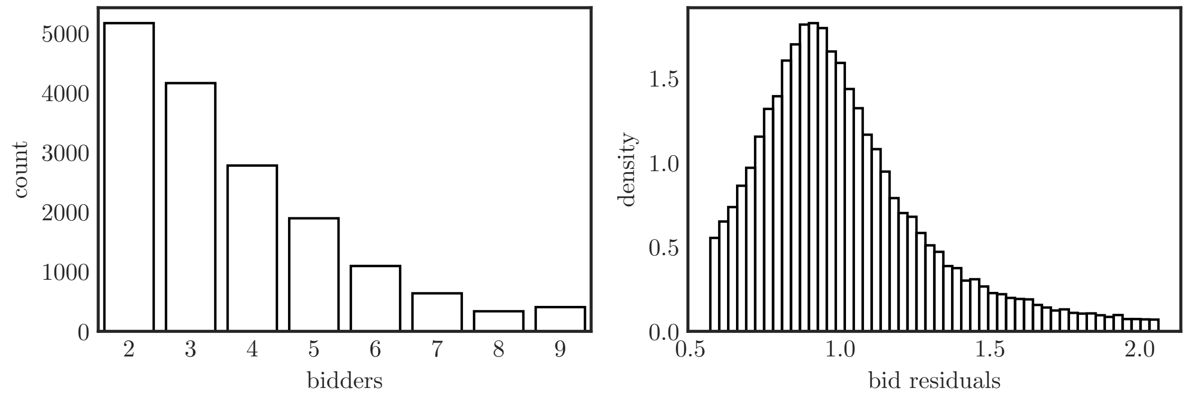


Figure 1: Distributions of the number of bidders (left) and of bid residuals (right).

To assess the change $\Delta(u^*)$ in the expected revenue, we use the quantile version of the revenue formula in Table 1, which yields

$$\Delta(u^*) := \text{Rev}(u^*) - \text{Rev}(0) = MaA_3(u^*)v(u^*) - \int_0^{u^*} (A_2'(x) + MaA_3'(x))\, v(x)\, dx,$$

where $u^* > 0$ is the counterfactual exclusion level (rank of the counterfactual reserve price in the distribution of valuations) and we used the fact that in our application $v(0) = \underline{r}$. Since $\Delta(u^*)$ is similar to the $T$-type functional (43), its estimator is

$$\hat{\Delta}_h(u^*) := \varphi(u^*)\hat{v}_h(u^*) - \left[\int_0^{u^*} \chi_\psi(u)\hat{Q}(u)\, du + \check{A}(u^*)\psi(u^*)\hat{Q}(u^*) - \check{A}(0)\psi(0)\hat{Q}(0)\right], \quad (73)$$

where $\varphi(x) := Ma\check{A}_3(x)$, $\psi(x) := A_2'(x) + MaA_3'(x)$ and $\chi_\psi$ is defined in (49).

We use the undersmoothing bandwidth $h = 1.06s \cdot n^{-0.34}$, where $s$ is the standard deviation of spacings of bid residuals, and evaluate $\hat{\Delta}_h(u^*)$ on the evenly spaced grid $\{i/n\}_{i=0}^n$. This bandwidth is slightly smaller than the Silverman rule of thumb bandwidth $h = 1.06s \cdot n^{-1/5}$.

To construct the confidence bands, we use the representation in Theorem 4 and Remark 5. First, 1000 realizations of the bid quantile density $\hat{q}_h^U(\cdot)$ are simulated, independently from the data, based on pseudo-bids from the uniform[0,1] distribution. Second, for a nominal confidence level $(1-\alpha)$, the critical value $c_{n,1-\alpha}$ is computed as the $(1-\alpha)$-quantile of $\sup_{u\in[h,1-h]}(\hat{q}_h^U(u) - 1)$. Finally the one-sided confidence band is computed as

$$\left(\hat{\Delta}_h(u) - MaA_3(u)\check{A}(u) \cdot \hat{q}_h(u) \cdot c_{n,1-\alpha},\ +\infty\right), \quad u \in [h, 1-h]. \quad (74)$$

We test the hypothesis of nonexistence of a counterfactual (positive) reserve price that

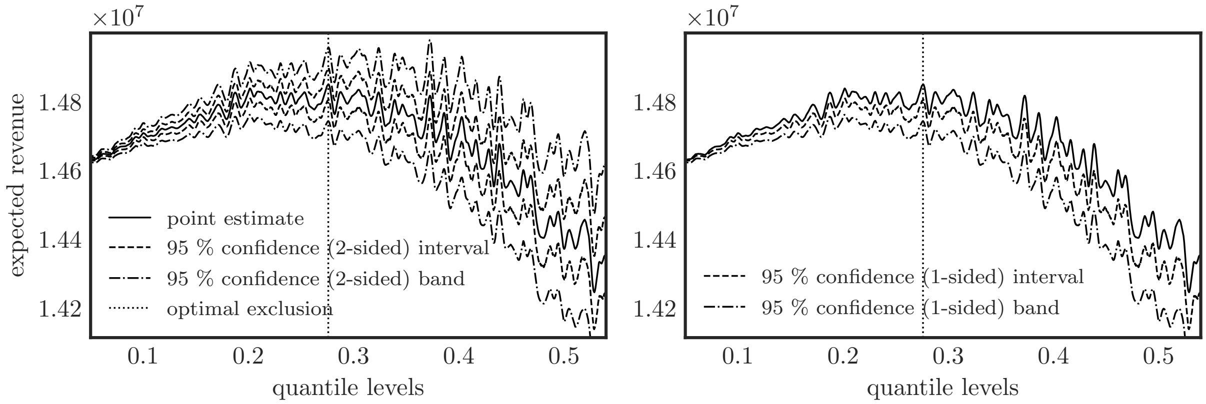


Figure 2: Confidence intervals and bands for the counterfactual expected revenue.

