Large sample properties for estimators based on the order statistics approach in auctions

Konrad Menzel
Department of Economics, New York University

Paolo Morganti
Department of Business Administration, Instituto Tecnologico Autonomo de Mexico

For symmetric auctions, there is a close relationship between distributions of order statistics of bidders' valuations and observable bids that is often used to estimate or bound the valuation distribution, optimal reserve price, and other quantities of interest nonparametrically. However, we show that the functional mapping from distributions of order statistics to their parent distribution is, in general, not Lipschitz continuous and, therefore, introduces an irregularity into the estimation problem. More specifically, we derive the optimal rate for nonparametric point estimation of, and bounds for, the private value distribution, which is typically substantially slower than the regular root-n rate. We propose trimming rules for the nonparametric estimator that achieve that rate and derive the asymptotic distribution for a regularized estimator. We then demonstrate that policy parameters that depend on the valuation distribution, including optimal reserve price and expected revenue, are irregularly identified when bidding data are incomplete. We also give rates for nonparametric estimation of descending bid auctions and strategic equivalents.

Keywords. Empirical auctions, order statistics, bounds, irregular identification, uniform consistency.

JEL classification. C13, C14, D44.

The order statistics approach has been very fruitful for deriving nonparametric identification results and bounds for auction models.[1] However, as we show in this paper, the central step of "inverting out" the distribution of bidders' valuations from the distribution of an order statistic introduces an irregularity into the estimation problem. We show that point estimation of, or construction of bounds for, the cumulative distribution function (c.d.f.) of valuations from order statistics is generally at a rate slower than

[1] See, for example, Athey and Haile (2007).

root-n, and that uniformly valid inference requires trimming or some other form of regularization.

In the literature on empirical auctions, the problem of inverting a distribution of order statistics arises in several settings. For one, we may not observe the complete bidding data: there are many relevant cases in which, by design, only a subset of the bids are observed, even to the auctioneer. Most importantly, in descending bid (Dutch) auctions, only the winning bid is observed, whereas in the popular ascending "button" auction format, the auction ends when the second-highest bidder drops out, so that the highest bid is not observed. Furthermore, in many cases and without regard to the auction format, the researcher may only have access to a data set in which only the transaction price and/or the winning bid is recorded.

On the other hand, the results of this paper are also relevant settings in which some of the conditions of the benchmark model for the auction are relaxed and the ordered set of bids does not represent order statistics for a distribution of valuations or bidding strategies, but only are used to construct bounds. Haile and Tamer (2003) analyzed ascending bid formats, where the highest bid recorded for a particular bidder over the course of a given auction need not necessarily correspond to the "idealized" bid described by the theoretical model at hand. Their bounds are calculated by inverting the distribution of each order statistic separately; see also Chernozhukov, Lee, and Rosen (forthcoming) for a treatment of the statistical problem of constructing this type of bounds in the regular case. Also, in a recent study, Aradillas-López, Gandhi, and Quint (2013) proposed a test for correlated private values that is based on estimators for the valuation distribution from transaction prices in auctions with different numbers of participants, and Armstrong (2011) derived bounds on policy parameters for first-price auctions that allow for unobserved heterogeneity.

As an example, consider the textbook version of a descending bid auction in which, starting from a very high initial value, the auctioneer announces a decreasing sequence of prices and awards the good to the first bidder to drop out of the auction at a transaction price equal to the current quote. In this format, only the winning bidder reveals information about her true valuation. Under the assumption of independent private values, an identification argument based on the relationship between the distribution of valuations and its order statistics leads to the intriguing conclusion that, in principle, observations of the highest bid alone are sufficient to recover the full private value distribution, regardless of the size of the auction. However, when it comes to estimation, our analysis shows that even very large samples will, in general, be quite uninformative about the lower tail of the distribution of valuations. More specifically, we show that even though the distribution of private values can be estimated nonparametrically at a root-n rate at any given point, the rate for estimating the distribution function (with respect to an appropriate functional norm) can be made arbitrarily slow by considering auctions with larger numbers of bidders.

While in this particular and very stylized example, the difficulty stems from estimating the lower tail of the private value distribution, we show, for example, that in set-identified problems, intersection bounds on the upper tail of the distribution constructed from the complete bidding data suffer from the same problems. The differ-

ence between pointwise and uniform convergence rates is not merely a theoretical concern, but we show that even for point-identified settings, economically relevant (finite-dimensional) policy parameters inherit irregular identification as functionals of the valuation distribution. More specifically, we analyze bounds on the rate of convergence for linear functionals, expected revenue, and optimal reserve price. In particular, it is shown that expected revenue and optimal reserve price are not estimable at a root-n rate in general, but the fastest possible rate may be significantly slower, depending on the number of bidders in the observed and the counterfactual auctions. In particular, trimming the problematic parts of the support without regard to sample size will, in general, not lead to consistent estimation of those functionals.

These difficulties are not restricted to estimators based on distributions of order statistics, but we show that for the leading cases of independent private values (IPV) second- and first-price auctions with incomplete bidding data, the derived rates cannot be improved on by any other nonparametric estimator. More specifically, we derive upper bounds on the convergence rates for point estimators and bounds for the private value c.d.f. $F_0(v)$ under different functional norms, and for policy relevant functionals of the distribution, such as projected expected revenue and optimal reserve price. In particular, for the IPV case, these bounds on rates imply that without imposing shape restrictions on the parent distribution, the slow rate of the estimator based on the order statistics approach is, in fact, also the optimal rate of nonparametric estimation in the sense of Stone (1980). However, in general, the best attainable rates are shown to depend on the degree of smoothness of $F_0(v)$, the number of bidders in the auction, and which particular bids are observed.

Much of the recent literature on nonparametric estimation of auctions has focused on identification (for a relatively recent survey, see Athey and Haile (2007)), where Athey and Haile (2002) and Komarova (2009) provided results on nonparametric identification from incomplete bidding data, and Haile and Tamer (2003) proposed a method of constructing nonparametric bounds on the distribution under weaker assumptions on bidding behavior by inverting the distribution of each bid separately. Guerre, Perrigne, and Vuong (2000) derived rate-optimal nonparametric estimators for first-price auctions when all bids are observed, a case for which the problem of inverting distributions of order statistics does not arise.

Also, in a parametric context, Laffont, Ossard, and Vuong (1995) proposed a simulated nonlinear least-squares estimator that uses, for computational reasons, only moments of the winning bid in a first-price auction. While their model is fully parametric and the resulting estimators converge at the root-n rate, our results suggest that identification using this approach may be fragile and estimators using the full set of bids may be less sensitive to misspecification of the valuation distribution.[2]

---

[2]Nonparametric identification of features of a model also permits interpretation of parametric procedures as a plausible statistical approximation rather than treating the parametric specification of the model as prior knowledge, and nonparametric nonidentification results are helpful to shed light on which particular features of a parametric model used for estimation are substantive for identification, as already argued by Roehrig (1988). However, this interpretation also implies that the properties of the corresponding nonparametric estimator are indicative of the quality of this approximation. In this fashion, if nonparametric

Even though this paper focuses on estimation of auctions, the statistical problem is a prototype for other settings in which only the minimum, maximum, or other element of an ordered sample of realizations of a vector of random variables is observed. Athey and Haile (2002) and Komarova (2009) also pointed out the similarities between the problem of estimating valuation distributions from observations of the highest bid and competing risks models in duration analysis, so that estimators for failure time models with nonparametric baseline hazards under endogenous censoring should be expected to share some of the statistical properties of the procedures analyzed in this paper.

There is a loose resemblance between our problem and the irregular identification of finite-dimensional parameters considered by Khan and Tamer (2010). Classical examples covered by their analysis include Manski's (1975) maximum score estimator (see also Horowitz (1992)), Lewbel's (2000) estimator for latent variable models with endogeneity, estimation of average treatment effects (see, for example, Imbens (2004) for an overview), and estimation of the intercept in a semiparametric censored regression model (Andrews and Schafgans (1998)). A common characteristic of these problems is that in each case, an "identification at infinity" argument translates to a failure of a support condition with respect to all or some conditioning variables in any finite sample. As we show, in estimation of auctions from incomplete bidding data, the failure of the support condition arises endogenously from the economic model in that only data on the "winning" (or other specifically ranked) unit are observed. This feature is also shared by nonparametric competing risks and multinomial choice models with a large number of goods; see, for example, Berry, Linton, and Pakes (2004). In contrast to the existing literature on irregular identification, the problem in this paper concerns estimation of an infinite-dimensional parameter, where the irregularity affects different parts of the distribution to be estimated to different degrees. While regular subcases of this estimation problem exist, rates of convergence are shown to depend on the norm for the corresponding parameter space.

We now formally state the estimation problem analyzed in this paper. Section 2 derives optimal convergence rates for nonparametric point estimators and bounds for the valuation distribution for second-price formats under various sets of assumptions, and Section 3 derives the asymptotic distribution of a regularized version of that estimator. Section 4 shows how irregular estimation of the private value distribution affects rates for nonparametric estimation of functionals, including expected revenue and optimal reserve price, and Section 5 shows how to extend the main rate result to the case of first-price auctions and related formats.

1. Description of the problem

In this paper, we consider estimation when we observe data from n independent auctions in which one indivisible object is auctioned. Each auction $i = 1, \ldots, n$ has K bidders and for each auction, the bidders' valuations $(V_1, \ldots, V_K)$ are drawn from a joint distribution $F_K \in \mathcal{F}_K$ with joint support $\mathcal{V}^K \subset \mathbb{R}_+^K$, where $F_K := F_{V_1, \ldots, V_K}(v_1, \ldots, v_K)$ and

---

estimation is possible only at a very slow rate of consistency, we should be very cautious in interpreting a root-n consistent parametric estimator as an approximation to the more complex "true" model.

$\mathcal{F}_K$ denotes a subset of the set of c.d.f.s on $\mathcal{V}^K$ (i.e., the set of upper semicontinuous, nondecreasing functions from $\mathcal{V}^K$ to the unit interval that attain the values 0 and 1 in the closure of $\mathcal{V}^K$). We also let $F_0(v) := \frac{1}{K}\sum_{k=1}^{K} F_{V_k}(v)$ denote the marginal distribution of bidders' valuations, where $F_0 \in \mathcal{F}_0$ the set of c.d.f.s on $\mathcal{V}$ that are consistent with a joint distribution $F_K \in \mathcal{F}_K$.

For our results on consistency and asymptotic distribution of estimates of the valuation distribution and/or bounds, we are not going to impose a particular relationship between the distributions of bidders' valuations and bids in the auction. While the economic interpretation for each of the estimands discussed below—including distributions, bounds or functionals of a private value distribution—crucially depends on more specific economic primitives of the auction, the large sample properties of the corresponding estimators are derived for a common condition on the joint distribution of observable bids.

ASSUMPTION 1.1 (Distributions of Valuations and Bids). *Each auction features a single good and K bidders, where the marginal distribution of bidders' valuations is $F_0(v)$ and the joint distribution for the ordered sample of the K bids $(B_1, \ldots, B_K)$ is given by*

$$G_{1,\ldots,K}(v_1,\ldots,v_K) \equiv G_{1,\ldots,K}(v_1,\ldots,v_K; F_K) := \mathbb{P}_{F_K}(B_1 \le v_1,\ldots,B_K \le v_K),$$

with marginal distributions $G_k(v_k)$ for $k = 1,\ldots,K$. The marginal probability density function (p.d.f.) $f_0(v)$ of V_k is bounded away from zero in the interior of the support and the first p derivatives of $f_0(v)$ are bounded.

Note that in our notation, the valuation V_k and the bid B_k do not necessarily correspond to the same bidder, but for the latter, the index k denotes the rank of the bid, whereas $V_1,\ldots,V_K$ denotes the unordered sample of valuations. Since there are no further restrictions on the joint distribution of V_k across bidders, Assumption 1.1 allows for common values, correlated private values, unobserved heterogeneity, and asymmetric or risk-averse bidders. For our results on the upper bounds on rates of convergence and optimal rates, we focus our attention mostly on estimation of the c.d.f. and functionals of private value distributions using data from second-price auctions under the IPV paradigm; the case of first-price and descending bid auctions is discussed in Section 5.

ASSUMPTION 1.2 (Second Price Auction). *Each auction features a single good and K risk-neutral bidders, where* (i) *participation is exogenous and* (ii) *the auction satisfies symmetric independent private values (IPV), $V_k \overset{i.i.d.}{\sim} F_0(v)$ for some $F_0 \in \mathcal{F}_0$, where* (iii) *any distribution $F \in \mathcal{F}_0$ is absolutely continuous with respect to the Lebesgue measure with density $f(v)$, where the p.d.f. $f_0(v)$ of V_k is bounded away from zero in the interior of the support, and the first p derivatives of $f_0(v)$ are bounded.* (iv) *The auction is sealed-bid second-price or a strategically equivalent format, and participants play weakly dominant strategy with bids $B_k \equiv b^*(V_k) = V_k$.*

Note that Assumption 1.2 implies Assumption 1.1, where the joint distribution $F_K(v_1,\ldots,v_K) = F_0(v_1)\cdots F_0(v_K)$, and the distribution of bids $G_{1,\ldots,K}(v_1,\ldots,v_K; F_K)$ is

given by the joint c.d.f. of the K order statistics for a sample of K independent and identically distributed (i.i.d.) draws from F_0 given in Appendix A. Except for the results on upper bounds on the rates of convergence, we work with the weaker Assumption 1.1 rather than imposing the IPV model in Assumption 1.2. In particular, our results on consistency and asymptotic distribution also apply to construction of bounds from order statistics under the weaker set of conditions. However, we are going to impose throughout this paper that bidders are symmetric, which is not essential for most of our qualitative insights, but a proper treatment would require substantially more notation and case distinctions, and make the exposition harder to follow.[3] We also do not explicitly analyze the case of observable heterogeneity, but the arguments presented in this paper are made conditional on observed auction-specific regressors and number of bidders.

To keep our results general, we allow the data set available to the econometrician to be any r-dimensional subvector of the complete vector $(B_{i1}, \ldots, B_{iK})$ of bids.

ASSUMPTION 1.3 (Observable Bids). *We observe the $k_1 < k_2 < \cdots < k_r$ lowest bids $\mathbf{B}_i = (B_{ik_1}, B_{ik_2}, \ldots, B_{ik_r})$ for n i.i.d. auctions of a single good with K bidders. In particular, $\mathbf{B}_1, \ldots, \mathbf{B}_n$ are i.i.d. draws from the distribution $G_{k_1,\ldots,k_r}(\cdot; F_K)$.*

For example, if Assumption 1.2 holds and we only record the transaction price for each of the n auctions, the observed bids correspond to $\mathbf{B}_i = B_{iK-1} = V_{iK-1}$, the second-highest valuation among potential buyers in the ith auction. Given the sample $\mathbf{B}_1, \ldots, \mathbf{B}_n$, we denote the marginal distribution of the kth bid by

$$\widehat{G}_{nk}(v) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{B_{ik} \leq v\}.$$

Convergence rates for estimators of the distribution with respect to the L_q norm for $1 \leq q < \infty$ and linear functionals turn out to depend on the behavior of the tails of the parent distribution $F_0(\cdot)$. To characterize that relationship, we assume that the p.d.f. $f_0(\cdot)$ is bounded from above by polynomial functions with exponents α_1 and α_2 in the quantile τ for values of τ close to 0 or 1, respectively. More specifically, we make the following assumption.

ASSUMPTION 1.4 (Tail Behavior). *Let $h(\tau; F_0) := f_0(F_0^{-1}(\tau))$, where f_0 is the p.d.f. associated with F_0. Then there exist constants α_1, α_2 such that for low quantiles τ, the behavior of the p.d.f. of V_k is characterized by*

$$\limsup_{\tau_1 \to 0} \tau_1^{-\alpha_1} h(\tau_1; F_0) < \infty$$

---

[3]Note that for the case of asymmetric bidders and observed bidder identities, all arguments made in this paper also hold conditional on the identity of the winner, the second-highest bidder, and so forth in each auction. In particular, if for all bidders $l \neq k$, the Radon–Nikodym derivative of the distribution of bidder k's valuation V_k with respect to that of bidder l's valuation V_l is bounded from above and away from zero, the optimal rate for estimating the distribution of V_k is the same as in the symmetric case.

and in the upper tail of the distribution is characterized by

$$\limsup_{\tau_2 \to 1}(1 - \tau_2)^{-\alpha_2} h(1 - \tau_2; F_0) < \infty$$

for all $F_0 \in \mathcal{F}_0$.

For example, if the p.d.f. of V is bounded from above and away from zero at the lower boundary of its support, the first part of Assumption 1.4 holds with $\alpha_1 = 0$, whereas if V follows a log-normal distribution, then the statement holds for any $\alpha_1 > \frac{9}{5}$.