| Number of bidders | 2 | 3 | 2-5 | 5-9 | 2-9 |
|---|---|---|---|---|---|
| sample size | 10328 | 12477 | 43387 | 26841 | 60758 |
| bandwidth | 0.01 | 0.009 | 0.006 | 0.007 | 0.006 |
| optimal exclusion $\tilde{u}$ | 0.274 | 0.305 | 0.293 | 0.311 | 0.276 |
| $H_0$ | reject | reject | reject | reject | reject |

Table 3: Test results at the 95% confidence level.

would increase the seller's expected revenue. Formally, we test the hypotheses $H_0$ against $H_1$, where

$$H_0 : \sup_{u^* \in [h, 1-h]} \Delta(u^*) = 0, \quad H_1 : \sup_{u^* \in [h, 1-h]} \Delta(u^*) > 0. \tag{75}$$

The corresponding test statistic is the maximal (over the grid) value of the lower end point function of the one-sided confidence band, and $H_0$ is rejected whenever this maximum is positive. We denote by $\tilde{u}$ the point at which the maximum is attained, i.e. the optimal exclusion level.

We test the hypothesis using subsamples of auctions with different numbers of bidders, see Table 3. We use subsamples with 2 and 3 bidders, and also with 2-5 (small auctions), 5-9 (large auctions), and 2-9 (all auctions) bidders.[15] Under all the specifications, $H_0$ is rejected at 95% confidence level, see Table 3 and Figure 2, meaning that the revenue gains at the optimal reserve price are statistically significant, albeit relatively small.

[15]Typically, a researcher would pick, for the sake of simplicity, a subsample of auctions with the same number of bidders. However, our methodology allows for a random number of bidders, so we can pool auctions with different numbers of bidders together.

# 7 Practical considerations

In this section, we briefly discuss some important technical aspects of our methodology.

**Choice of the grid.** While it is theoretically possible to evaluate our estimators at any quantile level, choosing the evenly-spaced grid $\{i/n\}_{i=2}^n$ has a massive impact on the computational complexity of the estimation procedure and its performance.

Note that, with this grid, the estimate of $\hat{q}_h(u)$ becomes a discrete convolution of the vector of spacings $\{b_{(i)} - b_{(i-1)}\}_{i=2}^n$ with a discrete filter corresponding to $K_h$. The discrete convolution is a remarkably fast and reliable procedure. Moreover, the counterfactuals estimators can be well approximated with the weighted cumulative sums of the vectors of spacings. Consequently, all our estimators can be thought of as combinations of elementary vector operations with sorting and convolution.

**Shape restrictions.** Since $v(\cdot)$ is a quantile function, one may want to impose monotonicity on $\hat{v}_h(\cdot)$ and the associated confidence bands. As suggested in Chernozhukov et al. (2009, 2010), an effective way of doing so is *smooth rearrangement* of the estimate and the confidence bands, whose discrete counterpart is merely a sorting algorithm. We leave the analysis of such shape-restricted estimators for future work. We note that there is emerging literature exploiting shape restrictions for auction counterfactuals, e.g. Henderson et al. (2012); Luo and Wan (2018); Pinkse and Schurter (2019); Ma et al. (2021).

**Competing estimators.** The main competitor to our estimator $\hat{q}_h(\cdot)$ of the bid quantile density is the reciprocal of the kernel estimator $\hat{f}_l(\cdot)$ of the bid density, as in the first step of the procedure of Guerre et al. (2000),

$$\tilde{q}_l\,(u) = \left(\hat{f}_l(b_{([un]+1)})\right)^{-1} = \left(\frac{1}{nl}\sum_{i=1}^{n} K\left(\frac{b_{([un]+1)} - b_i}{l}\right)\right)^{-1}, \quad u \in [0,1].$$

An insightful comparison of $\hat{q}_h$ and $\tilde{q}_l$ was carried out by Jones (1992) who showed that the variance components of the mean squared errors (MSE) of the estimators are equal if so-called *scale match-up* bandwidths are used. Namely, for a fixed point $u \in (0,1)$, the MSE of $\hat{q}(u)$ is only less than or equal to that of $\tilde{q}_l(u)$ when $q(u)q''(u) \leqslant 1.5q'(u)^2$ or, equivalently, $f(b)f''(b) \geqslant 1.5f'(b)^2$. Therefore, the reciprocal kernel density $\tilde{q}_l$ performs better close to the center of the distribution, while the kernel quantile density $\hat{q}_h$ is preferable at the tails.

Finally, we note that the algorithms to construct the estimates of $\tilde{q}_l$ and $\hat{q}_h$ on the grid, seem to have different computational complexity, which can be heuristically shown to be

$O(n^2)$ and $O(n \log n)$, respectively. This is due to the fact that the convolution algorithm, which the estimator $\hat{q}_l(u)$ relies on, has the complexity of roughly $O(n \log n)$ due to its usage of the fast Fourier transform.

# 8 Conclusion

In this paper, we develop a novel approach to estimation and inference on counterfactual functionals of interest (such as the expected revenue as a function of the rank of the counterfactual reserve price) in the standard nonparametric model of the first-price sealed-bid auction. We show that these counterfactuals can be written as continuous linear functionals of the quantile function of bidders' valuations, which can be recovered from the observed bids using a well-known explicit formula. We suggest natural estimators of the counterfactuals and show that their asymptotic behavior depends on their structure. In particular, we classify the counterfactuals into two types, one allowing for parametric (fast) convergence rates and standard inference, and the other exhibiting nonparametric (slow) convergence rates and lack of uniform convergence. For each of the types of counterfactuals, we develop simple, simulation-based algorithms for constructing pointwise confidence intervals and uniform confidence bands. We apply our results to assess the potential for revenue extraction by setting an optimal reserve price, using Phil Haile's USFS auction data.

Avenues for further research include incorporating auction heterogeneity, showing the minimax optimality of the counterfactual estimators, developing procedures for data-driven bandwidth selection, shape-restricted estimation of the valuation quantile function and associated functionals, and estimation and inference for the value density estimator based on our value quantile function estimator.

# Acknowledgements

We are grateful to Karun Adusumilli, Tim Armstrong, John Asker, Yuehao Bai, Zheng Fang, Sergio Firpo, Antonio Galvao, Andreas Hagemann, Vitalijs Jascisens, Hiroaki Kaido, Nail Kashaev, Michael Leung, Tong Li, Hyungsik Roger Moon, Hashem Pesaran, Guillaume Pouliot, Geert Ridder, Pedro Sant'Anna, Yuya Sasaki, Matthew Shum, Liyang Sun, Takuya Ura, Quang Vuong, Kaspar Wuthrich, and seminar participants at USC for valuable comments. All errors and omissions are our own.

# Appendix

## A Estimation and inference for value quantiles, proofs

### A.1 Proof of Theorem 1

First, we need the following two lemmas concerning expressions that appear further in the proof.

**Lemma 1.** *Suppose $K$ is a continuous function of bounded variation. Then*

$$\int_0^1 K_h(u-z)\,d\left(\hat{Q}(z)-Q(z)\right) = -\int_0^1 \left(\hat{Q}(z)-Q(z)\right)\,dK_h(u-z) + R_n^I(u), \tag{76}$$

*where* $\sup_{u\in[0,1]} |R_n^I(u)| = O_{a.s.}\left(\frac{1}{nh}\right)$.

*Proof.* Denote $\hat{\psi}(z) = \hat{Q}(z) - Q(z)$ and note that $\hat{\psi}$ is a function of bounded variation a.s. Using integration by parts for the Riemann-Stieltjes integral (see e.g. Stroock, 1998, Theorem 1.2.7), we have

$$\int_0^1 K_h(u-z)\,d\hat{\psi}(z) = -\int_0^1 \hat{\psi}(z)\,dK_h(u-z) + K_h(u-1)\hat{\psi}(1) - K_h(u)\hat{\psi}(0) \tag{77}$$

To complete the proof, note that $\hat{\psi}(1) = b_{(n)} - \bar{b} = O_{a.s.}(n^{-1})$, $\hat{\psi}(0) = b_{(1)} - \underline{b} = O_{a.s.}(n^{-1})$, $|K_h(u-1)| \leqslant h^{-1}K(0)$ and $|K_h(u)| \leqslant h^{-1}K(0)$. □

**Lemma 2.** *Suppose $K$ is a continuous function of bounded variation. Then, for every $u \in [0,1]$,*

$$Z_n^*(u) := \sqrt{nh}\int_0^1 (\hat{F}(Q(z)) - z)\,dK_h(u-z) = -\mathbb{G}_{n,h}(u), \tag{78}$$

$$\mathbb{G}_{n,h}(u) := \frac{\sqrt{nh}}{n}\sum_{i=1}^n \left[K_h(u-F(b_i)) - \mathbb{E}K_h(u-F(b_i))\right]. \tag{79}$$

*Proof.* Using integration by parts for the Riemann-Stieltjes integral (see e.g. Stroock, 1998, Theorem 1.2.7), we have

$$\begin{aligned}\int_0^1 (\hat{F}(Q(z)) - z)\,dK_h(u-z) &= -\int_0^1 K_h(u-z)\,d\left[\hat{F}(Q(z)) - z\right] + K_h(u-1)\left[\hat{F}(\bar{b}) - 1\right] + K_h(u)\hat{F}(0)\\ &= -\int_0^1 K_h(u-z)\,d\left[\hat{F}(Q(z)) - z\right],\end{aligned}$$

where we used the fact that $\hat{F}(\bar{b}) = 1$ a.s. and $\hat{F}(0) = 0$ a.s. We further write

$$\begin{aligned}\int_0^1 (\hat{F}(Q(z)) - z)\, dK_h(u-z) &= -\int_0^1 K_h(u-z)\, d\left[\hat{F}(Q(z)) - z\right] \\ &= -\int_0^{\bar{b}} K_h(u - F(x))\, d\left[\hat{F}(x) - F(x)\right] \\ &= -\frac{1}{n}\sum_{i=1}^n \left[K_h(u - F(b_i)) - \mathbb{E}K_h(u - F(b_i))\right],\end{aligned}$$

where in the second equality we used the change of variables $x = Q(z)$. □

We now proceed with the proof of Theorem 1.

Plug in the BK expansion (23) and use Lemma 1 to obtain

$$\hat{q}_h(u) - q_h(u) = \int_0^1 K_h(u-z)\, d\left[\hat{Q}(z) - Q(z)\right] \tag{80}$$

$$= \int_0^1 \left[\hat{Q}(z) - Q(z)\right]\, dK_h(u-z) + R_n^I(u) \tag{81}$$

$$= \int_0^1 q(z)(\hat{F}(Q(z)) - z)\, dK_h(u-z) + \int_0^1 R_n^{BK}(z)\, dK_h(u-z) + R_n^I(u). \tag{82}$$