Under Assumption 1.2, the observed bids are order statistics of samples of independent draws from the parent distribution F. For a given parent distribution $F \in \mathcal{F}_0$, denote the joint c.d.f. of the $(k_1, \ldots, k_r)$th order statistics by

$$G(\mathbf{b}; F) := G_{k_1, \ldots, k_r}(b_{k_1}, \ldots, b_{k_r}; F) := \mathbb{P}_F(B_{k_1} \leq b_{k_1}, \ldots, B_{k_r} \leq b_{k_r}).$$

For example, the c.d.f. for the kth order statistic can be expressed as

$$G_k(b_k; F) := \sum_{m=k}^{K} F(b_k)^m [1 - F(b_k)]^{K-m}$$

$$= \frac{K!}{(k-1)!(K-k)!} \int_0^{F(b_k)} s^{k-1}(1-s)^{K-k} \, dt,$$

whereas a pair of order statistics $B_{(k_1;K)}$, $B_{(k_2;K)}$ has the joint c.d.f.

$$G_{k_1, k_2}\big((b_{k_1}, b_{k_2}); F\big)$$
$$= N(k_1, k_2; K) \int_0^{F_1} \int_0^{F_2} s_1^{k_1-1}(s_2 - s_1)^{k_2-k_1-1}(1-s_2)^{K-k_2} \, ds_2 \, ds_1, \tag{1.1}$$

where for a tuple of integers $0 < k_1 < \cdots < k_r \leq K$, we denote $N(k_1, \ldots, k_r; K) = \frac{K!}{k_1!(k_2-k_1-1)!\cdots(K-k_r)!}$ and $F_s := F(b_{k_s})$; see, for example, David and Nagaraja (2003). We give an expression for the general case in the Appendix.

EXAMPLE 1.1. To frame thoughts, suppose that we observe the winning bid B_{iK} in a sealed-bid independent values second-price auction, which is the highest order statistic for K i.i.d. draws from the population distribution of valuations $F_0(v)$. In this case, the c.d.f. of the observed bid is given by $G_K(v, F) = [F(v)]^K$ and the maximum-likelihood estimator for the parent distribution is given by

$$\widehat{F}_n(v) = \sqrt[K]{\widehat{G}_{nK}(v)} =: \phi_K^{-1}\big(\widehat{G}_{nK}(v)\big),$$

where $\widehat{G}_{nK}(v) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{B_{iK} \leq v\}$ is the empirical c.d.f. of B_{iK}.

In this case, there is a closed form for the nonparametric maximum-likelihood estimator that is also guaranteed to be nondecreasing in v. Furthermore, from elementary arguments, $\widehat{F}_n(v)$ is pointwise consistent at a root-n rate and asymptotically efficient

at every $v \in \mathcal{V}$. However, it turns out that for the problem of inverting distributions of order statistics, pointwise asymptotic arguments can be very misleading when we are interested in estimating the distribution as a function or finite-dimensional parameters of the distribution nonparametrically. We consider convergence rates for estimating the c.d.f. and bounds implied by the different order statistics with respect to functional norms instead. In particular, the asymptotic squared bias of the nonparametric maximum-likelihood (ML) estimator of the c.d.f. from incomplete bidding data will typically be of the same order of magnitude as its variance, so that the question of asymptotic efficiency cannot be posed in a meaningful way.

In the case of Example 1.1, the problem arises because the mapping $\phi_K^{-1}(\tau) = \sqrt[K]{\tau}$ is not Lipschitz continuous in $\tau \in [0, 1]$ for any $K > 0$, which slows down the rates of convergence for $\widehat{F}_n(v)$ as a function.[4]

More generally, for any $k = 1, \ldots, K$, we define the mapping

$$\phi_k(\tau) := \frac{K!}{(k-1)!(K-k)!} \int_0^\tau s^{k-1}(1-s)^{K-k}\, ds, \tag{1.2}$$

which is strictly increasing by inspection. From the facts about order statistics given in Appendix A, the marginal c.d.f. of the kth order statistic in a sample of K i.i.d. observations from a distribution F can be written as

$$G_k(v; F) := \mathbb{P}_F(B_{ik} \leq v)$$
$$= \frac{K!}{(k-1)!(K-k)!} \int_0^{F(v)} s^{k-1}(1-s)^{K-k}\, ds = \phi_k\big(F(v)\big). \tag{1.3}$$

In the following discussion, we, therefore, consider the inverse of this mapping:

$$\psi_k(\tau) := \phi_k^{-1}(\tau) \quad \text{for any } \tau \in [0, 1]. \tag{1.4}$$

Under the symmetric IPV paradigm given in Assumption 1.2, the empirical c.d.f. of the k highest bid estimates the c.d.f. of the kth order statistic $G_k(v; F)$ in (1.3), so that by the usual invariance properties of maximum likelihood, the nonparametric maximum-likelihood estimator (MLE) for the c.d.f. of valuations can be obtained by applying the inverse of the mapping $\phi_k(\cdot)$ to the empirical distribution of the kth bid:

$$\widehat{F}_n(v) = \phi_k^{-1}\big(\widehat{G}_{nk}(v)\big) =: \psi_k(\widehat{G}_{nk}). \tag{1.5}$$

However, note that Assumption 1.1 does not state that the marginal distributions of bids $G_k(v)$ coincide with $G_k(v; F)$, but recent results on identification and inference

---

[4]Since the quantile transformation linking the distribution of observed bid to the parent distribution is a continuous function on the compact set $[0, 1]$, it is also uniformly continuous by the Heine–Cantor theorem, so that the mapping is also continuous with respect to the sup-norm. In this sense, the inverse problem of recovering the parent distribution from the joint distribution of observable bids is not ill-posed. However, to derive a uniformly valid distributional approximation to the estimator, it will be necessary to regularize this inverse, because the local linearization of the problem turns out to be ill-posed even though the original problem is not.

under departures from the benchmark model in Assumption 1.2 derive bounds based on $\psi_k(G_k(v))$.[5]

1.1 *Aggregates and bounds*

Depending on the interpretation of the function $\psi_k(G_k(v))$, the researcher is often interested in combining information from several bids: under the IPV paradigm in Assumption 1.2, we will show that to achieve the optimal rate of convergence for estimating the private value distribution, it is, in general, necessary to use the information on all observable bids, for example, by averaging trimmed versions of estimates obtained from the marginal distributions of the $k_1, \ldots, k_r$ bids. For bounds and hypothesis tests in settings that relax the standard IPV assumptions, the literature has been considering intersection bounds of the form

$$F_0(v) \in \big[\max\{\psi_{l_1}(G_{l_1}(v + \Delta)), \ldots, \psi_{l_s}(G_{l_s}(v + \Delta))\},$$
$$\min\{\psi_{m_1}(G_{m_1}(v)), \ldots, \psi_{m_t}(G_{m_t}(v))\}\big]$$

for all $v \in \mathcal{V}$, where $\{l_1, \ldots, l_s, m_1, \ldots, m_t\} \subset \{k_1, \ldots, k_r\}$ and $\Delta \geq 0$ is a minimal bid increment or other friction.

In addition to estimators of the private values distribution and its functionals in the point-identified case, we also consider estimation of bounds on the c.d.f. of valuations and other functions that aggregate the information from all observable bids under the weaker conditions in Assumption 1.1. In its most general form, we write the estimand as

$$H_0(v) := H^*\big(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v))\big).$$

Examples for aggregation functions $H^*(\cdot)$ include weighted averages

$$H^*\big(F_1(v), \ldots, F_r(v)\big) := w_1 F_1(v) + \cdots + w_r F_r(v), \quad \sum_{q=1}^{r} w_r = 1,$$

intersection bounds

$$H^*\big(F_1(v), \ldots, F_r(v)\big) := \max\{F_1(v), \ldots, F_r(v)\},$$

and Haile and Tamer's (2003) smoothed bounds

$$H^*\big(F_1(v), \ldots, F_r(v)\big) := \sum_{q=1}^{r} F_q(v) \frac{\exp\{\lambda F_q(v)\}}{\displaystyle\sum_{p=1}^{r} \exp\{\lambda F_p(v)\}}$$

for some large value $\lambda > 0$. Our results in Section 2 show that if the IPV framework in Assumption 1.2 is assumed to hold, estimators using trimming can achieve a rate

---

[5]For example, Haile and Tamer (2003) allowed for jump-bidding in ascending auctions, and Aradillas-López, Gandhi, and Quint (2013) considered unobserved heterogeneity.

convergence corresponding to that of the "best" component in $\psi_{k_1}(G_{k_1}), \ldots, \psi_{k_r}(G_{k_r})$, whereas without the IPV assumption, the optimal rate depends on the least precise component.

2. Estimation of the c.d.f. and optimal rates

In this section, we give bounds on the rate of convergence of the nonparametric estimator for the parent distribution $F_0(v)$. Due to the nature of the problem, convergence rates will depend on the norm on the function space $\mathcal{F}_0$. More specifically, we consider convergence with respect to the L_∞ or sup-norm $\|h\|_\infty := \sup_{v \in \mathcal{V}} |h(v)|$ for any $h \in \mathcal{F}_0$, and the L_q norm with respect to Lebesgue measure is defined as $\|h\|_q := (\int |h(v)|^q \, dv)^{1/q}$ for $1 \le q < \infty$.

Following Stone (1980), we say that r_n is an *upper bound to the rate of convergence* of $\widehat{F}_n$ under the norm $\| \cdot \|$ if

$$\liminf_n \sup_{F_K \in \mathcal{F}_K} P_{F_K}\big(\|\widehat{F}_n - F\| > cr_n^{-1}\big) > 0 \tag{2.1}$$

for any sequence of estimators $\{\widehat{F}_n\}_{n \ge 0}$ and

$$\lim_{c \to 0} \liminf_n \sup_{F_K \in \mathcal{F}_K} P_{F_K}\big(\|\widehat{F}_n - F\| > cr_n^{-1}\big) = 1. \tag{2.2}$$

These bounds are not specific to any given estimator $\widehat{F}_n$ in the problem. We establish these bounds on the rate of convergence by constructing a worst-case scenario in terms of a true distribution $F_0 \in \mathcal{F}_0$ and a local perturbation that cannot be distinguished with certainty by any statistical procedure. In principle, this "hardest" estimation problem may be different for different estimators and/or different measures of distance, but it turns out that for our purposes, the form of the perturbations that determine the sharpest bound on the rate is the same for all problems we consider.

Also, r_n is called an *achievable rate of convergence* if we can construct a sequence $\{\widehat{F}_n\}$ of estimators that satisfy

$$\lim_{c \to \infty} \limsup_n \sup_{F_K \in \mathcal{F}_K} P_F\big(\|\widehat{F}_n - F\| > cr_n^{-1}\big) = 0. \tag{2.3}$$

If a rate r_n is both achievable and an upper bound to the rate of convergence, we say that it is the *optimal rate of convergence* for all nonparametric estimators of $F_0 \in \mathcal{F}_0$.

The statement in (2.3) can be read as a requirement that the estimator normalized by r_n^{-1} is concentrated on a compact set with respect to the relevant norm for large samples or, equivalently, that the rescaled distance between the estimator and the estimand is asymptotically tight.

The notion of optimality for the rate of convergence is of course very weak and in many cases, we will be able to construct several distinct and equally reasonable estimators, all of which achieve the optimal rate. However, our main motivation for establishing rate optimality is to demonstrate that the upper bound cannot be improved on and, therefore, constitutes a useful measure for the difficulty of the nonparametric estimation problem at hand.

2.1 *Consistent estimation of the valuation distribution*

For clarity of the exposition, we focus on the case in which only the kth lowest bid out of K is observed. We argue later on in this section that this is without loss of generality for a discussion of the achievable rate of convergence for nonparametric estimators of the valuation distribution. In particular, we will show that the upper bound on the rate in Theorem 2.3 can be achieved by an estimator that combines the inverted empirical marginal c.d.f.s of the observed bids in a straightforward manner.

By the invariance principle, the nonparametric maximum-likelihood estimator for the c.d.f. of valuations can be obtained by applying the inverse of the mapping $\phi_k(\cdot)$ to the empirical distribution of the kth bid,

$$\widehat{F}_n(v) = \phi_k^{-1}(\widehat{G}_{nk}(v)) =: \psi_k(\widehat{G}_{nk}), \tag{2.4}$$

where $\psi_k(\tau) := \phi_k^{-1}(\tau)$ for $\tau \in [0, 1]$. To illustrate the main idea behind the slow rates of convergence, we first look at the setting in Example 1.1.

EXAMPLE 2.1 (Example 1.1, Continued). For the problem of estimating the parent distribution from the highest order statistic, the nonparametric maximum-likelihood estimator is

$$\widehat{F}_n(v) = \sqrt[K]{\widehat{G}_{nK}(v)}$$

for the empirical c.d.f. of the highest bid, $\widehat{G}_{nK}$. To see why this estimator can be consistent at best at a rate $n^{-1/K}$ with respect to the sup-norm, consider the estimate around the lowest observation in a given sample, $\tilde{B}_n := \min\{B_{1K}, \ldots, B_{nK}\}$. The size of the jump in the empirical c.d.f. $\widehat{G}_{nK}$ at $\tilde{B}_n$ is $\frac{1}{n}$, which translates into a jump in the estimate by $\widehat{F}_n(\tilde{B}_n^+) - \widehat{F}_n(\tilde{B}_n^-) = (\frac{1}{n})^{1/K}$. However, if the true c.d.f. $F_0(v)$ is continuous, the limits at $\tilde{B}_n$ from the right and from the left coincide so that we can use the triangle inequality to bound

$$\sup_{v \in \mathcal{V}} |\widehat{F}_n(v) - F_0(v)| \geq \max\{|\widehat{F}_n(\tilde{B}_n^+) - F_0(\tilde{B}_n^+)|, |\widehat{F}_n(\tilde{B}_n^-) - F_0(\tilde{B}_n^-)|\} \geq \frac{1}{2}\left(\frac{1}{n}\right)^{1/K}$$

for any realization of the sample. Theorems 2.3 and 2.1 below imply that this is, in fact, also the optimal uniform rate if we impose no further smoothness restrictions on the derivatives of $F(v)$.

Alternatively, consider the pointwise asymptotic mean-squared error (MSE) of this estimator, which is given by

$$\text{AsyMSE}(\tau) := \mathbb{E}\left[\frac{n}{K^2}\left(\widehat{G}_{nK}(F_0^{-1}(\tau)) - G_K(F_0^{-1}(\tau); F_0)\right)^2\left(G_K(F_0^{-1}(\tau); F_0)\right)^{2/K-2}\right]$$

$$= \frac{(1 - \tau^K)\tau^{2-K}}{K^2}$$

using the delta rule. We show in Section 3 that this approximation to the MSE is not uniform in τ and, furthermore, we can see that for any $K > 2$, the pointwise MSE along a sequence of quantiles $\tau_n \to 0$, $\frac{1}{n}\text{AsyMSE}(\tau_n)$ converges to zero only if $\tau_n n^{1/K} \to \infty$.

As this example illustrates, the difficulty consists in that the inverse mapping from the distribution of observable bids to the parent distribution may not be Lipschitz continuous in some cases, but may have divergent slope in the tails of the distribution. The problem is mitigated by the fact that the variance of the empirical distribution decreases linearly as we move out into the tails, but persists unless we observe bids that are close to the lowest and highest order statistics in a sufficiently large number of auctions.

As discussed earlier, the nonparametric MLE for observations on a single observed bid B_{ik} involves the inverse mapping $\psi_k(\cdot)$, which is not Lipschitz continuous since its first derivative $\psi'_k(\tau) = \frac{1}{\phi'_k(\psi_k(\tau))} = \frac{1}{N(k,K)}\psi_k(\tau)^{1-k}(1-\psi_k(\tau))^{k-K}$, which behaves like $\tau^{1/k-1}$ for τ close to zero and behaves like $(1-\tau)^{1/(K-k+1)-1}$ for τ close to 1, which diverge to infinity if $k > 1$ or $K - k > 2$, respectively.

We now give a general uniform consistency result for nonlinear transformations of the empirical c.d.f. with finitely many singularities of this form, which we then apply to the problem of estimating the c.d.f. of valuations from data on a particular bid in n i.i.d. auctions.

CONDITION 2.1. Let $\psi \in \mathcal{C}^2([0,1])$ be bounded and suppose that (i) there are only finitely many points $\tau_1^* < \tau_2^* < \cdots < \tau_S^* \in [0,1]$ such that $\psi'(\tau)$ diverges in a neighborhood of those points and (ii) there exist finite constants $A_s > 0$ and $\delta_1, \ldots, \delta_S$ such that for all $s = 1, \ldots, S$, $\lim_{\tau \to \tau_s^*} |\frac{\psi(\tau)-\psi(\tau_s^*)}{|\tau-\tau_s^*|^{\delta_s}}| = A_s$. Also assume that $\psi'(\tau)$ is monotone and does not switch sign on any interval of the form (τ_{i-1}^*, τ_i^*) for two adjacent singular points.