**First term in** (82).

Since $f$ is bounded away from zero, $|q'| \leqslant M < \infty$ for some constant $M$, and hence $|q(z) - q(u)| \leqslant M|z-u|$. The first term in (82) can then be rewritten as

$$\int_0^1 q(z)(\hat{F}(Q(z)) - z)\, dK_h(u-z) = q(u)\int_0^1 (\hat{F}(Q(z)) - z)\, dK_h(u-z) + R_n^{II}(u), \tag{83}$$

where

$$\left|R_n^{II}(u)\right| = \left|\int_0^1 (q(z) - q(u))(\hat{F}(Q(z)) - z)\, dK_h(u-z)\right| \tag{84}$$

$$\leqslant Mh\left|\int_0^1 (\hat{F}(Q(z)) - z)\, dK_h(u-z)\right| = Mh\left|(nh)^{-1/2}Z_n^*(u)\right|. \tag{85}$$

By Lemma 2, $Z_n^*(u) = -\mathbb{G}_{n,h}(u)$, where the process $\mathbb{G}_{n,h}(u) = O_{a.s.}(\log h)$ uniformly in $u \in [0,1]$ (see e.g. Silverman, 1978; Stute, 1984), and hence

$$R_n^{II}(u) = O_{a.s.}\left(\frac{h\log h}{\sqrt{nh}}\right) \text{ uniformly over } u \in (0,1). \tag{86}$$

Applying Lemma 2 to the first term in (83) allows us to rewrite

$$\int_0^1 q(z)(\hat{F}(Q(z)) - z)\, dK_h(u-z) = -q(u)(nh)^{-1/2}\mathbb{G}_{n,h}(u) + O_{a.s.}\left(\frac{h \log h}{\sqrt{nh}}\right). \tag{87}$$

**Second term in** (82).

This term can be upper bounded as follows,

$$\sup_u \left| \int_0^1 R_n^{BK}(z)\, dK_h(u-z) \right| \leqslant \sup_u \int_0^1 \left| R_n^{BK}(z) \right| \left| dK_h(u-z) \right| \tag{88}$$

$$\leqslant \sup_z |R_n^{BK}(z)| TV(K_h) = O_{a.s.}\left(n^{-3/4}\ell(n)\right) h^{-1} TV(K) = O_{a.s.}\left(h^{-1} n^{-3/4} l(n)\right), \tag{89}$$

where we used the properties of total variation in the first inequality and in the second equality.

Plugging (87) and (89) into (82) and multiplying by $\sqrt{nh}$ yields

$$\sqrt{nh}\left(\hat{q}_h(u) - q_h(u)\right) = -q(u)\mathbb{G}_{n,h}(u) + O_{a.s.}\left(h \log h + h^{-1/2} n^{-1/4} \ell(n)\right), \tag{90}$$

where we disregarded the term $\sqrt{nh} R_n^I(u)$, since it has the uniform order $O_{a.s.}(n^{-1/2}h^{-1/2})$, which is smaller than $O_{a.s.}\left(h^{-1/2} n^{-1/4} \ell(n)\right)$.

Note that, for $u \in [h, 1-h]$, there exists $\zeta(u, z)$ lying between $u$ and $z$ such that

$$q_h(u) = \int_0^1 q(z) K_h(u-z)\, dz = \int_0^1 \left(q(u) + q'(\zeta(u,z))(u-z)\right) K_h(u-z)\, dz \tag{91}$$

$$= q(u) + O(h). \tag{92}$$

Combining this with (90) yields

$$\sqrt{nh}\left(\hat{q}_h(u) - q(u)\right) = -q(u)\mathbb{G}_{n,h}(u) + O_{a.s.}\left(n^{1/2}h^{3/2} + h \log h + h^{-1/2} n^{-1/4} \ell(n)\right), \tag{93}$$

uniformly in $u \in [h, 1-h]$. Using $\mathbb{G}_{n,h}(u) = O_{a.s.}(\log h)$ again, we conclude that

$$\sqrt{nh}\left(\hat{q}_h(u) - q(u)\right) = O_{a.s.}\left(\log h + n^{1/2} h^{3/2}\right) \tag{94}$$

(note that we dropped the terms $h \log h$ and $h^{-1/2} n^{-1/4} \ell(n)$ since they are smaller than $\log h$) or, dividing by $\sqrt{nh}$,

$$\hat{q}_h(u) - q(u) = O_{a.s.}\left(\frac{\log h}{\sqrt{nh}} + h\right), \text{ uniformly in } u \in [h, 1-h]. \tag{95}$$

Now we replace $q(u)\mathbb{G}_{n,h}(u)$ by $\hat{q}_h(u)\mathbb{G}_{n,h}(u)$ in (93), which leads to the approximation error

$$\mathbb{G}_{n,h}(u)\left(\hat{q}_h(u) - q(u)\right) = O_{a.s}\left(\frac{(\log h)^2}{\sqrt{nh}} + h\log h\right), \text{ uniformly in } u \in [h, 1-h], \tag{96}$$

as follows from (95). Hence, (93) becomes

$$\sqrt{nh}\left(\hat{q}_h(u) - q(u)\right) = -\hat{q}_h(u)\mathbb{G}_{n,h}(u) + O_{a.s.}\left(n^{1/2}h^{3/2} + h\log h + h^{-1/2}n^{-1/4}\ell(n)\right), \tag{97}$$

uniformly in $u \in [h, 1-h]$, where we dropped the term $\frac{(\log h)^2}{\sqrt{nh}}$ since it is smaller than $h^{-1/2}n^{-1/4}\ell(n)$.