Note that by standard arguments, this condition also implies that for $\delta_s \neq 0$, the first derivative of $\psi(\cdot)$ satisfies

$$\lim_{\tau \to \tau_s^*} \left| \frac{\psi'(\tau) - \psi'(\tau_s^*)}{|\tau - \tau_s^*|^{\delta_s-1}} \right| = A_s$$

for the same constants as in the statement of Condition 2.1.

THEOREM 2.1. Suppose that Condition 2.1 holds and let $\delta := \min\{\delta_1, \ldots, \delta_S, \frac{1}{2}\}$. Then the rate $r_n = n^\delta$ is achievable for an estimator of the function $\psi(G_0(v))$ with respect to the sup-norm.

See Appendix B for a proof. We can now use this result to establish a uniform rate of consistency for the estimator in (2.4).

PROPOSITION 2.1. Suppose Assumptions 1.2 and 1.3 hold with $r = 1$ and $k_1 = k$. Then the estimator $\widehat{F}_n$ in (2.4) achieves the rate $r_n = n^\lambda$ with $\lambda = \min\{\frac{1}{k}, \frac{1}{K-k+1}, \frac{1}{2}\}$ under the sup-norm.

PROOF. By Assumption 1.2, we have that for the distribution of the observable bid B_{ik}, $\psi_k(G_k(v; F)) = F(v)$ for any $v \in \mathcal{V}$. From the previous discussion, it is straightforward to verify that Condition 2.1 holds for the mapping $\psi_k(\cdot)$ with $\tau_1^* = 0$, $\tau_2^* = 1$, $\delta_1 = \frac{1}{k}$, and

$\delta_2 = \frac{1}{K-k+1}$, where $A_1 = \frac{1}{N(k,K)^{1/k}}$ and $A_2 = \frac{1}{N(k,K)^{1/(K-k+1)}}$. Hence we can apply Theorem 2.1 with $\delta = \min\{\frac{1}{k}, \frac{1}{K-k+1}\}$, so that the estimator in (2.4) is uniformly consistent with rate $r_n = n^\delta$, which is at the same time the upper bound on the rate of convergence established in Theorem 2.3 for the special case of a single observable bid for each auction. □

Note that the consistency results in this section so far were only about the case of a single observed bid per auction. However, as the next subsection shows, it is straightforward to extend the argument and establish achievability of the bound on the rate established in Theorem 2.3 below by considering a procedure that combines the estimators obtained from inverting the marginal distribution of each bid separately.

2.2 *Aggregation of estimators*

We now extend the statement from Proposition 2.1 to averages and other aggregates of trimmed estimates of the form described in Assumption 2.1. Denote the empirical c.d.f. of the k highest bid with $\widehat{G}_{nk}$, and assume that for all $k = k_1, \ldots, k_r$, the mapping $\psi_k(\cdot)$ satisfies Condition 2.1 with exponents δ_{ks} for $s = 0, \ldots, S$ for a common grid of singular points $\tau_0^*, \ldots, \tau_S^*$. For trimming sequences $b_{nks} \to 0$, define $D_{nk}(v) := \mathbb{1}\{|\widehat{G}_{nk}(v) - \tau_s^*| \geq b_{nks}$ for all $s = 0, \ldots, S\}$. We will give conditions on the trimming sequences b_{nks} below; in particular, the optimal rates depend on whether the standard IPV framework is assumed to hold or not.

We now consider estimators of the form

$$\widehat{H}_n(v) := H\big(\psi_{k_1}(\widehat{G}_{nk_1}(v)), \ldots, \psi_{k_r}(\widehat{G}_{nk_r}(v)), D_{nk_1}(v), \ldots, D_{nk_r}(v)\big),$$

where $D_{nk}(v)$ is an indicator variable that equals zero if v falls into the trimming range for the kth component $\psi_k(\widehat{G}_{nk}(v))$. We also assume that $H(\cdot, \cdot)$ and $H^*(\cdot)$ satisfy the following conditions.

ASSUMPTION 2.1. (i) *The aggregation function* $H(F_{k_1}, \ldots, F_{k_r}, D_{k_1}, \ldots, D_{k_r})$ *is Lipschitz continuous with Lipschitz constant* $M_H < \infty$ *in* F_k *if* $D_k = 1$ *for all* $k = k_1, \ldots, k_r$, (ii) *satisfies* $H(F_1, \ldots, F_k, 1, \ldots, 1) = H^*(F_1, \ldots, F_k)$, (iii) $H^*(F, \ldots, F) \equiv \tilde{H}(F)$ *for all* F *and some function* $\tilde{H}(\cdot)$, *and* (iv) *is constant in* ψ_k *in a neighborhood of radius* $r > 0$ *around* $\psi^{-1}(\tau_s^*)$ *for all* $s = 1, \ldots, S$ *if* $D_k = 0$.

Assumption 2.1 allows for trimmed maxima, minima, weighted averages, and Haile and Tamer's (2003) smoothed bounds on the trimmed estimates, where, for the average,

$$H(F_{k_1}, \ldots, F_{k_r}, D_{k_1}, \ldots, D_{k_r})$$

$$:= \begin{cases} \dfrac{1}{r}\displaystyle\sum_{q=1}^{r}\psi_{k_q}(\tau_{s_q}^*) & \text{if } \big|\psi_{k_q}^{-1}(F_{k_q}) - \tau_{s_q}^*\big| < b_{nk_q} \text{ for all } q = 1, \ldots, r, \\[4mm] \dfrac{\displaystyle\sum_{k=k_1,\ldots,k_r} D_k F_k}{\displaystyle\sum_{k=k_1,\ldots,k_r} D_k} & \text{otherwise,} \end{cases}$$

and the maximum

$$H(F_{k_1}, \ldots, F_{k_r}, D_{k_1}, \ldots, D_{k_r})$$
$$:= \max\{D_{k_1}F_{k_1} + (1 - D_{k_1})\psi_{k_1}(\tau_{s_1}^*), \ldots, D_{k_r}F_{k_r} + (1 - D_{k_r})\psi_{k_r}(\tau_{s_r}^*)\},$$

where $\tau_{s_k}^* = \arg\min\{|\psi_k^{-1}(F_k) - \tau| : \tau \in \{\tau_1^*, \ldots, \tau_S^*\}\}$.

For the achievable rate of the aggregate estimator and the optimal trimming strategy, we have to distinguish two cases in the following theorem.

THEOREM 2.2. *Let* $B_{i1}, \ldots, B_{iK}$, $i = 1, \ldots, n$, *be a sample of i.i.d. observations with marginal distributions* $G_1(v_1), \ldots, G_K(v_1)$. *Suppose that Assumption* 2.1 *holds, and that the mappings* $\psi_k : [0, 1] \to [0, 1]$ *satisfy Condition* 2.1 *with constants* δ_{ks}, $s = 0, \ldots, S$ *and* $k = k_1, \ldots, k_r$. *Then* $\widehat{H}_n(v)$ *is consistent for* $H_0(v)$ *under the sup-norm with rate* $r_n = n^\lambda$, *where the following statements hold:*

(a) $\lambda = \max_{k=k_1,\ldots,k_r} \min\{\delta_{k1}, \ldots, \delta_{kS}, \frac{1}{2}\}$ *if* $\psi_k(G_k(v)) = F_0(v)$ *for* $k = k_1, \ldots, k_r$ *for some function* $F_0(v)$, *and the trimming sequences* $b_{nks}n^{-\delta_{ks}/\bar{\delta}_s} \to b_{ks} \in (0, \infty)$, *where* $\bar{\delta}_s := \max\{\delta_{k_1s}, \ldots, \delta_{k_rs}\}$.

(b) $\lambda = \min_{k=k_1,\ldots,k_r} \min\{\delta_{k1}, \ldots, \delta_{kS}, \frac{1}{2}\}$ *if* $\psi_k(G_k(v)) \neq \psi_{k'}(G_{k'}(v))$ *for some* k, k', *and the trimming sequences* $b_{nks}n^{-1} \to b_{ks} \in [0, \infty)$. *In particular, this rate is attainable without trimming.*

In terms of applications, we can think of case (a) as having a correct model for the marginal distributions of the different bids and then combining the estimates of a common implied c.d.f. in a way that achieves the optimal rate. The second case includes the scenarios of interest for bounds applications in which different bids imply different bounds on the marginal distribution of valuations.

The conditions on the trimming rates have to be different between cases (a) and (b) for the following reason: in part (a), the individual estimates are "substitutes" for each other and we can simply ignore the estimates with large "local" variance, whereas in the setting of (b) the sharpest bound at a given value of v may be determined by the least precisely estimated component $\psi_k(\widehat{G}_{nk}(v))$. However, note that since the rate for $\widehat{H}_n$ in part (b) is determined by this worst case, trimming sequences that satisfy $b_{nks}n^{-1} \to b_{ks} > 0$ will yield a strict improvement over no trimming in all other cases.

Noting that under IPV for sealed-bid second price auctions, Assumption 1.2, the implied private value distributions $\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v))$ from the different observable bids all coincide, we can directly apply the two cases (a) and (b) in Theorem 2.2 to derive implications for estimation of auctions both with and without the IPV assumptions.

COROLLARY 2.1. *Suppose Assumptions* 1.3 *and* 2.1 *hold for an auction model, and let* $S = 1$ *and* $\tau_0 = 0$, $\tau_1 = 1$. *Then* $\widehat{H}_n(v)$ *is consistent for* $H_0(v)$ *at the rate* n^λ, *where the following statements hold:*

(a) $\lambda = \min\{\frac{1}{k_1}, \frac{1}{K-k_r+1}, \frac{1}{2}\}$ *if, in addition, Assumption* 1.2 *(i.e., the canonical IPV paradigm) holds, and the trimming sequences satisfy* $b_{nk0}n^{-k_1/k} \to b_{k0} < \infty$ *for the lower tail and* $b_{nk1}n^{-(K-k_r+1)/(K-k+1)} \to b_{k1} < \infty$ *for the upper tail of the empirical distribution* $\widehat{G}_{kn}$.

(b) $\lambda = \min\{\frac{1}{k_r}, \frac{1}{K-k_1+1}, \frac{1}{2}\}$ *if, in addition, Assumption* 1.1 *holds, and the trimming sequences satisfy* $b_{nk0}n^{-k_r/k} \to b_{k0} < \infty$ *for the lower tail and* $b_{nk1}n^{-(K-k_1+1)/(K-k+1)} \to b_{k1} < \infty$ *for the upper tail of the empirical distribution* $\widehat{G}_{kn}$.

Note that the roles between k_1 and k_r have switched between cases (a) and (b): since in part (a), the estimands $\psi_k(G_k)$ are the same for all $k = k_1, \ldots, k_r$ by assumption, with appropriate trimming rates, they are driven by the most informative bid at every value v, whereas for bounds, for example, a maximum or minimum may be attained at any of the components alone, so that the overall rate of convergence is driven by the *least* precisely estimable component at every given point. For example, the bounds estimated by Haile and Tamer (2003, Figure 10), are based on all bids for auctions of sizes up to $K = 8$. While the resulting estimated bounds on the private value distribution are clearly informative, assuming that auctions of all sizes increase at the same rate with n, the optimal rate of convergence with respect to the *sharp* bounds implied by Theorems 2.2 and 2.4 is, in fact, as slow as $r_n = n^{-1/8}$. To put the result in perspective, note that this is the same rate as for nonparametric estimation of a 12-dimensional density under the default assumption of bounded second partial derivatives. It should also be noted that the rate is equally slow if we restrict our attention only to either tail of the distribution, since the sharp bound on the upper tail may be determined by one of the lower bids.

2.3 *Upper bounds on the rate for estimating* $F_0(v)$

The slow rates of consistency in Proposition 2.1 are, in fact, not an artifact of the order statistics approach, but we now show that without additional smoothness assumption on the distribution of valuations, no nonparametric estimator can possibly achieve a faster rate. To give an intuition for this result, consider again the problem of estimation if only the winning bid is observed.

EXAMPLE 2.2. Consider again the setting from Example 1.1, where only the highest of K bids is observed in n i.i.d. auctions. To estimate $F_0(v)$ at or below the τ quantile, our sample has to contain a sufficient number of observations of valuations less than $v_\tau := F_0^{-1}(\tau)$ to be able to distinguish the true distribution F_0 from any alternative $\tilde{F}$ that coincides with F_0 for all values of $v \geq v_\tau$. Without smoothness restrictions, one possible alternative $\tilde{F}$ in this class puts an atom of mass τ on v_τ and sets $\tilde{F}(v) = 0$ for all $v < v_\tau$, so that $\sup_{v \in \mathcal{V}} |\tilde{F}(v) - F_0(v)| \geq \tau$. However, for the highest bid for the ith observed auction, we have that $B_{iK} \leq v_\tau$ with probability equal to $G_K(v_\tau) = \tau^K$, which vanishes much faster than the size of the perturbation as we let τ shrink to 0. For example, for $K = 6$ and $\tau = 0.25$, we can see from a simple back of the envelope calculation that with probability 78.4%, a sample of $n = 1000$ i.i.d. auctions does not contain one single observation $B_{iK} \leq v_{0.25}$ and is, therefore, completely uninformative about F_0 for values $v \leq v_{0.25}$.

This simple calculation illustrates that what constitutes the "tails" of a distribution for practical purposes depends crucially on the implied convergence rate. For the magnitudes considered in this previous example, the irregularity implied by the functional mapping $\psi_k(\cdot)$ affects estimation even at the third or fourth deciles of the parent distribution.

With n i.i.d. IPV second-price auction for which we observe a vector $(B_{ik_1}, \ldots, B_{ik_r})$ of r different bids, we can establish the following upper bound on the rate of convergence.

THEOREM 2.3. *Let $\widehat{F}_n$ be an estimator for F and let p be the number of bounded derivatives of f as defined in Assumption 1.2. Then under Assumptions 1.2–1.4, $r_n = n^\lambda$ is an upper bound on the rate of convergence satisfying (2.1) and (2.2), where the following equalities hold:*

(a) $\lambda = \min\{\frac{p+1}{k_1+p}, \frac{p+1}{K-k_r+1+p}, \frac{1}{2}\}$ *for the sup-norm.*

(b) $\lambda = \min\{\frac{q(p+1)-\alpha_1+1}{q(k_1+p)}, \frac{q(p+1)-\alpha_2+1}{q(K-k_r+1+p)}, \frac{1}{2}\}$ *for the $L_q(\mu)$ norm.*

(c) $\lambda = \frac{1}{2}$ *under the norm $\|\cdot\|_q$ for any $q \geq 1$ including $q = \infty$, if we restrict the function to a compact subset $A \subset \operatorname{int} \mathcal{V}$.*

In particular, if $p = 0$, the convergence rate in part (a) *of Theorem 2.2 is optimal in the sense of Stone (1980).*

It is important to notice that this bound depends crucially on the number K of bidders in the auction and on which particular bids are observed. In particular, for the setting of Example 1.1, for $p = 0$, the upper bound on the uniform rate of convergence implied by Theorem 2.3 is $r_n = n^{1/K}$, which is the same as the result of our informal discussion of the problem before. However, the discussion in Example 2.1 also demonstrates that for $p \geq 1$, the unrestricted nonparametric MLE does not achieve the upper bound on the rate in part (a) of this theorem.

Note that the shape of the tails of the parent distribution—parameterized by α_1, α_2— affects the rate only under the L_q norm, but not under the sup-norm. This should not come as a surprise since the sup-norm on a function space is invariant to one-to-one transformations of the domain—in particular the integral transformation $v \mapsto F_0(v)$— whereas the L_q norm is not. In particular, for q large, the difference between the expressions for the respective exponents λ in parts (a) and (b) of the statement becomes small.

Part (c) of Theorem 2.3 essentially states that the problem of irregularity is restricted to the tails of the distribution. In particular, if the observed bids (e.g., winning bid or transaction price) are from the upper tail of the distribution, the difficulty consists mainly in estimating the c.d.f. for the lower tail, whereas many policy parameters such as expected revenue or optimal reserve price are more sensitive to properties of its upper tail. However, we show in Section 4 that it is wrong to conclude from this that those functionals of the private value distribution were unaffected by the irregularity in general, and we derive (generally slower than parametric) bounds on the respective rates of convergence for several scalar policy parameters.

The considerations behind the result in this theorem extend to the case of asymmetric bidders, but a clean formulation of the corresponding results and proofs would require additional notation and a different exposition, and we therefore leave this to future research.