Finally, write

$$\sqrt{nh}\left(\hat{v}_h(u) - v(u)\right) = \sqrt{nh}\left(\hat{Q}(u) - Q(u)\right) + \check{A}(u)\sqrt{nh}\left(\hat{q}_h(u) - q(u)\right). \tag{98}$$

Since $\hat{Q}(u) - Q(u) = O_{a.s.}(n^{-1/2})$ uniformly in $u \in (0,1)$, we have

$$\sqrt{nh}\left(\hat{v}_h(u) - v(u)\right) = O_{a.s.}(h^{1/2}) + \check{A}(u)\sqrt{nh}\left(\hat{q}_h(u) - q(u)\right). \tag{99}$$

Combining this with (97) and noting that $h\log h$ is smaller than $h^{1/2}$ yields

$$\sqrt{nh}\left(\hat{v}_h(u) - v(u)\right) = -\check{A}(u)\hat{q}_h(u)\mathbb{G}_{n,h}(u) + O_{a.s.}\left(n^{1/2}h^{3/2} + h^{1/2} + h^{-1/2}n^{-1/4}\ell(n)\right), \tag{100}$$

uniformly in $u \in [h, 1-h]$. Dividing by $\hat{q}_h(u)$, which is bounded away from zero w.p.a. 1, completes the proof. □

## A.2 Proof of Theorem 2

A key ingredient of the proof is to note that Lemmas 2.3 and 2.4 of Chernozhukov et al. (2014) continue to hold even if their random variable $Z_n$ does not have the form $Z_n = \sup_{f\in\mathcal{F}_n}\mathbb{G}_n f$ for the standard empirical process $\mathbb{G}_n$, but instead is a generic random variable admitting a strong sup-Gaussian approximation with a sufficiently small remainder.

For completeness, we provide the aforementioned trivial extensions of the two lemmas here.

Let $X$ be a random variable with distribution $P$ taking values in a measurable space $(S, \mathcal{S})$. Let $\mathcal{F}$ be a class of real-valued functions on $S$. We say that a function $F : S \to \mathbb{R}$ is an *envelope* of $\mathcal{F}$ if $F$ is measurable and $|f(x)| \leqslant F(x)$ for all $f \in \mathcal{F}$ and $x \in S$.

We impose the following assumptions (A1)-(A3) of Chernozhukov et al. (2014).

(A1) The class $\mathcal{F}$ is *pointwise measurable*, i.e. it contains a coutable subset $\mathcal{G}$ such that for every $f \in \mathcal{F}$ there exists a sequence $g_m \in \mathcal{G}$ with $g_m(x) \to f(x)$ for every $x \in S$.

(A2) For some $q \geqslant 2$, an envelope $F$ of $\mathcal{F}$ satisfies $F \in L^q(P)$.

(A3) The class $\mathcal{F}$ is $P$-pre-Gaussian, i.e. there exists a tight Gaussian random variable $G_P$ in $l^\infty(\mathcal{F})$ with mean zero and covariance function

$$\mathbb{E}[G_P(f)G_P(g)] = \mathbb{E}[f(X)g(X)] \text{ for all } f, g \in \mathcal{F}.$$

**Lemma 3** (A trivial extension of Lemma 2.3 of Chernozhukov et al. (2014)). *Suppose that Assumptions (A1)-(A3) are satisfied and that there exist constants $\underline{\sigma}$, $\bar{\sigma} > 0$ such that $\underline{\sigma}^2 \leqslant Pf^2 \leqslant \bar{\sigma}^2$ for all $f \in \mathcal{F}$. Moreover, suppose there exist constants $r_1, r_2 > 0$ and a random variable $\tilde{Z} = \sup_{f\in\mathcal{F}} G_P f$ such that $\mathbb{P}(|Z - \tilde{Z}| > r_1) \leqslant r_2$. Then*

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(Z \leqslant t) - \mathbb{P}(\tilde{Z} \leqslant t)\right| \leqslant C_\sigma r_1 \left\{\mathbb{E}\tilde{Z} + \sqrt{1 \vee \log(\underline{\sigma}/r_1)}\right\} + r_2,$$

*where $C_\sigma$ is a constant depending only on $\underline{\sigma}$ and $\bar{\sigma}$.*

*Proof.* For every $t \in \mathbb{R}$, we have

$$\begin{aligned}\mathbb{P}(Z \leqslant t) &= \mathbb{P}(\{Z \leqslant t\} \cap \{|Z - \tilde{Z}| \leqslant r_1\}) + \mathbb{P}(\{Z \leqslant t\} \cap \{|Z - \tilde{Z}| > r_1\})\\ &\leqslant \mathbb{P}(\tilde{Z} \leqslant t + r_1) + r_2\\ &\leqslant \mathbb{P}(\tilde{Z} \leqslant t) + C_\sigma r_1 \left\{\mathbb{E}\tilde{Z} + \sqrt{1 \vee \log(\underline{\sigma}/r_1)}\right\} + r_2,\end{aligned}$$

where Lemma A.1 of Chernozhukov et al. (2014) (an anti-concentration inequality for $\tilde{Z}$) is used to deduce the last inequality. A similar argument leads to the reverse inequality, which completes the proof. □

**Lemma 4** (A trivial extension of Lemma 2.4 of Chernozhukov et al. (2014)). *Suppose that there exists a sequence of $P$-centered classes $\mathcal{F}_n$ of measurable functions $S \to \mathbb{R}$ satisfying assumptions (A1)-(A3) with $\mathcal{F} = \mathcal{F}_n$ for each $n$, where in the assumption (A3) the constants $\underline{\sigma}$ and $\bar{\sigma}$ do not depend on $n$. Denote by $B_n$ the Brownian bridge on $\ell^\infty(\mathcal{F}_n)$, i.e. a tight Gaussian random variable in $\ell^\infty(\mathcal{F}_n)$ with mean zero and covariance function*

$$\mathbb{E}[B_n(f)B_n(g)] = \mathbb{E}[f(X)g(X)] \textit{ for all } f, g \in \mathcal{F}_n.$$

*Moreover, suppose that there exists a sequence of random variables $\tilde{Z}_n = \sup_{f\in\mathcal{F}_n} B_n(f)$ and a sequence of constants $r_n \to 0$ such that $|Z_n - \tilde{Z}_n| = O_P(r_n)$ and $r_n\mathbb{E}\tilde{Z}_n \to 0$. Then*