The previous result on rates for estimation of the c.d.f. can be modified to cover construction of bounds on the private value distribution and other ways to aggregate the information from several bids when the IPV assumption fails. However, since the arguments leading to bounds of this type do not assume one single marginal valuation distribution $F_0(v)$, we have to be careful how we adjust the likelihood-based argument for the upper bounds on rates in Theorem 2.3. We summarize the conditions on the bounds and the joint distribution of bids and valuations in the following condition.

CONDITION 2.2. Suppose we observe the $k_1 < \cdots < k_r$ lowest bids and let $G(b_{k_1}, \ldots, b_{k_r})$ be the joint distribution of those observed bids. Furthermore, there exist $\underline{u}, \overline{v}$ in the interior of $\mathcal{V}$ such that for all $u \leq \underline{u}$ and all $v \geq \overline{v}$, there exists a joint distribution of observable bids $G(b_{k_1}, \ldots, b_{k_r}; F_K)$ and a marginal valuation distribution $F_0(v)$ with density $f_0(v)$ such that the first p derivatives of $f_0(v)$ are bounded and the following conditions hold:

(i) $G_{k_r}(u; F_K) \leq \kappa F_0(u)^{k_r}$ and $G_{k_1}(v; F_K) \geq (1 - F_0(v))^{k_1}$ for some constant $\kappa > 0$.

(ii) $G_k(v; F_K)$ is nondecreasing in F_0 for each $k = k_1, \ldots, k_r$, and $\psi_{k_r}(G_{k_r}(u; F_K + \varepsilon h)) - \psi_{k_r}(G_{k_r}(u; F_K)) \geq \kappa \varepsilon h(u)$ and $\psi_{k_1}(G_{k_1}(v; F_K + \varepsilon h)) - \psi_{k_1}(G_{k_1}(v; F_K)) \geq \kappa \varepsilon h(v)$ for $|\varepsilon|$ small enough and any function $h(v)$ with $|h(v)| \leq 1$ such that $F_K + \varepsilon h \in \mathcal{F}_K$.

(iii) $H^*(\cdot)$ is nondecreasing in each of its component, and the one-sided derivatives of $H^*(\psi_{k_1}(G_{k_1}(u; F_K)), \ldots, \psi_{k_r}(G_{k_r}(u; F_K)))$ with respect to the rth component and of $H^*(\psi_{k_1}(G_{k_1}(v; F_K)), \ldots, \psi_{k_r}(G_{k_r}(v; F_K)))$ with respect to the first component exist and are bounded from below by $b_H > 0$ in a neighborhood of $G_{k_1}(v; F_K))$, and $\psi_{k_r}(G_{k_r}(u; F_K))$.

In particular, this class of distributions also satisfies Assumptions 1.1 and 1.3, so that in the case $p = 0$, Corollary 2.1 establishes achievability of the upper bound stated below. Loosely speaking, parts (i)–(iii) imply that the class of models for which we construct the bounds includes a case for which the joint distribution of bids behaves similarly to a joint distribution of order statistics, and part (iv) requires that the sharp bounds are attained by the least precisely estimable component of the intersection bound for some submodel satisfying (i)–(iii).

THEOREM 2.4. *Suppose Condition 2.2 holds. Then $r_n = n^\lambda$ is an upper bound on the rate of convergence for estimating $H_0(v) := H^*(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v)))$ with respect to the sup-norm, where $\lambda = \min\{\frac{p+1}{k_r+p}, \frac{p+1}{K-k_1+1+p}, \frac{1}{2}\}$. In particular, if $p = 0$, the convergence rate in part* (b) *of Theorem 2.2 is optimal.*

This result applies to the bounds derived in Haile and Tamer (2003) and Aradillas-López, Gandhi, and Quint (2013) if the relevant class of distributions of bids and valuations includes parametric models of jump bidding, endogenous entry of bidders, or unobserved heterogeneity, respectively, that satisfy Condition 2.2.

3. ASYMPTOTIC DISTRIBUTION

As argued before, the linearized version of inverting the distribution of an order statistic is always ill-posed for $K > 1$, even though the original problem is not. Since Gaussian approximations to distributions of estimators rely on such a linear approximation, we have to regularize the functional mapping $\psi_k(\cdot)$ between the empirical c.d.f. of the order statistic and the parent distribution defined in (1.4) so as to control its curvature as we approach the singular points of the mapping.

EXAMPLE 3.1 (Example 1.1, Continued). From Donsker's theorem, the empirical c.d.f. for the highest bid, $\sqrt{n}(\widehat{G}_{nK}(v) - G_K(v)) \xrightarrow{d} N(0, G(v)(1 - G(v)))$ uniformly in v. The quantile transformation $\psi_K(\tau) := \sqrt[K]{\tau}$ is strictly monotone and uniformly continuous in $\tau \in [0, 1]$, but the corresponding functional mapping $\psi_K(G)$ is not Hadamard differentiable. From the delta rule,

$$\sqrt{n}(\widehat{F}_n(v) - F_0(v)) \xrightarrow{d} N\big(0, G(v)\big(1 - G(v)\big)\big|\psi_K'\big(G(v)\big)\big|^2\big)$$

pointwise in $v \in \mathcal{V}$. However, this convergence is generally not uniform in v, the pointwise approximation becomes worse as we approach the lower bound of the support, and $\psi_K'(G(v))$ diverges. The fact that the variance of $\widehat{G}_{nK}(v)$ decreases linearly in $G(v) \to 0$ mitigates, but does not resolve, this problem as long as $K \geq 3$.

To address this difficulty, notice that in the proof of the functional delta method (see, e.g., Theorem 20.8 in van der Vaart (1998)), the requirement of Hadamard differentiability can be weakened to approximability by a sequence of functions that satisfy the following condition.

CONDITION 3.1. (i) There is an estimator $\widehat{G}_n$ for G_0 such that $r_n(\widehat{G}_n - G_0) = O_P(1)$, and (ii) for the sequence of maps $\tilde{\psi}_n$ between normed spaces $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ and $(\Gamma, \|\cdot\|_\Gamma)$, there exist continuous linear maps $\tilde{\psi}_n' : \mathbb{B} \to \Gamma$ such that for every fixed h,

$$\lim_{n\to\infty}\big\|r_n\big(\tilde{\psi}_n\big(G_0 + r_n^{-1}h_n\big) - \tilde{\psi}_n(G_0)\big) - \tilde{\psi}_n'(h)\big\|_\Gamma \to 0$$

for all $h_n \to h$ such that the map $\psi_n(G_0 + r_n^{-1}h_n)$ is defined.

In other words, if we find an appropriate way to smooth the mapping $\psi_k(\cdot)$ depending on sample size, we can control the error in the linear approximation. We can then replace Hadamard differentiability of $\psi(\cdot)$ in the original proof with the weaker requirement from Condition 3.1 to obtain the same conclusion, which is stated in the following lemma.

LEMMA 3.1. *Suppose Condition 3.1 holds for a sequence of mappings $\tilde{\psi}_n$. Then for every $v \in [\underline{v}, \overline{v}]$,*

$$r_n\big(\tilde{\psi}_n(\widehat{G}_n) - \tilde{\psi}_n(G_0)\big) \rightsquigarrow \mathbb{G}$$

uniformly in $v \in [\underline{v}, \overline{v}]$, where $\mathbb{G}$ is a Brownian bridge.

We are now going to propose a regularization of the estimation problem that leads to an asymptotically Gaussian estimator $\widehat{F}_n$ of F_0. For the case of one observed bid corresponding to the kth order statistic, the nonparametric MLE is given by $\psi_k(\widehat{G}_{k,n}(v))$, where $\widehat{G}_{k,n}$ is the empirical c.d.f. of the kth lowest bid, and for every $k = 1, \ldots, K$, $\psi_k(\tau) := \phi_k^{-1}(\tau)$ is a strictly monotonic continuous one-to-one mapping $\psi_k : [0, 1] \to [0, 1]$ from the unit interval onto itself, which can be obtained from inverting (1.2). For any number of bidders greater than two, this mapping is uniformly continuous, but not Lipschitz continuous on $[0, 1]$. Therefore, $\psi_k(\tau)$ is, in particular, not Hadamard differentiable, so the functional delta rule does not apply to $\psi_k(\cdot)$ itself.

However, it will be possible to approximate the mapping with a regularized transformation $\tilde{\psi}_k(\tau; a_n)$, where a_n is a sequence of tuning parameters and where the regularized mapping satisfies $\sup_{\tau \in [0,1]} |\tilde{\psi}'_k(\tau, a_{nk})| \leq a_{nk}$. More specifically, we define

$$\tau^*_{1k}(a) := \min\{\tau \in [0, 1] : \psi'_k(\tau) \leq a\}$$

and

$$\tau^*_{2k}(a) := \min\{\tau \in [0, 1] : \psi'_k(1 - \tau) \leq a\}$$

for any $a \geq 1$, and propose the modification

$$\tilde{\psi}_k(\tau, a) := \begin{cases} a\tau & \text{if } \tau \leq \tau^*_{1k}(a), \\ 1 - a\tau & \text{if } \tau \geq 1 - \tau^*_{2k}(a), \\ a\tau + w_k(\tau, a)(\psi_k(\tau) - a\tau) & \text{otherwise,} \end{cases} \tag{3.1}$$

where for every a, $w_k(\tau, a)$ is twice continuously differentiable in τ, $w_k(\tau^*_{jk}(a), a) = w'_k(\tau^*_{jk}(a), a) = 0$ for $j = 1, 2$, $0 \leq w'(\tau, a) \leq \frac{a}{\phi(\tau) - a}$, and $w_k(\frac{1}{2}, a) = 1$. This specification ensures that $\tilde{\psi}_k(G, a)$ applied to any c.d.f. again yields a valid c.d.f. Furthermore, $\tilde{\psi}_k(\tau, a_n)$ is differentiable for any a and is Lipschitz with constant a.

For a given choice of a_{nk}, we can now define the estimator from inverting the empirical c.d.f. of the k highest bid by

$$\widehat{F}_{nk}(v) := \tilde{\psi}_k(\widehat{G}_{nk}(v), a_{nk}) \tag{3.2}$$

for any $k = k_1, \ldots, k_r$. As with the estimator with trimming introduced in the previous section, we can aggregate these r different estimators into

$$\widehat{F}_n(v) := \frac{1}{r} \sum_{s=1}^{r} \widehat{F}_{nk_s}(v).$$

Compared to the estimator with trimming, this smoothed estimator has the advantage that there are no discontinuous jumps at the boundaries of the trimming intervals; furthermore, it can be seen easily that this estimator is guaranteed to be nondecreasing.

It is crucial to notice that this regularization, including the choice of the tuning parameter a_{nk}, only depends on properties of the mapping $\psi_k(\cdot)$, which are known, but not on any features of the data-generating process. In particular when conditioning on

a vector of observed covariates X_i, the choice for a_{nk} should not be made dependent on the value of X_i.

To characterize the distribution of the joint estimator, define

$$\widehat{S}_n(v) := \sum_{s,t=1}^{r} \left\{ \left[\psi'_{k_s}(\widehat{G}_{nk_s}(v)) \psi'_{k_t}(\widehat{G}_{nk_t}(v)) \right]^{-1} \right.$$
$$\left. \times \frac{1}{n} \sum_{i=1}^{n} \left(\mathbb{1}\{B_{ik_s} \leq v\} - \widehat{G}_{nk_s}(v) \right) \left(\mathbb{1}\{B_{ik_t} \leq v\} - \widehat{G}_{nk_t}(v) \right) \right\}.$$

We can now give rates for the bound a_{nk} of the slope that ensure a uniform Gaussian approximation to the distribution of the (regularized) estimators $\widehat{F}_{nk}$ and $\widehat{F}_n$.

THEOREM 3.1. *Suppose Assumptions* 1.1 *and* 1.3 *hold, and that for the regularized estimator in* (3.2), a_{nk} *satisfies* $\limsup_n a_{nk} n^{-\lambda} = 0$ *for all* $k = k_1, \ldots, k_r$, *where* $\lambda = \frac{(k^*-1)^2}{2k^*(2k^*-1)}$ *and* $k^* := \max\{k, K - k + 1\}$. *Then the estimator* $\widehat{F}_{nk}(\tau)$ *satisfies*

$$\sqrt{n}(|\psi'_k|)^{-1}(\widehat{F}_{nk} - \tilde{\psi}_k(G_k)) \rightsquigarrow \mathbb{G}_{F_0},$$

a Gaussian process with covariance kernel $H(v_1, v_2) = G_K(v_1; F_0)(1 - G_K(v_2; F_0))$ *for* $v_1 \leq v_2$. *Furthermore, the estimator* $\widehat{F}_n$ *can be expressed as*

$$\sqrt{n}\widehat{S}_n(v)^{-1/2}(\widehat{F}_{nk}(v) - \tilde{\psi}_k(G_k(v))) \overset{d}{\to} N(0, 1)$$

uniformly in $v \in \mathcal{V}$.

Note in particular that the rate on a_{nk} implies that the "pasting points" implicitly defined in (3.1), $\tau^*_{1k}(a_{nk})$ and $\tau^*_{2k}(a_{nk})$, converge to zero at a rate that is slower than needed to achieve the optimal rate of convergence for the estimator derived in Section 2. In particular, the regularization bias in the trimmed regions of the support is of a larger order than the sampling variation of the estimator. Also, the rate on a_{nk} is slower for small values of k, which is a consequence of the relative rates at which the slope and the curvature of $\psi_k(\cdot)$ diverge as we approach the critical points of the mapping.

In a given sample, the choice a_{nk} has to trade off the regularization bias (for a_{nk} too small) and the error in the distributional approximation by a Gaussian process (from a_{nk} diverging too fast). As for most other regularization problems in econometrics, the sufficient condition on the asymptotic rate for the regularization parameter a_{nk} obviously does not imply a particular value for a finite sample, but a prescription has to balance the effects of trimming the support against improved coverage. The resolution of this trade-off will, in general, vary a lot across different applications, and we leave the construction of specific criteria for choosing the regularization parameter for future research.

The results in Theorem 3.1 can also be used directly to approximate the distribution of linear functionals of the valuation distribution $F_0(v)$, which will be discussed in more detail in the next section. However, a distribution theory for bounds on the private value distributions also has to take into account that the mapping $H^*(\cdot)$ is not differentiable;

a theory for intersection bounds based on regular estimators for individual bounds has been developed by Chernozhukov, Lee, and Rosen (forthcoming), but we leave the problem of incorporating trimmed or smoothed versions of irregular estimators into that framework for future research.

4. FUNCTIONALS OF THE VALUATION DISTRIBUTION

In empirical research on auctions, the distribution of valuations is often only of derived interest, but the researcher may want to use an estimator for F_0 to approximate other characteristics of the auction that can be characterized as functionals of the underlying distribution. In this section, we give bounds on the rate of convergence for estimators of general linear functionals of F_0 as well as expected revenue and the optimal reserve price for an auction of arbitrary size $\tilde{K}$.

4.1 *Linear functionals*

Consider linear functionals of the valuation distribution

$$T(F) := \int_0^\infty w(v) F(dv)$$

for a weighting function $w(v)$. We also define the weighting function in terms of quantiles of the valuation distribution,

$$\omega(\tau; F_0) := w(F_0^{-1}(\tau)).$$

ASSUMPTION 4.1. (i) *There are $\underline{\tau} \in [0, 1]$ and $\overline{\tau} \in [0, 1]$ such that $\omega(\tau; F)$ does not change sign on $[0, \underline{\tau}]$ or $[\overline{\tau}, 1]$. (ii) Furthermore, there exist constants β_1 and β_2 such that for all $F \in \mathcal{F}_0$, the behavior of $\omega(\tau; F)$ is described by*

$$\lim_{\tau \to 0} \tau^{-\beta_1} \omega'(\tau; F) < \infty \quad and \quad \lim_{\tau \to 1} (1 - \tau)^{-\beta_2} \omega'(\tau; F) < \infty.$$

We can also state this condition in terms of primitive assumptions on the p.d.f.: by the chain rule, $\omega'(\tau; F) = \frac{d}{d\tau} w(F^{-1}(\tau)) = \frac{w'(F^{-1}(\tau))}{h(\tau; F)}$, so that β_1 depends implicitly on the tail behavior of $h(\tau; F)$ given in Assumption 1.4.

EXAMPLE 4.1 (Allocative Efficiency). As a measure of the change in social surplus from the auction relative to assigning the object to a bidder at random, consider the expected difference between the auction winner's valuation of the object and that of the average bidder,

$$\Delta W(F) := \mathbb{E}[B_{iK}] - \mathbb{E}_F[V] = \int_0^\infty v \, dG_K(v) - \int_0^\infty v \, dF(v).$$

If we observe the winning bid B_{iK} for a number of IPV sealed-bid second-price auctions of size K, the first part of that difference can be estimated by the sample mean of observed bids, whereas the second part is a linear functional $T(F_0)$ of the private value distribution with weighting function $w(v) = v$. Hence, if F_0 satisfies Assumption 1.4 with tail parameters α_1 and α_2, then $T(F)$ satisfies Assumption 4.1 with $\beta_1 = -\alpha_1$ and $\beta_2 = -\alpha_2$.