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(Z_n \leqslant t) - \mathbb{P}(\tilde{Z}_n \leqslant t)\right| \to 0.$$

*Proof.* Take $\beta_n \to \infty$ sufficiently slowly such that $\beta_n r_n(1 \vee \mathbb{E}\tilde{Z}_n) = o(1)$. Then since $\mathbb{P}(|Z_n - \tilde{Z}_n| > \beta_n r_n) = o(1)$, by Lemma 3, we have

$$\sup_{t\in\mathbb{R}} \left|\mathbb{P}(Z_n \leqslant t) - \mathbb{P}(\tilde{Z}_n \leqslant t)\right| = O\left(r_n(\mathbb{E}\tilde{Z}_n + |\log(\beta_n r_n)|)\right) + o(1) = o(1).$$

This completes the proof. □

**Lemma 5.** *Let*

$$W_n^* = \sup_{u\in[0,1]} \xi(u)\sqrt{nh} \cdot \frac{1}{n}\sum_{i=1}^{n} [K_h(U_i - u) - \mathbb{E}K_h(U_i - u)] \tag{101}$$

*for some smooth function $\xi : [0,1] \to \mathbb{R}$. Then there exists a tight centered Gaussian random variable $B_n$ in $\ell^\infty([0,1])$ with the covariance function*

$$\mathbb{E}[B_n(u)B_n(v)] = \xi(u)\xi(v) \cdot Cov\left(K_h(U-u), K_h(U-v)\right), \quad u, v \in [0,1], \tag{102}$$

*such that, for $\tilde{W}_n = \sup_{u\in[0,1]} B_n(u)$, we have the approximation*

$$W_n^* = \tilde{W}_n + O_p\left((nh)^{-1/6}\log n\right). \tag{103}$$

*Proof.* Define the class of functions

$$\mathcal{F}_n = \{[0,1] \ni x \mapsto \xi(u)K_h(u-x), \quad u \in [0,1]\} \tag{104}$$

and note that

$$W_n^* = \sqrt{h}\|\mathbb{G}_n\|_{\mathcal{F}_n}. \tag{105}$$

Let us apply Chernozhukov et al. (2014, Proposition 3.1) to obtain a sup-Gaussian approximation of $W_n^*$. Indeed, in the notation of Chernozhukov et al. (2014, Section 3.1), take

$$g \equiv 1, \quad \mathcal{G} = \{g\}, \quad \mathcal{I} = [0,1], \quad c_n(u,g) = \xi(u). \tag{106}$$

Then the representation (8) in Chernozhukov et al. (2014) holds, i.e.

$$W_n^* = \sup_{(u,g)\in\mathcal{I}\times\mathcal{G}} c_n(u,g)\sqrt{nh}\cdot\frac{1}{n}\sum_{i=1}^{n}\left[K_h(U_i-u)-\mathbb{E}K_h(U_i-u)\right]. \tag{107}$$

It is now trivial to check that the assumptions of Chernozhukov et al. (2014, Proposition 3.1) hold and the statement of the lemma follows. □

Let us now go back to the proof of Theorem 2. Use Lemma 5 with $\xi(u)=u$ and note that Lemma 4 and Chernozhukov et al. (2014, Remark 3.2) then imply

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(W_n^*\leqslant t)-\mathbb{P}(\tilde{W}_n\leqslant t)\right|\to 0. \tag{108}$$

On the other hand, by Theorem 1 we have

$$W_n = W_n^* + O_{a.s.}\left(h^{1/2}+h^{-1/2}n^{-1/4}l(n)\right). \tag{109}$$

Substituting (103) into this equation, we obtain

$$W_n = \tilde{W}_n + O_p\left((nh)^{-1/6}\log n + h^{1/2}+h^{-1/2}n^{-1/4}l(n)\right). \tag{110}$$

Assumption 3 then implies $W_n-\tilde{W}_n = o_p(\log^{-1/2} n)$. Chernozhukov et al. (2014, Remark 3.2) now implies

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(W_n\leqslant t)-\mathbb{P}(\tilde{W}_n\leqslant t)\right|\to 0. \tag{111}$$

Given (108) and (111), applying the triangle inequality finishes the proof. □

# B Estimation and inference for counterfactuals, proofs

## B.1 Proof of Theorem 3

First, write

$$\hat{S}(u^*)-S(u^*) = \int_{u^*}^{1}\varphi(u)(\hat{Q}(u)-Q(u))\,du \tag{112}$$

$$-\check{A}(u^*)\psi(u^*)(\hat{Q}(u^*)-Q(u^*))+\check{A}(1)\psi(1)(\hat{Q}(1)-Q(1)). \tag{113}$$

Using the classical BK expansion (23), we obtain

$$\hat{S}(u^*) - S(u^*) = -\int_{u^*}^{1} \varphi(u)q(u)\left[\hat{F}(Q(u)) - u\right] du \tag{114}$$

$$+ \check{A}(u^*)\psi(u^*)q(u^*)\left[\hat{F}(Q(u^*)) - u^*\right] + R_n(u^*), \tag{115}$$

where the composite error term

$$R_n(u^*) = \int_{u^*}^{1} \varphi(u)r_n(u)\,du - \check{A}(u^*)\psi(u^*)r_n(u^*) + \check{A}(1)\psi(1)(\hat{Q}(1) - Q(1)) \tag{116}$$

$$= O_{a.s.}\left(n^{-3/4}\ell(n)\right), \tag{117}$$

uniformly in $u^* \in [0,1]$. The latter rate follows from (23) and the fact that $\hat{Q}(1) - Q(1) = O_{a.s.}(n^{-1})$.

Denoting $U_i = F(b_i)$, we can write

$$\sqrt{n}\left(\hat{S}(u^*) - S(u^*)\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[f_{u^*}(U_i) - \mathbb{E}f_{u^*}(U_i)\right] + O_{a.s.}\left(n^{-1/4}\ell(n)\right). \tag{118}$$

Let us show that the class $\{f_{u^*} \mid u^* \in [0,1]\}$ is Donsker.

Since the sum of a finite number of Donsker classes is Donsker (see Alexander, 1987), and also a constant times a Donsker class is Donsker, it suffices to show that the following (uniformly bounded) classes are Donsker,

$$\mathcal{H} = \{h_{u^*} : U \mapsto \int_{u^*}^{1} \varphi(u)q(u)1(U \leqslant u)\,du \mid u^* \in [0,1]\}, \tag{119}$$

$$\mathcal{G} = \{g_{u^*} : U \mapsto \check{A}(u^*)\psi(u^*)q(u^*)1(U \leqslant u^*) \mid u^* \in [0,1]\}. \tag{120}$$

Indeed, for $u^*, v^* \in [0,1]$,

$$|h_{u^*}(U) - h_{v^*}(U)| = \left|\int_{\min(u^*,v^*)}^{\max(u^*,v^*)} \varphi(x)q(x)1(U \leqslant x)\,dx\right| \leqslant |u^* - v^*| \sup_{u\in[0,1]} |\varphi(u)q(u)|, \tag{121}$$

i.e. $\mathcal{H}$ is Lipschitz in parameter. Its Donskerness follows by Theorem 2.7.11 and Theorem 2.5.6 in Vaart and Wellner (1996).

On the other hand, $\mathcal{G}$ is Donsker since it is a product of the set of constant functions $U \mapsto \check{A}(u^*)\psi(u^*)q(u^*)$ (which is trivially Donsker) and the VC class $\{1(\cdot \leqslant u^*),\ u^* \in [0,1]\}$. $\square$

## B.2 Proof of Theorem 4

We have

$$\sqrt{nh}\left(\hat{T}_h(u^*) - T(u^*)\right) = \check{\varphi}(u^*)\left(\hat{v}_h(u^*) - v(u^*)\right) + \sqrt{nh}\left(\hat{S}_{\check{\psi}}(u^*) - S_\psi(u^*)\right) \tag{122}$$

$$= \check{\varphi}(u^*)\check{A}(u^*)\hat{q}_h(u^*)\left(-\mathbb{G}_{n,h}(u^*) + R_n(u)\right) + O_p\left(h^{1/2}\right) \tag{123}$$

$$= -\check{\varphi}(u^*)\check{A}(u^*)\hat{q}_h(u^*)\mathbb{G}_{n,h}(u^*) + O_p\left(n^{1/2}h^{3/2} + h^{1/2} + h^{-1/2}n^{-1/4}\ell(n)\right), \tag{124}$$

uniformly in $u^* \in [h, 1-h]$, where the last two equations use Theorem 1 and Theorem 3. Dividing by $\hat{q}_h(u^*)$, which is bounded away from zero w.p.a. 1, finishes the proof. □

## B.3 Proof of Theorem 5

Use Lemma 5 with $\xi(u) = u\varphi(u)$ and note that Lemma 4 and Remark 3.2 in Chernozhukov et al. (2014) then imply

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(W_n^{T*} \leqslant t) - \mathbb{P}(\tilde{W}_n^T \leqslant t)\right| \to 0. \tag{125}$$

On the other hand, by Theorem 4 we have

$$W_n = W_n^* + O_{a.s.}\left(h^{1/2} + h^{-1/2}n^{-1/4}l(n)\right). \tag{126}$$

Substituting (103) into this equation, we obtain

$$W_n^T = \tilde{W}_n^T + O_p\left((nh)^{-1/6}\log n + h^{1/2} + h^{-1/2}n^{-1/4}l(n)\right). \tag{127}$$

Under the assumption of the theorem, $h$ decays polynomially, and hence $W_n^T - \tilde{W}_n^T = o_p(\log^{-1/2} n)$. Remark 3.2 of Chernozhukov et al. (2014) now implies

$$\sup_{t\in\mathbb{R}}\left|\mathbb{P}(W_n^T \leqslant t) - \mathbb{P}(\tilde{W}_n^T \leqslant t)\right| \to 0. \tag{128}$$

Given (125) and (128), applying the triangle inequality finishes the proof. □

# References

Alexander, K. S. (1987): "The central limit theorem for empirical processes on Vapnik-Cervonenkis classes," *The Annals of Probability*, 178–203.

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

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

Bahadur, R. R. (1966): "A note on quantiles in large samples," *The Annals of Mathematical Statistics*, 37, 577–580.

Bickel, P. J. and M. Rosenblatt (1973): "On some global measures of the deviations of density function estimates," *The Annals of Statistics*, 1071–1095.

Bloch, D. A. and J. L. Gastwirth (1968): "On a simple estimate of the reciprocal of the density function," *The Annals of Mathematical Statistics*, 39, 1083–1085.

Chernozhukov, V., D. Chetverikov, and K. Kato (2014): "Gaussian approximation of suprema of empirical processes," *The Annals of Statistics*, 42, 1564–1597.

Chernozhukov, V., I. Fernandez-Val, and A. Galichon (2009): "Improving point and interval estimators of monotone functions by rearrangement," *Biometrika*, 96, 559–575.

Chernozhukov, V., I. Fernández-Val, and A. Galichon (2010): "Quantile and Probability Curves Without Crossing," *Econometrica*, 78, 1093–1125.

Csörgő, M., L. Horváth, and P. Deheuvels (1991): "Estimating the quantile-density function," in *Nonparametric Functional Estimation and Related Topics*, Springer, 213–223.