Given the condition on the weighting function in Assumption 4.1, we can now give the following bound on the rate for linear functionals.

PROPOSITION 4.1. *Suppose Assumptions* 1.2, 1.3, *and* 4.1 *hold. Then*

$$r_n = n^{\min\{1/2,(2+p+\beta_1)/(k_1+p),(2+p+\beta_2)/(K-k_r+1+p)\}}$$

is an upper bound to the rate of convergence for estimating the linear functional

$$T(F) = \int_0^\infty w(v)F(dv) = \int_0^1 F^{-1}(s)\omega(s;F)\,ds.$$

It should be noted that this result is closely related to the bound on the rate for estimating F_0 with respect to the L_2 norm, which is essentially the relevant metric for a function space if we are interested in linear functionals. The role of the tail condition for the weighting function $w(v)$ in Assumption 4.1 is completely parallel to that on Assumption 1.4 for estimating the private value distribution in Theorem 2.3.

EXAMPLE 4.2. Consider again the problem in Example 4.1, where we observe the transaction price of n i.i.d. second-price auctions with K bidders and we are interested in estimating the expectation of V_k, $w(v) \equiv v$. Hence, $\omega'(\tau; F) = \frac{1}{h(\tau;F)}$. If, in addition, we assume that the support of V_k is bounded and the p.d.f. $f_0(v)$ is bounded away from zero, $\beta_1 = \beta_2 = 0$. Then by Proposition 4.1, a nonparametric estimator for the expectation of V_k can at best achieve the rate $r_n = n^{2/(K-1)}$ for any $K > 5$.

On the other hand, if we observe all K bids for each auctions, for example, as in the framework of Guerre, Perrigne, and Vuong (2000), we can estimate the expected valuation directly as the sample average of all bids across all auctions and, as expected, the bound for this scenario corresponds to a root-n rate.

4.2 *Expected revenue*

Next we are going to perform the following thought experiment: suppose we observe the k highest bid from n repeated sealed-bid second-price auctions of K bidders with independent private values, and based on these data, we want to predict expected revenue, that is, the expectation of the second highest bid, for an auction of the same format with $\tilde{K}$ bidders. Clearly if $\tilde{K} = K$ and $k = K - 1$, that is, we observe the second-highest bid for the observed auctions, then the sample average of observed bids is a root-n consistent estimator for expected revenue, even in the absence of any structural assumptions on the problem.

In all other cases, from our assumptions on the format of the auction and its equilibrium, the distribution of the transaction price is that of the $(\tilde{K} - 1)$st order statistic in a sample of $\tilde{K}$ i.i.d. draws, and we can, for example, use an estimator of the parent c.d.f. to approximate that distribution. Note that in contrast to the previous case, this type of extrapolation also relies crucially on our structural model, both for the observed and the counterfactual auction.

The following result gives the bound on the rate for nonparametric estimation of the expectation of the k highest out of $\tilde{K}$ bids based on observations of the $k_1, \ldots, k_r$ highest

bids out of K bidders:

PROPOSITION 4.2. *Suppose Assumptions* 1.2–1.4 *hold. Then*

$$r_n = n^{\min\{1/2,(k(1+p)+1-\alpha_1)/(k_1+p),((\tilde{K}-k)(1+p)+1-\alpha_2)/(K-k_r+1+p)\}}$$

is an upper bound to the rate of convergence for estimating the expectation of the k highest bid in a second-price auction of $\tilde{K}$ i.i.d. bidders.

It is interesting to note that in the case $k_1 = k_r = k$, this bound does not rule out the possibility that expected revenue can be estimated at a root-n rate unless $\tilde{K}$ is substantially smaller—less than half as large, to be precise—than K, even though from the previous proofs, these bounds appear to be sharp. However, it is important to point out that this result does not imply root-n estimability for expected revenue, even if $\tilde{K} > \frac{K-1}{2}$. In particular, a "naive" plug-in estimator of expected revenue using an untrimmed estimator for the parent distribution is likely not going to achieve that rate, though this remains to be shown formally.

4.3 *Optimal reserve price*

Suppose we observe the transaction price for n i.i.d. second-price IPV auctions with K risk-neutral bidders and we are interested in estimating the seller's optimal reserve price p^* to maximize the seller's surplus. By a standard result from auction theory (see, e.g., Riley and Samuelson (1981)), the seller's expected profit can be written as

$$\pi(p; F) = v_0 F^K(p) + K \int_p^\infty \left(vf(v) - \left(1 - F(v)\right)\right) F^{K-1}(v)\, dv,$$

where v_0 is the seller's valuation of the object.

Clearly $p^* > v_0$ for any distribution $F \in \mathcal{F}_0$, so that if $v_0 > \underline{v}$, then perturbations of the lower tail of the distribution do not affect the optimal reserve price. By Theorem 2.3, the estimator $\widehat{F}_n$ proposed in Section 2 converges to F_0 at the root-n rate uniformly in $v \in [v_0, \overline{v}]$, and since $\pi(p; F)$ is Lipschitz in $F(p)$, $\pi(p; \widehat{F}_n)$ is also root-n consistent for $\pi(p; F_0)$ uniformly in $v \in [v_0, \overline{v}]$.

We can now inspect the first-order conditions for a maximum of $\pi(p, F)$,

$$0 = \frac{d}{dp}\pi(p; F) = K(v_0 - p)F^{K-1}(p)f(p)K\left(1 - F(p)\right)F^{K-1}(p)$$

$$\Leftrightarrow \quad p = v_0 + \frac{1 - F(p)}{f(p)},$$

so that the optimal reserve price does not depend on the number of bidders in the "counterfactual" auction. It is now easy to verify that if $v_0 \le \underline{v}$ and for the class $\mathcal{F}_0$, there is no common upper bound on the density $f(v)$ in the lower tail of the support of V, for any $\tau \in [0, 1]$, we can find a distribution $\tilde{F} \in \mathcal{F}_0$ such that $\tilde{p}^* := \arg\max_p \pi(p; \tilde{F})$ is at or below the τ-quantile of that distribution, $\tilde{p}^* \le \tilde{F}^{-1}(\tau)$.

Given that distribution $\tilde{F}$, we can perturb the distribution below the τ quantile such that the corresponding optimal reserve price changes by at least $\frac{1}{4}\tau^{p+1}$. Now note that by Lemma C.1 in Appendix C, for a sample of n i.i.d. auctions of K bidders, the smallest quantile at which we can reliably detect such a perturbation is of the order $\tau_{1n} := \overline{\tau}n^{-\min\{1/2,1/(K-2+p)\}}$ for some constant $\overline{\tau} > 0$. Combining these two observations, we can obtain an upper bound on the rate of convergence for estimating the optimal reserve price, where it is also immediate that the rate cannot be faster than root-n and the possibility of perturbations on the upper tail does not impose further restrictions on the rate.

PROPOSITION 4.3. *Suppose Assumptions 1.2 and 1.3 hold with $p = 0$, and that the seller's valuation is $v_0 \leq \underline{v}$. Then without further restrictions on $\mathcal{F}_0$, $r_n = n^{\min\{1/2,1/k_1\}}$ is an upper bound on the rate of convergence for any nonparametric estimator of the optimal reserve price p^* for an auction of $\tilde{K}$ bidders from transaction price data from n i.i.d. auctions with K bidders. However, if $v_0 > \underline{v}$, then p^* can be estimated at a rate $r_n = n^{1/2}$.*

Note that the bound on the rate for the optimal reserve price implied by this proposition is always slower than the parametric rate if $K > 3$. Also, it is clear from the argument, that shape restrictions on the distributions in $\mathcal{F}_0$ can mitigate this problem, for example, if there is a (common) upper bound for the p.d.f. of v. Using the same argument, it is also possible to show that a risk-neutral participant in a first-price auction who has access to incomplete bidding data from past second-price auctions can estimate her equilibrium bid only at that same rate.

In one of their examples, Haile and Tamer (2003) imposed concavity on the seller's surplus as a function of the reserve price to show how to derive bounds on p^*. It is interesting to note that this is one special case for which the parametric rate is, in fact, attainable for the optimal reserve price: if $\pi(p, F)$ is concave in p for all $F \in \mathcal{F}_0$, and we have an estimator for $\widehat{\pi_n(p; F)}$ such that $\widehat{\pi_n(p; F)}$ is concave with probability 1 at all n and is root-n consistent for $\pi(p; F)$ at every $p \in \mathcal{V}$. By a slight modification of Theorem 10.8 in Rockafellar (1972), pointwise convergence of a concave function at rate $n^{1/2}$ implies uniform convergence at the $n^{1/2}$ rate, so that by Theorem 3.4.1 in van der Vaart and Wellner (1996), $\widehat{p}^* := \arg\min_{p \in \mathcal{V}} \widehat{\pi_n(p; F)}$ converges to p^* at the root-n rate. However, note that this argument does not work for estimators that do not impose concavity on $\widehat{\pi_n(p; F)}$ in a given sample; furthermore, it is generally difficult to justify such a restriction in terms of economic primitives for the auction problem.

5. FIRST-PRICE AND DESCENDING BID AUCTIONS

So far, all our results have been about the conceptually more straightforward case of second-price auctions. However, one class of settings for which the problem of incomplete bidding data is most salient are descending bid auctions. In this format, an auctioneer announces a descending sequence of prices, and the object is won by the first bidder who is willing to accept the current price. In particular, the remaining $K - 1$ potential buyers do not reveal their types, so that if bidding strategies are strictly monotone, only the bid that corresponds to the highest valuation is known to the econometrician.

Under the IPV assumption and if bidders are risk-neutral, this format is strategically equivalent to a sealed-bid first-price auction. In this last section, we show how some of our insights for the second-price format apply to first-price and strategically equivalent formats.

It is known from standard results in auction theory that given the valuation distribution F, the bidding strategy $b(v; F)$ in a symmetric Bayesian Nash equilibrium is characterized by the differential equation

$$b(v; F) := v - \frac{1}{K-1} \frac{F(v)b'(v; F)}{f(v)}. \tag{5.1}$$

That is, strategic bidders do not bid their actual valuations, but shade their bids by a factor that depends on the number of competing potential buyers and the private values distribution. More specifically, for risk-neutral bidders, it is optimal to bid the expectation of the next-highest bid conditional on winning the object. There is, in general, no closed-form expression for the optimal bid, and the shading factor depends on the unknown private value distribution. However, we show below that there is an easy two-step approach to obtain "pseudo-true" valuations, even when bidding data are incomplete.

We now replace the model for the second-price sealed-bid auction from Assumption 1.2 with a new assumption.

ASSUMPTION 5.1 (First-Price Auction). *Assumption* 1.2(i)–(iii) *holds and* (iv') *the auction is sealed-bid first-price or any other format that is strategically equivalent under the remaining assumptions, and participants play the symmetric Bayesian Nash equilibrium with bidding functions that satisfy* (5.1).

Now denote $g(b; F) := \frac{f(b^{-1}(b;F))}{b'(b;F)}$, the marginal distribution of bids with the corresponding c.d.f. $G(b; F) = F(b^{-1}(b; F))$, so that we can rewrite (5.1) as

$$b^{-1}(b; F) = b + \frac{1}{K-1} \frac{G(b; F)}{g(b; F)}.$$

We can now use this characterization of the inverse bidding function and the underlying valuation distribution to derive an upper bound on the convergence rate for nonparametric estimators as defined in (2.1) and (2.2).

PROPOSITION 5.1. *Let* $\widehat{F}_n$ *be an estimator for* F_0. *Then under Assumptions* 1.3, 1.4, *and* 5.1, $r_n = n^\lambda$ *is an upper bound on the rate of convergence under the sup-norm, where*

$$\lambda = \min\left\{ \frac{p}{2p+1}, \frac{1+p}{k_1+p}, \frac{p}{K-k_r+1+p} \right\}.$$

This rate result is not sharp and we conjecture that it can be strengthened slightly to $r_n^* = \min\{(\frac{n}{\log n})^{p/(2p+1)}, n^{(1+p)/(k_1+p)}, n^{p/(K-k_r+1+p)}\}$ using standard arguments on global convergence rates of nonparametric estimators; see Stone (1983). Note that if the convergence rate is determined in the tails of the distribution, this bound on the convergence rate is exactly the same as for second-price auctions, and the nonparametric

adjustment for the shading factor to account for strategic behavior only affects the over-all bound on the rate if the tails can otherwise be estimated with reasonable precision.

While establishing formally that the rate r_n^* is, in fact, achievable is beyond the scope of this paper, in the case in which only the highest bid is observed, it is possible to adapt the nonparametric plug-in approach from Guerre, Perrigne, and Vuong (2000) and obtain a distribution of estimated quasivaluations $b^{-1}(B_{iK}; F)$ of the highest bidders. We now give a brief explanation of how such a procedure can be designed.

Using the formulae for p.d.f.s and c.d.f.s of order statistics, one can verify that the ratio of the c.d.f. and the p.d.f. of the highest order statistic of bids equals $\frac{G_K^K(v;F)}{g_K^K(v;F)} = \frac{F(v)b'(v;F)}{f(v)}$ for all $v \in \mathcal{V}$. Hence, it is possible to express the inverse bidding function directly in terms of the p.d.f. $g_K(\cdot; F)$ and the c.d.f. $G_K^K(\cdot; F)$ of the *observed* bid:

$$b^{-1}(b_k; F) = b_k + \frac{1}{K-1} \frac{G(b_k; F)}{g(b_k; F)} = b_k + \frac{1}{K-1} \frac{G_K^K(b_k; F)}{g_K^K(b_k; F)}. \tag{5.2}$$

We can, therefore, estimate the sample distribution of the Kth order statistic of valuations across the n auctions by plugging nonparametric estimators for the density and the c.d.f. of the (observed) highest bids into this expression in a first step, and estimate the marginal c.d.f. of valuations $F(v)$ in a second step by inverting the distribution of the estimated quasivaluations.

6. Discussion

This paper establishes optimal rates for nonparametric estimation of the valuation distribution from incomplete bidding data in sealed-bid second-price auctions and strategic equivalents. If the econometrician only observes the highest bid or the transaction price, these rates may be very slow, even for auctions of a moderate size. These results suggest that a lot may be gained from combining different bids or data from auctions of different sizes. Also, the rates for bounds on the distribution are in many cases substantially slower than for estimating the distribution function in the point-identified case.

There are several cases in which the irregularity in estimating the tails of the distribution does not affect convergence rates for policy parameters: if the functional of interest only depends on the value of the distribution function at a point (e.g., the probability that the valuation of a random bidder exceeds the reserve price) or puts sufficiently low weight on the problematic tails (e.g., for projecting expected revenue for an auction with a sufficiently large number of bidders as in Proposition 4.2), root-n consistency may be preserved. But as a cautionary note, the calculations in Example 2.2 suggest that even in these cases, the irregularity may still lead to poor finite-sample performance of the resulting estimator.

The performance of nonparametric estimators could be enhanced significantly by imposing shape restrictions or a parametric structure for very low and/or high quantiles, depending on which bids are observed. Constraints of this type can generally be imposed by smoothing the empirical c.d.f. or two-step procedures (see, e.g., Aït-Sahalia and Duarte (2003) or Mammen and Thomas-Agnan (1999)), which can be solved at a

computational cost that is of the same order as that for the unconstrained problem. However, strategies for variance and bias reduction in this setting have to be different from the well known trade-offs in estimating probability densities in that smoothing only has to be local and depends on ordinal properties of the sample. However, deriving practical smoothing procedures that impose these shape restrictions and establishing their convergence rates is not straightforward and will be left for future research. But we conjecture that the bounds on the rate under smoothness restrictions $p \geq 1$ derived in Section 2 may, in fact, be sharp.

Finally, it should be noted that the difficulties in inverting distributions of order statistics to obtain the parent distribution also appear to apply to inference for other auction formats. A particularly relevant case is that of descending auctions in which, by construction, only the highest bidder reveals her type. Optimal rates for estimating first-price auctions when all bids are observed have been derived by Guerre, Perrigne, and Vuong (2000), but the behavior of nonparametric estimators with incomplete bidding data remains an open question.