Elyakime, B., J. J. Laffont, P. Loisel, and Q. Vuong (1994): "First-price sealed-bid auctions with secret reservation prices," *Annales d'Economie et de Statistique*, 115–141.

Enache, A. and J.-P. Florens (2017): "A quantile approach to the estimation of first-price private value auction," *Available at SSRN 3522067*.

Falk, M. (1986): "On the estimation of the quantile density function," *Statistics & Probability Letters*, 4, 69–73.

Gimenes, N. and E. Guerre (2021): "Quantile regression methods for first-price auctions," *Journal of Econometrics*.

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

——— (2009): "Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions," *Econometrica*, 77, 1193–1227.

Guerre, E. and C. Sabbah (2012): "Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function," *Econometric Theory*, 87–129.

Haile, P., H. Hong, and M. Shum (2003): "Nonparametric tests for common values at first-price sealed-bid auctions," .

Haile, P. A. (2001): "Auctions with resale markets: An application to US forest service timber sales," *American Economic Review*, 91, 399–427.

Hall, P. (1991): "On convergence rates of suprema," *Probability Theory and Related Fields*, 89, 447–455.

——— (2013): *The bootstrap and Edgeworth expansion*, Springer Science & Business Media.

Henderson, D. J., J. A. List, D. L. Millimet, C. F. Parmeter, and M. K. Price (2012): "Empirical implementation of nonparametric first-price auction models," *Journal of Econometrics*, 168, 17–28.

Horowitz, J. L. (2001): "The bootstrap," in *Handbook of econometrics*, Elsevier, vol. 5, 3159–3228.

Ingraham, A. T. (2005): "A test for collusion between a bidder and an auctioneer in sealed-bid auctions," *Available at SSRN 712881*.

Jones, M. C. (1992): "Estimating densities, quantiles, quantile densities and density quantiles," *Annals of the Institute of Statistical Mathematics*, 44, 721–727.

Kiefer, J. (1967): "On Bahadur's representation of sample quantiles," *The Annals of Mathematical Statistics*, 38, 1323–1342.

Krishna, V. (2009): *Auction theory*, Academic press.

Li, T. and I. Perrigne (2003): "Timber sale auctions with random reserve prices," *Review of Economics and Statistics*, 85, 189–200.

Li, T., I. Perrigne, and Q. Vuong (2000): "Conditionally independent private information in OCS wildcat auctions," *Journal of Econometrics*, 98, 129–161.

Loertscher, S. and L. M. Marx (2020): "Asymptotically optimal prior-free clock auctions," *Journal of Economic Theory*, 187, 105030.

Luo, Y. and Y. Wan (2018): "Integrated-quantile-based estimation for first-price auction models," *Journal of Business & Economic Statistics*, 36, 173–180.

Ma, J., V. Marmer, and A. Shneyerov (2019): "Inference for first-price auctions with Guerre, Perrigne, and Vuong's estimator," *Journal of Econometrics*, 211, 507–538.

Ma, J., V. Marmer, A. Shneyerov, and P. Xu (2021): "Monotonicity-constrained nonparametric estimation and inference for first-price auctions," *Econometric Reviews*, 40, 944–982.

Marmer, V. and A. Shneyerov (2012): "Quantile-based nonparametric inference for first-price auctions," *Journal of Econometrics*, 167, 345–357.

Marra, M. (2020): "Sample Spacings for Identification: The Case of English Auctions With Absentee Bidding," *Available at SSRN 3622047.*

Matzkin, R. L. (2013): "Nonparametric identification in structural economic models," *Annual Review of Economics*, 5, 457–486.

Montiel Olea, J. L. and M. Plagborg-Møller (2019): "Simultaneous confidence bands: Theory, implementation, and an application to SVARs," *Journal of Applied Econometrics*, 34, 1–17.

Paarsch, H. J., H. Hong, et al. (2006): "An introduction to the structural econometrics of auction data," *MIT Press Books*, 1.

Paul, A. and G. Gutierrez (2004): "Mean sample spacings, sample size and variability in an auction-theoretic framework," *Operations Research Letters*, 32, 103–108.

Perrigne, I. and Q. Vuong (2019): "Econometrics of Auctions and Nonlinear Pricing," *Annual Review of Economics*, 11, 27–54.

Pinkse, J. and K. Schurter (2019): "Estimation of auction models with shape restrictions," *arXiv preprint arXiv:1912.07466.*

Riley, J. G. and W. F. Samuelson (1981): "Optimal Auctions," *The American Economic Review*, 71, 381–392.

Rio, E. (1994): "Local invariance principles and their application to density estimation," *Probability Theory and Related Fields*, 98, 21–45.

Siddiqui, M. M. (1960): "Distribution of quantiles in samples from a bivariate population," *Journal of Research of the National Bureau of Standards*, 64, 145–150.

Silverman, B. W. (1978): "Weak and strong uniform consistency of the kernel estimate of a density and its derivatives," *The Annals of Statistics*, 177–184.

Smirnov, N. V. (1950): "On the construction of confidence regions for the density of distribution of random variables," in *Doklady Akad. Nauk SSSR*, vol. 74, 189–191.

Stroock, D. W. (1998): *A concise introduction to the theory of integration*, Springer Science & Business Media.

Stute, W. (1984): "The oscillation behavior of empirical processes: The multivariate case," *The Annals of Probability*, 361–379.

Vaart, A. W. and J. A. Wellner (1996): *Weak convergence and empirical processes*, Springer.

Welsh, A. (1988): "Asymptotically efficient estimation of the sparsity function at a point," *Statistics & Probability Letters*, 6, 427–432.

Zincenko, F. (2021): "Estimation and Inference of Seller's Expected Revenue in First-Price Auctions," *Available at SSRN 3966545.*