Appendix A: Joint distribution of order statistics

The joint p.d.f. of the $(k_1, \ldots, k_r)$th order statistics is given by[6]

$$
\begin{aligned}
g_{k_1,\ldots,k_r}&(\mathbf{v}; F) \\
&= N(k_1, \ldots, k_r; K)\big[F(v_{k_1})\big]^{k_1-1} f(v_{k_1})\big[F(v_{k_2}) - F(v_{k_1})\big]^{k_2-k_1-1} f(v_{k_2}) \cdots \\
&\quad \times \big[1 - F(v_{k_r})\big]^{K-k_r} f(v_{k_r}),
\end{aligned}
$$

where $N(k_1, \ldots, k_r; K) = \frac{K!}{(k_1-1)!(k_2-k_1-1)!\cdots(K-k_r)!}$, $v_{k_1} \leq v_{k_2} \leq \cdots \leq v_{k_r}$. We can then obtain the joint c.d.f. by integrating the joint p.d.f. from the lower bound of the support, $(\underline{v}, \ldots, \underline{v})$ to $(v_{k_1}, \ldots, v_{k_r})$,

$$
\begin{aligned}
G_{k_1,\ldots,k_r}&(\mathbf{v}; F) \\
&= N(k_1, \ldots, k_r; K) \\
&\quad \times \int_{\underline{v}}^{v_{k_1}} \int_{\underline{v}}^{v_{k_2}} \cdots \int_{\underline{v}}^{v_{k_r}} \Big\{ \big[F(u_1)\big]^{k_1-1} f(u_1)\big[F(u_2) - F(u_1)\big]^{k_2-k_1-1} f(u_2) \cdots \\
&\quad \times \big[1 - F(u_r)\big]^{K-k_r} f(u_r) \Big\} \, d\mathbf{u} \\
&= (k_1, \ldots, k_r; K) \int_{\mathcal{I}^r(\mathbf{v})} s_1^{k_1-1} (s_2 - s_1)^{k_2-k_1-1} \cdots \\
&\quad \times (1 - s_r)^{K-k_r} \, ds_r \cdots ds_1,
\end{aligned}
\tag{A.1}
$$

where $\mathcal{I}^r(\mathbf{v}) := [0, F(v_{k_1})] \times [0, F(v_{k_2})] \times \cdots \times [0, F(v_{k_r})] \subset [0, 1]^r$ and the second expression follows from the change of variables formula.

---

[6]See, for example, David and Nagaraja (2003, p. 12).

APPENDIX B: PROOF OF CONSISTENCY RESULTS

PROOF OF THEOREM 2.1. The main problem we have to deal with in this proof is that the nonlinear transformation $\psi_k(\cdot)$ of the empirical c.d.f. is not asymptotically linear uniformly in $v \in \mathcal{V}$. We will, therefore, show uniform convergence separately in two data-dependent regions $\mathcal{C}_n$ and $\mathcal{V}/\mathcal{V}_n(\eta)$, and show that the probability that the union of those two regions covers all of $\mathcal{V}$ converges to 1 as n grows. In the following, we write $\widehat{G}_n(v) = \widehat{G}_{nk}(v)$ since k is taken to be fixed. We also use the notation $\lesssim$ for "smaller than up to a universal constant."

Fix $\eta > 0$ and let

$$\mathcal{V}_n(\eta) := \{v \in \mathcal{V} : |G_0(v) - \tau_s^*| \leq \eta n^{-1} \text{ for all } s = 1, \ldots, S\}.$$

Also define the data-dependent subset of as

$$\mathcal{C}_n := \{v \in \mathcal{V} : |\widehat{G}_n(v) - \tau_s^*| \leq n^{-1} \text{ for all } s = 1, \ldots, S\}.$$

For a given choice of $\eta > 1$, we define the event

$$\mathcal{A}_n(\eta) := \left\{ \sup_{v \in \mathcal{V}_n(\eta)} \frac{|G_0(v) - \tau_s^*|}{|\widehat{G}_n(v) - \tau_s^*|} \leq \eta \text{ for all } s = 1, \ldots, S \right\}.$$

Note that $\mathcal{A}_n(\eta) = \{\mathcal{V}_n(\eta) \subset \mathcal{C}_n\}$.

Also denote the event

$$\mathcal{B}_n(c) := \left\{ \sup_{v \in \mathcal{C}_n} |\psi(\widehat{G}_n(v)) - \psi(G_0(v))| > cr_n^{-1} \right\}.$$

We will now establish that (i) the limiting probability of $\mathcal{A}_n(\eta)$ can be made arbitrarily close to 1 by choosing η sufficiently large and that (ii) the probability of $\mathcal{B}_n(c)$ can be arbitrarily small for c large enough.

To show that $\lim_{\eta \to \infty} \lim_{n \to 0} \mathbb{P}(\mathcal{A}_n(\eta)) = 1$, consider the class $\mathcal{F}_s$ of functions

$$\mathcal{F}_{sn}(\eta) := \left\{ \frac{\mathbb{1}\{v \in (-\infty, v]\}}{|G_0(v) - \tau_s^*|} \bigg| v \in \mathcal{V}_n(\eta) \right\},$$

with the envelope function $F_{ns}(v; \eta) = \frac{\mathbb{1}\{v \in \mathcal{V}_n(\eta)\}}{|G_0(v) - \tau_s^*|}$.

We can bound the norm of the envelope function by

$$\begin{aligned}
\|F_{ns}(\eta)\|_{P,2}^2 &= \int_0^{\max\{0, \tau_s^* - \eta n^{-1}\}} \frac{1}{(t - \tau_s^*)^2}\, dt + \int_{\min\{1, \tau_s^* + \eta n^{-1}\}}^1 \frac{1}{(t - \tau_s^*)^2}\, dt \\
&= \frac{2}{\eta n^{-1}} - \min\left\{ \frac{1}{\tau_s^*}, \frac{1}{\eta n^{-1}} \right\} - \min\left\{ \frac{1}{1 - \tau_s^*}, \frac{1}{\eta n^{-1}} \right\} \leq \frac{2}{\eta n^{-1}}.
\end{aligned} \tag{B.1}$$

Using standard notation from empirical process theory (see, e.g., van der Vaart and Wellner (1996)), for a given value of $\varepsilon > 0$, we define the bracketing number $N_{[]}(\varepsilon, \mathcal{F}, \|\cdot\|)$ of a class of functions $\mathcal{F}$ as the smallest number of brackets $[l, u] := \{f : l(v) \leq f(v) \leq$

$u(v)$ for all $v \in \mathcal{V}\}$ with $\|u - l\| < \varepsilon$ with respect to a norm $\| \cdot \|$ needed to cover $\mathcal{F}$. Also let the entropy integral be

$$J_{[]}\big(\delta, \mathcal{F}, \| \cdot \|\big) := \int_0^\delta \sqrt{1 + \log N_{[]}\big(\varepsilon\|F\|, \mathcal{F}, \| \cdot \|\big)}\, d\varepsilon$$

for any $\delta > 0$.

For the class $\mathcal{F}_{ns}$, we can construct ε brackets of the form $[l, u]$ with $l(v) := \frac{\mathbb{1}\{v \leq v_U\}}{\min_{v \in \{v_L, v_U\}} |G_0(v) - \tau_s^*|}$ and $u(v) := \frac{\mathbb{1}\{v \leq v_L\}}{\min_{v \in \{v_L, v_U\}} |G_0(v) - \tau_s^*|}$, where $v_L < v_U$ satisfies $|\psi'(v_L)|^2 \times (G_0(v_U) - G_0(v_L)) < \varepsilon^2$. We first show that for fixed ε, the bracketing number $N_{[]}(\varepsilon\|F_{sn}\|_{P,2}, \mathcal{F}_{sn}, \| \cdot \|_{P,2})$ is uniformly bounded in n, where $\|f\|_{P,2} := (\int f^2\, dP)^{1/2}$ denotes the $L_2(P)$ norm of f.

For notational simplicity, consider only the case $s = 1$ with $\tau_1^* = 0$. Then the lowest $\varepsilon\|F_{sn}\|$ bracket can be chosen as described above with $v_{L1} = G_0^{-1}(\eta n^{-1})$ and some $v_{U1} \geq \varepsilon G_0(v_{L1})\|F_{1n}\|_{P,2} = \varepsilon n^{1/2 - 1/2} = \varepsilon$. Hence, the upper bound for the next-higher bracket does not decrease in n. Hence, we can bound the bracketing number by $N_{[]}(\varepsilon\|F_{sn}\|_{P,2}, \mathcal{F}_{sn}, \| \cdot \|_{P,2}) \leq 1 + \frac{1}{\varepsilon}$, so that

$$J_{[]}\big(1, \mathcal{F}_{sn}, L_2(P)\big) = \int_0^1 \sqrt{1 + \log N_{[]}\big(\varepsilon\|F_{sn}\|_{P,2}, \mathcal{F}_{sn}, \| \cdot \|_{P,2}\big)}\, d\varepsilon$$

$$\leq \int_0^1 \sqrt{1 + \log(2) - \log(\varepsilon)}\, d\varepsilon < \infty,$$

where the finite upper bound does not depend on $s = 1, \dots, S$ or $n = 1, 2, \dots$.

Since by Assumption 1.3, $\mathbf{B}_1, \dots, \mathbf{B}_n$ are i.i.d., we can use Theorem 2.14.2 in van der Vaart and Wellner (1996) to bound

$$\mathbb{E}^* \sup_{v \in \mathcal{V}_n(\eta)} \left| \frac{\widehat{G}_n(v) - G_0(v)}{G_0(v) - \tau_s^*} \right| \lesssim n^{-1/2} J_{[]}\big(1, \mathcal{F}_{sn}, L_2(P)\big)\|F_{sn}\|_{P,2}, \tag{B.2}$$

where $\mathbb{E}^* X$ denotes the outer expectation of X.

Since $J_{[]}(1, \mathcal{F}_{sn}, L_2(P))$ is finite, for any $\eta > 1$, we can use Markov's inequality to bound

$$P\big(\mathcal{A}_n^C(\eta)\big)$$

$$= 1 - P\left(\sup_{v \in \mathcal{V}_n(\eta)} \frac{|G_0(v) - \tau_s^*|}{|\widehat{G}_n(v) - \tau_s^*|} \leq \eta \text{ for all } s = 1, \dots, S \right)$$

$$\leq \sum_{s=1}^S P\left(\inf_{v \in \mathcal{V}_n(\eta)} \frac{|\widehat{G}_n(v) - G_0(v) + G_0(v) - \tau_s^*|}{|G_0(v) - \tau_s^*|} \leq \frac{1}{\eta} \right)$$

$$\leq \sum_{s=1}^S \left\{ P\left(\inf_{v \in \mathcal{V}_n(\eta)} \frac{|\widehat{G}_n(v) - G_0(v)|}{|G_0(v) - \tau_s^*|} \leq \frac{1}{\eta} - 1 \right) \right. \tag{B.3}$$

$$\left. + P\big(\text{sign}(\widehat{G}_n(v) - \tau_s^*) \neq \text{sign}(G_0(v) - \tau_s^*) \text{ for some } v \in \mathcal{V}_n(\eta)\big) \right\}$$

$$\leq 2 \sum_{s=1}^{S} P\left(\inf_{v \in \mathcal{V}_n(\eta)} \frac{|\widehat{G}_n(v) - G_0(v)|}{|G_0(v) - \tau_s^*|} \leq -\frac{\eta - 1}{\eta} \right)$$

$$\lesssim \frac{2SJ_{[]}(1, \mathcal{F}_{sn}, L_2(P))}{(\eta - 1)},$$

where the last step uses Markov's inequality together with (B.2) and (B.1). This bound on the probability can be made arbitrarily small by choosing a sufficiently large value of $\eta > 1$.

Next, we bound the probability of $\mathcal{B}_n(c)$. First note that since $\psi(\tau)$ is differentiable, at every $v \in \mathcal{V}$, a mean-value expansion gives

$$\psi(\widehat{G}_n(v)) - \psi(G_0(v)) = \psi'(\overline{G}_n(v))(\widehat{G}_n(v) - G_0(v)), \tag{B.4}$$

where $\overline{G}_n(v)$ is an intermediate value between $\widehat{G}_n(v)$ and $G_0(v)$. In this approximation, the term $\psi'(\overline{G}_n(v))$ is not guaranteed to be bounded, but $\overline{G}_n(v)$ may be arbitrarily close to τ_s^* for some $s = 1, \ldots, S$ with positive probability.

However, by monotonicity of $\psi'(\tau)$ between the critical points τ_s^*, $s = 1, \ldots, S$, we can bound

$$\left| \psi'(\overline{G}_n(v)) \right| \leq \max\left\{ \left| \psi'(\widehat{G}_n(v)) \right|, \left| \psi'(G_0(v)) \right| \right\} \leq \left| \psi'(\eta^{-1} G_0(v)) \right| \tag{B.5}$$

for all values $v \in \mathcal{C}_n$.

Now define the class of functions

$$\mathcal{H}_n := \left\{ \psi'(G_0(t)) \mathbb{1}\{v \leq t\} | t \in \mathcal{C}_n \right\},$$

with envelope function $H_n(v) := |\psi'(G_0(v))|$.

Noting that for any exponent $\delta_s > 0$ in Condition 2.1, $|\psi'(\tau)|$ is dominated by $\frac{1}{\tau - \tau_s^*}$ for values of τ close to τ^*, we can use the same reasoning as before to establish that the bracketing integral $J_{[]}(1, \mathcal{H}_n, L_2(P))$ is bounded. To bound the norm of the corresponding envelope functions, let $\delta_s < 1$ without loss of generality. Then for n sufficiently large, by Condition 2.1, we can bound

$$\int_{\tau_s^* + n^{-1}}^{(\tau_s^* + \tau_{s+1}^*)/2} |\psi'(t)|^2 \, dt \leq 2A_s \int_{\tau_s^* + n^{-1}}^{(\tau_s^* + \tau_{s+1}^*)/2} |t - \tau_s^*|^{2\delta_s - 2} \, dt \leq 2A_s n^{1 - 2\delta_s}.$$

We can now use (B.4), (B.5), and Theorem 2.14.2 in van der Vaart and Wellner (1996) to bound

$$\mathbb{E}^*\left[\sup_{v \in \mathcal{C}_n} |\psi(\widehat{G}_n(v)) - \psi(G_0(v))| \right] \leq \mathbb{E}^* \sup_{v \in \mathcal{C}_n} |\psi'(G_0(v))| |\widehat{G}_n(v) - G_0(v)|$$

$$\leq n^{-1/2} J_{[]}(1, \mathcal{H}_n, L_2(P)) \|H_n\|_{P,2} \tag{B.6}$$

$$\leq 2A_s J_{[]}(1, \mathcal{F}_{sn}, L_2(P)) n^{-\delta_s}$$

for all $s = 1, \ldots, S$ and n large enough.

Hence, denoting the complement of the event $\mathcal{A}_n(\eta)$ by $\mathcal{A}_n^C(\eta)$, we can use Markov's inequality together with the law of total probability and (B.3) to obtain

$$\mathbb{P}\left(\sup_{v\in\mathcal{V}_n(\eta)}\left|\psi(\widehat{G}_n(v))-\psi(G_0(v))\right|>cr_n\right)$$

$$\leq\mathbb{P}(\mathcal{B}_n(c))+\mathbb{P}(\mathcal{A}_n^C(\eta))$$

$$\leq\max_{s=1,\dots,S}\left\{\frac{2A_sJ_{[]}(1,\mathcal{H}_{sn},L_2(P))}{c}\right\}+\frac{2\sum_{s=1}^{S}J_{[]}(1,\mathcal{F}_{sn},L_2(P))}{(\eta-1)},$$

which can be made arbitrarily small by choosing η and c large enough.

Furthermore, from Condition 2.1, it follows that

$$\sup_{v\in\mathcal{V}\setminus\mathcal{V}_n(\eta)}\min_{\tau\in\{\tau_1^*,\dots,\tau_S^*\}}\left|\tau-G_0(v)\right|\leq(\eta n)^{-\delta}. \tag{B.7}$$

Since by Condition 2.1, $\psi(\tau)$ is monotone on the intervals $[\tau_s^*-\eta n^{-1},\tau_s^*]$ and $[\tau_s^*,\tau_s^*+\eta n^{-1}]$ for every η, by (B.7) and (B.7), there exists $c(\eta)<\infty$ such that conditional on the event $\mathcal{A}_n(\eta)\cap\mathcal{B}_n^C(c)$,

$$\sup_{v\in\mathcal{V}\setminus\mathcal{V}_n(\eta)}\left|\psi(\widehat{G}_n(v))-\psi(G_0(v))\right|<2cn^{-\delta}$$

with probability approaching 1, which completes the proof.   □

PROOF OF PROPOSITION 2.1. By Assumption 1.1, the distribution of the observable bid B_{ik} is given by $G_k(v)$. From the previous discussion, it is straightforward to verify that Condition 2.1 holds for the mapping $\psi_k(\cdot)$ with $\tau_1^*=0$, $\tau_2^*=1$, $\delta_1=\frac{1}{k}$, and $\delta_2=\frac{1}{K-k+1}$, where $A_1=\frac{1}{N(k,K)^{1/k}}$ and $A_2=\frac{1}{N(k,K)^{1/(K-k+1)}}$. Hence, we can apply Theorem 2.1 with $\delta=\min\{\frac{1}{k},\frac{1}{K-k+1}\}$, so that the estimator in (2.4) is uniformly consistent with rate $r_n=n^\delta$.   □

PROOF OF THEOREM 2.2. Suppose n is large enough that $b_{nks}<r$ for all $s=1,\dots,S$ and $k=k_1,\dots,k_r$, and fix $\eta>0$. For each $k=k_1,\dots,k_r$, we can define the sets

$$\mathcal{V}_{nk}(\eta):=\left\{v\in\mathcal{V}:\left|G_k(v)-\tau_s^*\right|\leq\eta b_{nks}\text{ for all }s=1,\dots,S\right\}$$

and the trimmed support of the estimate from the kth bid

$$\mathcal{C}_{nk}:=\left\{v\in\mathcal{V}:\left|\widehat{G}_{nk}(v)-\tau_s^*\right|\geq b_{nks}\text{ for all }s=1,\dots,S\right\},$$

and denote the event that $\mathcal{V}_{nk}(\eta)\subset\mathcal{C}_{nk}$ by

$$\mathcal{A}_{nk}(\eta):=\left\{\sup_{v\in\mathcal{V}_{nk}(\eta)}\frac{|G_k(v)-\tau_s^*|}{|\widehat{G}_{nk}(v)-\tau_s^*|}\leq\eta\text{ for all }s=1,\dots,S\right\}.$$

Also let $r_{nk}=n^{\min\{\delta_{k1},\dots,\delta_{kS}\}}$ for every $k=k_1,\dots,k_r$ and denote the event

$$\mathcal{B}_{nk}(c):=\left\{\sup_{v\in\mathcal{C}_{nk}}\left|\psi(\widehat{G}_{nk}(v))-\psi(G_k(v))\right|>cr_{nk}^{-1}\right\}$$

in an analogous fashion as in the proof of Theorem 2.1.

We prove the result by bounding the sup of the estimation error separately on the sets $\mathcal{C}_{nk}(\eta)$ and $\mathcal{V}/\mathcal{V}_{nk}$, and then establishing that the union of the two sets equals $\mathcal{V}$ with probability approaching 1. More specifically, we need to show that the probabilities (i) $P(\mathcal{B}_{nk}(c))$ and (ii) $P((\mathcal{V}/\mathcal{V}_{nk}(\eta)) \cap \mathcal{C}_{nk}(\eta) = \emptyset)$ can be made arbitrarily small by choosing η large enough, and that (iii) the approximation error from trimming the estimate on the set $\mathcal{V}/\mathcal{V}_{nk}(\eta)$ is bounded by cr_n.

For part (i) of this argument, note that on $\mathcal{C}_{nk}$, $|\psi'_k(\widehat{G}_{nk}(v))| \leq b_{nks}^{\delta_{ks}}$ almost surely for all k and s, so that similar to the argument for the previous result, we can find the upper bound

$$\mathbb{P}(\mathcal{B}_{nk}(c)) \leq \frac{2\max\limits_{s=1,\ldots,S}\{A_s J_{[]}(1, \mathcal{F}_{sn}, L_2(P)))b_{nks}\}}{cn},$$

where $\mathcal{F}_{ns}$ and the bracketing integral $J_{[]}(\cdot)$ are the same as in the proof of Theorem 2.1. Note that by assumption, $\lim n^{-1}b_{nks} < \infty$, so that the expression on the right-hand side is bounded by a constant that does not depend on n and can be made arbitrarily small by choosing c large enough.

For (ii), note that

$$1 - \mathbb{P}(\mathcal{V}_{nk}(\eta) \subset \mathcal{C}_{nk}(\eta)) \leq \sum_{k=k_1,\ldots,k_r} \mathbb{P}(\mathcal{A}_{nk}(\eta)) \leq \frac{2\sum\limits_{s=1}^{S} J_{[]}(1, \mathcal{F}_{sn}, L_2(P))}{\eta - 1},$$

where the second inequality follows from the same arguments used to bound the probability for the event $\mathcal{A}_n(\eta)$ in the proof of Theorem 2.1.

Since the set $k_1, \ldots, k_r$ is finite, we can use the bounds from inverting the distribution of each order statistic to bound the sup of the aggregate with trimming, $\sup_{v \in \mathcal{V}} |\widehat{F}_n(v) - F_0(v)|$, so that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{v \in \mathcal{V}}|\widehat{F}_n(v) - F_0(v)| > cr_n\right) \\
&\leq \mathbb{P}\left(\bigcup_k \mathcal{B}_{nk}(c)\right) + \mathbb{P}\left(\bigcup_k \mathcal{A}_{nk}^C(\eta)\right) \\
&\leq \sum_k \mathbb{P}(\mathcal{B}_{nk}(c)) + \sum_k \mathbb{P}(\mathcal{A}_{nk}(\eta)) \\
&\leq \sum_k \left\{\frac{2\max\limits_{s=1,\ldots,S}\{A_s J_{[]}(1, \mathcal{F}_{sn}, L_2(P)))b_{nks}\}}{cn} + \frac{2\sum\limits_{s=1}^{S} J_{[]}(1, \mathcal{F}_{sn}, L_2(P))}{\eta - 1}\right\}.
\end{aligned} \tag{B.8}$$

By the same argument as in the proof of the main theorem, each of the summands in this bound can be made arbitrarily small by choosing η and c large enough.

It remains to be shown that we can bound the error from trimming the estimates outside the set $\mathcal{V}_{nk}(\eta)$ on the estimator $\widehat{H}_n$: For part (a) of Theorem 2.2, we have $r_n =$

$\max\{r_{nk_1}, \ldots, r_{nk_r}\}$, so that by our assumptions on the function $H(\cdot)$,

$$
\sup_{v \in \mathcal{V}} \Big| H\big(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v)), \mathbb{1}_{v \in \mathcal{V}_{nk_1}(\eta)}, \ldots, \mathbb{1}_{v \in \mathcal{V}_{nk_r}(\eta)} \big)
$$

$$
- h\big(F_0(v)\big) \Big| \tag{B.9}
$$

$$
\leq \min_{k=k_1, \ldots, k_r} \sup_{v \in \mathcal{V}/\mathcal{V}_{nk}(\eta)} \min_{s=1, \ldots, S} \big| \tilde{H}\big(\psi_k(\tau_s^*)\big) - \tilde{H}\big(F_0(v)\big) \big| \leq M_H \eta^{\delta^*} r_n^{-1}.
$$

For part (b), $\tilde{r}_n = \max\{r_{nk_1}, \ldots, r_{nk_r}\}$ and we choose any $\tau_{ns_q}^*(v)$ in $\{\tau_1^*, \ldots, \tau_S^*\}$ that satisfies $|G_{k_q}(v) - \tau_{ns_q}^*(v)| \leq \eta b_{nks}$. Note that such a value exists for any $v \in \mathcal{V}/\mathcal{V}_{nk_q}(\eta)$ from the definition of $\mathcal{V}_{nk}(\eta)$. Then we have

$$
\sup_{v \in \mathcal{V}} \Big| H\big(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v)), \mathbb{1}_{v \in \mathcal{V}_{nk_1}(\eta)}, \ldots, \mathbb{1}_{v \in \mathcal{V}_{nk_r}(\eta)} \big)
$$

$$
- H^*\big(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v))\big) \Big|
$$

$$
\leq \max_{q=1, \ldots, r} \sup_{v \in \mathcal{V}/\mathcal{V}_{nk_q}(\eta)} \big| H^*\big(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_q}(\tau_{ns_q}^*(v)), \ldots, \psi_{k_r}(G_{k_r}(v))\big) \tag{B.10}
$$

$$
- H^*\big(\psi_{k_1}(G_{k_1}(v)), \ldots, \psi_{k_r}(G_{k_r}(v))\big) \big|
$$

$$
\leq M_H \eta^{\delta^*} r_n^{-1}.
$$

Hence the convergence rates in parts (a) and (b) of the theorem follow from (B.9), and (B.10), respectively, together with the bound in (B.8). □

Appendix C: Proofs for upper bounds on the rate of convergence

We use the notation $\lesssim$ for "smaller than up to a universal constant." Choose some $\tau_0 \in (0, 1)$ and let $\alpha_n = an^{-1/2}$ for some positive $a < \min\{\tau_0, 1 - \tau_0\}$, and let $\tau_{1n} = \tau_0 n^{-1/(k_1+p)}$ and $\tau_{2n} = \tau_0 n^{-1/(K-k_r+p)}$ be sequences of numbers between 0 and 1 that converge to zero.

Let $\xi(t)$ be a nonnegative function with support $[0, \frac{1}{2}]$ with $\int_0^{1/2} \xi(v)\, dv = \frac{1}{2}$, such that for some positive constant $B < \infty$, $\sup_{t \in \mathbb{R}} \xi(t) < B$, and the first p derivatives of $\xi(\cdot)$ are bounded in absolute value by B uniformly in t. To obtain an upper bound to the rate of convergence, we consider a sequence of distributions F_{0n} with corresponding p.d.f.s f_{0n} and perturbations of the true p.d.f. that are of the form

$$
f_{jn}(v) := f_{0n}(v)\big[1 + \xi_{jn}(v)\big] \tag{C.1}
$$

for $j = 1, 2, 3$, where we define

$$
\xi_{1n}(v) := \tau_{1n}^p \left(\xi\left(\tau_{1n}^{-1} F_{0n}^{-1}\left(\frac{\tau_{1n}}{2} - F_{0n}(v) \right) \right) \right.
$$

$$
\left. - \beta_{1n} \xi\left(\tau_{1n}^{-1} F_{0n}^{-1}\left(F_{0n}(v) - \frac{\tau_{1n}}{2} \right) \right) \right), \tag{C.2}
$$

$$\xi_{2n}(v) := \tau_{2n}^{p}\left(\xi\left(\tau_{2n}^{-1}F_{0n}^{-1}\left(F_{0n}(v) - 1 + \frac{\tau_{2n}}{2}\right)\right)\right.$$

$$\left. - \beta_{2n}\xi\left(\tau_{2n}^{-1}F_{0n}^{-1}\left(1 - \frac{\tau_{2n}}{2} - F_{0n}(v)\right)\right)\right),$$

$$\xi_{3n}(v; \alpha) := \alpha\left(\xi(F_{0n}(v) - \tau_0) - \beta_{3n}\xi(\tau_0 - F_{0n}(v))\right),$$

and the sequences β_{jn} are chosen in a way such that $\int \xi_{jn}(v)f_{0n}(v)\,dv = 0$. Note that from the choice of B, it follows that $\frac{1}{B} \leq \beta_{jn} \leq B$ for all n, so that $f_{jn}(v) = f_{0n}(v)[1 + \xi_{jn}(v)]$ is indeed a proper density. Also, the normalization by τ_{jn}^{p} ensures that the first p derivatives of $f_{jn}(v)$ are uniformly bounded.

Consider a nonnegative mapping $\varrho : \mathcal{F}_0 \times \mathcal{F}_0 \to \mathbb{R}_+$, the nonnegative real numbers such that $\varrho(F, F) = 0$ for any $F \in \mathcal{F}_0$. For most purposes of this paper, $\varrho(F, G)$ can be take to be a semimetric on the space $\mathcal{F}_0$, but we are not going to require the mapping to be symmetric in its arguments, which is important when we analyze the rate of convergence of functionals of the valuation distribution.[7]

LEMMA C.1. *Consider perturbations F_{1n} and F_{2n} of the sequence of c.d.f.s $F_{0n}(v)$ that are of the form as in* (C.1). *Suppose that for constants $\gamma_1, \gamma_2, \gamma_3 > 0$, $\varrho(F_{1n}, F_{0n}) \gtrsim \tau_n^{\gamma_1}$, $\varrho(F_{2n}, F_{0n}) \gtrsim \tau_n^{\gamma_2}$, and $\varrho(F_{3n}, F_{0n}) \gtrsim \alpha_n^{\gamma_3}$ for all $\tau_0 \in (0, 1)$ and $\delta < \tau_0$. Then under Assumptions* 1.2 *and* 1.3,*

$$\limsup_{n} P\left(\varrho(\widehat{F}_n, F_{0n}) > cn^{-\min\{\gamma_3/2, \gamma_1/(k_1+p), \gamma_2/(K-k_r+p)\}}\right) > 0$$

and

$$\lim_{c \to 0} \limsup_{n} P\left(\varrho(\widehat{F}_n, F_{0n}) > cn^{-\min\{\gamma_3/2, \gamma_1/(k_1+p), \gamma_2/(K-k_r+p)\}}\right) = 1.$$

PROOF. Without loss of generality, suppose the sequence $F_{0n} \equiv F_0(v)$ is constant at one fixed element $F_0 \in \mathcal{F}_0$. Consider local alternatives of the form $f_n(v) = f_{1n}(v)$ as defined in (C.1). The c.d.f. $F_{1n}(v)$ that corresponds to $f_{1n}(v)$ is given by

$$F_{1n}(v) := \int_0^v f_0(s)\left(1 + \xi_{1n}(s)\right)ds,$$

which is equal to $F_0(v)$ for all $v > F_0^{-1}(\tau_{1n})$.

To construct the likelihood ratio, note that

$$\frac{f_{1n}(v)}{f_0(v)} = \begin{cases} 1 + \tau_{1n}^{p}\xi\left(\tau_{1n}^{-1}F_0^{-1}\left(\frac{\tau_{1n}}{2} - F_0(v)\right)\right) & \text{if } 0 \leq F_0(v) < \frac{\tau_{1n}}{2}, \\ 1 - \beta_{1n}\tau_{1n}^{p}\xi\left(\tau_{1n}^{-1}F_0^{-1}\left(F_0(v) - \frac{\tau_{1n}}{2}\right)\right) & \text{if } \frac{\tau_{1n}}{2} \leq F_0(v) < \tau_{1n}, \\ 1 & \text{otherwise.} \end{cases}$$

---

[7]A semimetric on a space X $\varrho(x, y)$ is a map $\varrho : X \times X \to [0, \infty)$ such that for any $x, y, z \in X$, (i) $\varrho(x, y) = \varrho(y, x)$ and (ii) $\varrho(x, z) \leq \varrho(x, y) + \varrho(y, z)$.

Also note that for any pair of valuations $v_{k_t} > v_{k_s}$,

$$1 - B\tau_{1n}^p \leq \frac{F_{1n}(v_{k_t}) - F_{1n}(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})} \leq 1 + B\tau_{1n}^p. \tag{C.3}$$

To avoid an additional case distinction, we define $k_0 := 0$, $k_{r+1} := K + 1$, $B_{i0} := \inf \mathcal{V}$, and $B_{ir+1} := \sup \mathcal{V}$. Note that this is without loss of generality even if the support of V is not bounded, since the likelihood ratio only depends on the realizations of V through the c.d.f. $F_0(v)$, where $F_0(B_{i0}) = 0$ and $F_0(B_{ir+1}) = 1$.

Now by Assumptions 1.2 and 1.3, the likelihood ratio for an observation $\mathbf{B}_i := (B_{ik_1}, \ldots, B_{ik_r})$ is given by the Radon–Nikodym derivative

$$
\begin{aligned}
L_{1n}(B_{ik_1}, \ldots, B_{ik_r}) &= \frac{dG(B_{ik_1}, \ldots, B_{ik_r}; F_{1n})}{dG(B_{ik_1}, \ldots, B_{ik_r}; F_0)} \\
&= \left(\frac{F_{1n}(B_{ik_1})}{F_0(B_{ik_1})}\right)^{k_1 - 1} \left(\frac{F_{1n}(B_{ik_2}) - F_{1n}(B_{ik_1})}{F_0(B_{ik_2}) - F_0(B_{ik_1})}\right)^{k_2 - k_1 - 1} \cdots \\
&\quad \times \left(\frac{1 - F_{1n}(B_{ik_r})}{1 - F_0(B_{ik_r})}\right)^{K - k_r} \prod_{s=1}^{r} \frac{f_{1n}(B_{ik_s})}{f_0(B_{ik_s})} \\
&= \prod_{s=0}^{r} \left(\frac{F_{1n}(B_{ik_{s+1}}) - F_{1n}(B_{ik_s})}{F_0(B_{ik_{s+1}}) - F_0(B_{ik_s})}\right)^{k_{s+1} - k_s - 1} \prod_{s=1}^{r} \frac{f_{1n}(B_{ik_s})}{f_0(B_{ik_s})}.
\end{aligned}
$$

Since by Assumption 1.3, the observations $\mathbf{B}_1, \ldots, \mathbf{B}_n$ are i.i.d. across auctions, we can write the likelihood ratio of the sample between the models f_0 and f_{1n} as

$$\mathcal{L}_{1n} := \prod_{i=1}^{n} L_{1n}(B_{ik_1}, \ldots, B_{ik_r}).$$

Also define random variables $\chi_{is1} := \mathbb{1}\{F_0(B_{ik_s}) < \frac{\tau_{1n}}{2}\}$ and $\chi_{is2} := \mathbb{1}\{\frac{\tau_{1n}}{2} \leq F_0(B_{ik_s}) < \tau_n\}$ for $s = 1, \ldots, r$, and set $\chi_{i01} := \chi_{i11}$ and $\chi_{i02} = \chi_{ir+12} = \chi_{ir+12} = 0$. Taking logs, we obtain

$$
\begin{aligned}
l_{1n}(B_{ik_1}, \ldots, B_{ik_r}) &:= \log\big(L_{1n}(B_{ik_1}, \ldots, B_{ik_r})\big) \\
&= \sum_{s=1}^{r} \Bigg\{ \chi_{is1} \log\Bigg(1 + \tau_{1n}^p \xi\bigg(\tau_{1n}^{-1} F_0^{-1}\Big(\frac{\tau_{1n}}{2} - F_0(v)\Big)\bigg)\Bigg) \\
&\quad + \chi_{is2} \log\Bigg(1 - \beta_{1n} \tau_{1n}^p \xi\bigg(\tau_{1n}^{-1} F_0^{-1}\Big(F_0(v) - \frac{\tau_{1n}}{2}\Big)\bigg)\Bigg)\Bigg\} \\
&\quad + \sum_{s=0}^{r} (k_{s+1} - k_s - 1)(\chi_{is1} + \chi_{is2}) \log\left(\frac{F_{1n}(v_{k_t}) - F_{1n}(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})}\right) \\
&= \sum_{s=1}^{r} \Bigg\{ \chi_{is1} \tau_{1n}^p \xi\bigg(\tau_{1n}^{-1} F_0^{-1}\Big(\frac{\tau_{1n}}{2} - F_0(v)\Big)\bigg) \\
&\quad - \chi_{is2} \beta_{1n} \tau_{1n}^p \xi\bigg(\tau_{1n}^{-1} F_0^{-1}\Big(F_0(v) - \frac{\tau_{1n}}{2}\Big)\bigg)\Bigg\}
\end{aligned}
$$

$$+ \sum_{s=0}^{r}(k_{s+1} - k_s - 1)(\chi_{is1} + \chi_{is2})\log\left(\frac{F_{1n}(v_{k_t}) - F_{1n}(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})}\right)$$

$$+ o_P(\tau_{1n}^p \max\{1, \beta_{1n}\})$$

from a Taylor expansion of the log around 1.

From (C.3), we can see that if $\chi_{is1} = \chi_{is2} = 0$ for all $s = 1, \ldots, r$, then $l_{1n}(B_{ik_1}, \ldots,$ $B_{ik_r}) = 0$, so that the ith observation only contributes to the likelihood ratio if $B_{i1} < \tau_{1n}$. Also by inspection, we can bound $|\log(\frac{F_{1n}(v_{k_t}) - F_{1n}(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})})| \leq B\tau_{1n}^p$ for $v_{k_t} > v_{k_s}$ and any $\tau_n \in [0, 1]$. Hence, for any realization of $(B_{ik_1}, \ldots, B_{ik_r})$, $|l_{1n}(B_{ik_1}, \ldots, B_{ik_r})| \leq (K + 1)B\tau_{1n}^p$.

Since the log likelihood depends on the realization of $(B_{ik_1}, \ldots, B_{ik_r})$ only through the marginal quantile of each component, it follows from a change of variables under the integral that its expectation is given by

$$\mathbb{E}_{F_0}\left|\sum_{i=1}^{n} l_{1n}(B_{ik_1}, \ldots, B_{ik_r})\right|$$

$$= \int_{(\mathcal{V}^r)^n}\left|\sum_{i=1}^{n} l_{1n}(B_{ik_1}, \ldots, B_{ik_r})\right| \bigotimes_{i=1}^{n} dG(B_{ik_1}, \ldots, B_{ik_r}; F_0)$$

$$\leq n(K + 1)B\tau_{1n}^p \mathbb{E}_{F_0}\left[\sum_{s=0}^{r}(\chi_{is1} + \chi_{is2})(k_{s+1} - k_s)\right] \qquad (C.4)$$

$$\leq 2n(K + 1)B\tau_{1n}^p \sum_{s=1}^{r} k_s \tau_{1n}^{k_s},$$

where the first inequality follows from the triangle inequality and the last inequality uses that $(\chi_{is1} + \chi_{is2})$ is nonincreasing in s with probability 1. Hence, if $\limsup_n \tau_{1n} n^{1/(k_1+p)} < \infty$, we have

$$\limsup_n \mathbb{E}_{F_0}\left|\log(\mathcal{L}_{1n})\right| = \limsup_n \mathbb{E}_{F_0}\left|\sum_{i=1}^{n} l_{1n}(B_{ik_1}, \ldots, B_{ik_r})\right| < C \qquad (C.5)$$

for a positive constant $C < \infty$. Similarly, for a perturbation of the type $f_n(v) = f_{2n}(v)$, we need $\limsup_n \tau_{2n} n^{1/(K-k_r+p)} < \infty$ for (C.5) to hold.

Using (C.5), we can now adapt the argument from Stone (1980) to show that the rate implied by the sequences τ_{1n} and τ_{2n} is indeed an upper bound on the rate of convergence for a nonparametric estimator of $F_0(v)$. For completeness of the exposition, we now restate his argument: suppose the rate τ_n was not an upper bound on the rate of convergence. Then for $j = 1, 2$, there would be a statistical procedure to decide between $f_{jn}(v)$ and $f_0(v)$ such that the lim sup of the probability of a statistical error is equal to zero.

In particular, if we put prior probability $\frac{1}{2}$ on each f_{jn} and f_0, respectively, the posterior probability of f_{jn} is

$$\pi\big(f = f_{jn}|\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}\big) = \frac{\mathcal{L}_{jn}}{1 + \mathcal{L}_{jn}}.$$

Given the constant in (C.5), choose $\varepsilon = (1 + \exp(C/2))^{-1} > 0$. Then by (C.5) and the Markov inequality,

$$\mathbb{P}\big(\varepsilon < \pi\big(f = f_{jn}|\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}\big) < 1 - \varepsilon\big)$$

$$= \mathbb{P}\left(\varepsilon < \frac{\mathcal{L}_{jn}}{1 + \mathcal{L}_{jn}} < 1 - \varepsilon\right)$$

$$= \mathbb{P}\left(\frac{1}{1 + \exp(C/2)} < \frac{\mathcal{L}_{jn}}{1 + \mathcal{L}_{jn}} < \frac{\exp(C/2)}{1 + \exp(C/2)}\right) \tag{C.6}$$

$$= \mathbb{P}\big(\exp(-C/2) < \mathcal{L}_{jn} < \exp(C/2)\big)$$

$$= \mathbb{P}\big(|\log\mathcal{L}_{jn}| < C/2\big) \geq 1 - \frac{\mathbb{E}|\log\mathcal{L}_{jn}|}{C}.$$

Hence, taking limits yields

$$\liminf_n \mathbb{P}_{F_0}\big(\varepsilon < \pi\big(f = f_n|\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}\big) < 1 - \varepsilon\big) \geq 1 - \limsup_n \frac{\mathbb{E}|\log\mathcal{L}_{jn}|}{C} > \frac{1}{2},$$

so that the error probability of any decision rule between F_{1n} and F_0 has to be at least $\frac{\varepsilon}{4}$.

Now consider the following decision rule δ between F_{1n} and F_0 based on the candidate estimator $\widehat{F}_n$: we set $\delta_n(\widehat{F}_n) := F_0$ if $\varrho(\widehat{F}_n, F_0) < \frac{1}{2}\varrho(F_{1n}, F_0)$ and set $\delta_n(\widehat{F}_n) = F_{1n}$ otherwise. Suppose also that $\limsup_n \tau_{1n} n^{1/(k_1+p)} < \infty$. Then by the previous argument, this decision rule must have error probability $\frac{\varepsilon}{4}$ or greater, so that

$$P\big(\varrho(\widehat{F}_n, F_0) > cn^{-\gamma_1/(k_1+p)}\big)$$

$$\geq \frac{1}{2}\mathbb{P}_{F_0}\left(\frac{1}{2}\varrho(\widehat{F}_n, F_0) > cn^{-\gamma_1/(k_1+p)}\right) + \frac{1}{2}\mathbb{P}_{F_{1n}}\left(\frac{1}{2}\varrho(\widehat{F}_n, F_0) > cn^{-\gamma_1/(k_1+p)}\right)$$

$$\geq \frac{1}{2}\mathbb{P}_{F_0}\left(\frac{1}{2}\varrho(\widehat{F}_n, F_0) > c\tau_{1n}^{\gamma_1}\right) + \frac{1}{2}\mathbb{P}_{F_{1n}}\left(\frac{1}{2}\varrho(\widehat{F}_n, F_0) > c\tau_{1n}^{\gamma_1}\right) \tag{C.7}$$

$$\geq \frac{1}{2}\mathbb{P}_{F_0}\big(\delta_n(\widehat{F}_n) = F_{1n}\big) + \frac{1}{2}\mathbb{P}_{F_{1n}}\big(\delta_n(\widehat{F}_n) = F_0\big) > \frac{\varepsilon}{4}.$$

Applying the same argument to the perturbation

$$F_{2n}(v) := \int_0^v f_0(s)\big(1 + \xi_{2n}(s)\big)\,ds,$$

we also obtain

$$\liminf_n \mathbb{P}_{F_0}\big(\varrho(\widehat{F}_n, F_0) > cn^{-\gamma_2/(K-k_r+p)}\big) > \frac{\varepsilon}{4}. \tag{C.8}$$

Finally, consider a perturbation of $F_0(v)$ in the interior of $\mathcal{V}$ that is of the form

$$F_{3n}(v; \alpha) := \int_0^v f_0(s)\big(1 + \xi_{3n}(s; \alpha)\big)\, ds.$$

By a mean-value expansion around $\alpha = 0$, under F_0 the log likelihood satisfies

$$\sum_{i=1}^n l_{3n}(B_{ik_1}, \ldots, B_{ik_r}; \alpha)$$

$$= \log L_{3n}(B_{ik_1}, \ldots, B_{ik_r}; \alpha) \tag{C.9}$$

$$= 0 + \alpha \sum_{i=1}^n \left\{ \frac{\partial}{\partial \alpha} \log L_{3n}(B_{ik_1}, \ldots, B_{ik_r}; 0) + \frac{\alpha^2}{2} \frac{\partial^2}{\partial \alpha^2} \log L_{3n}(B_{ik_1}, \ldots, B_{ik_r}; \bar{\alpha}) \right\},$$

where $\bar{\alpha} \in [0, \alpha]$.

Now choose a sequence α_n such that $\limsup_n \sqrt{n} \alpha_n < \infty$. Note that the corresponding statistical experiment $L_{3n}(B_{ik_1}, \ldots, B_{ik_r}; \alpha) := \log \frac{dG(B_{ik_1}, \ldots, B_{ik_r}; F_{3n}(\cdot; \alpha))}{dG(B_{ik_1}, \ldots, B_{ik_r}; F_{0n})}$ is differentiable in quadratic mean with respect to α. Moreover, it can be verified that the score identity with respect to α holds at $\alpha = 0$ for every $n = 1, 2, \ldots$, so that

$$\limsup_n \mathbb{E}_{F_0} \left| \sum_{i=1}^n l_{3n}(B_{ik_1}, \ldots, B_{ik_r}; \alpha_n) \right| \leq \limsup_n \big(\sqrt{n} \alpha_n |O(1)| + n \alpha_n^2 |O(1)|\big) < \infty$$

and by the same line of reasoning as for the first case, so that

$$\liminf_n \mathbb{P}_{F_0}\big(\varrho(\widehat{F}_n, F_0) > cn^{-\gamma_3/2}\big) > \frac{\varepsilon}{4}. \tag{C.10}$$

Taken together, (C.7), (C.8), and (C.10) establish the first assertion of the lemma.

For the second part of Lemma C.1, consider a decision problem in which we put $\frac{1}{M}$ prior probability on each of the distributions

$$F_{nm,M}(v) = F_0(v) + \frac{m}{M-1}\big(F_{1n}(v) - F_0(v)\big)$$

for some $M > 1$ and $m = 0, \ldots, M - 1$. Again following the reasoning in Stone (1980) and adapting the arguments leading to (C.4) and (C.6), we can show that the overall error probability of any procedure $\delta_{nM} : \mathcal{F}_0 \to \{F_{n0,M}, \ldots, F_{nM,M}\}$ of classifying F into the M points based on $\widehat{F}_n$ is at least $1 - \frac{2}{M}$ can be bounded from below by

$$\liminf_n \mathbb{P}_{F_0}\big(\varrho(\widehat{F}_n, F_0) > cn^{-\gamma_1/(k_1+p)}\big) \geq \liminf_n \mathbb{P}_{F_0}\big(\delta_{nM}(\widehat{F}_n) \neq F_0\big) > 1 - \frac{2}{M}.$$

Since for any $M > 1$, we can pick $c > 0$ small enough such that for large n, $\varrho(F_{n1,M}, F_0) > cn^{-\gamma_1/(k_1+p)}$, we can make the probability on the right-hand side of this inequality arbitrarily close to 1 as we take the limit $c \to 0$. Applying the same argument to the perturbations F_{2n} and F_{3n}, we establish the second claim. □

PROOF OF THEOREM 2.3. Without loss of generality, consider the case $K - k_r \leq k_1$.

For part (a), note that by Assumption 1.4 for the local alternatives defined in (C.1), $\sup_{v \in \mathcal{V}} |F_{1n}(v) - F_0(v)| \geq \frac{\tau_{1n}^{p+1} \inf_{v \in \mathcal{V}} F_0(v)}{2}$, $\sup_{v \in \mathcal{V}} |F_{2n}(v) - F_0(v)| \geq \frac{\tau_{2n}^{p+1} \inf_{v \in \mathcal{V}} F_0(v)}{B}$, and $\sup_{v \in \mathcal{V}} |F_{3n}(v) - F_0(v)| = \alpha_n \delta$, so that for $\varrho(F, G) := \sup_{v \in \mathcal{V}} |F(v) - G(v)|$, $\gamma_1 = \gamma_2 = p + 1$, so that by Lemma C.1, (2.1) and (2.2) hold with $r_n = cn^{\min\{1/2, (p+1)/(k_1+p), (p+1)/(K-k_r+p)\}}$.

Next we establish part (b). From the definition of $\psi(\cdot)$ and the lower bound on the density $f_0(v)$ from Assumption 1.4, there exist $\eta_1, \eta_2 \in (0, \frac{1}{2})$ and $\kappa > 0$ such that $\int_{\inf \mathcal{V}}^{F_0^{-1}(\tilde{\tau})} \psi_{jn}(v) f_0(v) \, dv > \kappa \tilde{\tau}^{1+p}$ for all $\tilde{\tau} \in [\eta_1 \tau_{1n}, \eta_2 \tau_{1n}]$. Hence, we have by a change of variables,

$$
\begin{aligned}
\left(\left\| F_{1n}(v) - F_0(v) \right\|_q \right)^q &= \int_{-\infty}^{\infty} \left(F_{1n}(v) - F_0(v) \right)^q \mu(dv) \\
&= \int_0^{\tau_{1n}} \left(F_{1n}\left(F_0^{-1}(s)\right) - s \right)^q h(s; F_0)^{-1} \, ds \\
&\geq \int_{\eta_1 \tau_{1n}}^{\eta_2 \tau_{1n}} \kappa^q s^{q(1+p)} h(s; F_0)^{-1} \, ds \\
&\gtrsim \tau_{1n}^{q(1+p) - \alpha_1 + 1}
\end{aligned}
$$

for small values of τ_{1n} using the rates imposed in Assumption 1.4. Hence,

$$
\left\| F_{1n}(v) - F_0(v) \right\|_q \gtrsim \tau_{1n}^{(q - \alpha_1 + 1)/q} \asymp n^{-(q(1+p) - \alpha_1 + 1)/(q(k_1 + p))}.
$$

We can apply an analogous argument to the perturbations F_{2n} and F_{3n} so that by Lemma C.1, conditions (2.1) and (2.2) hold with

$$
r_n = cn^{\min\{1/2, (q(1+p) - \alpha_1 + 1)/(q(k_1 + p)), (q(1+p) - \alpha_2 + 1)/(q(K - k_r + 1 + p))\}}.
$$

Part (c) follows immediately from part (b), noticing that by restricting the function to any compact subset A of the interior of $\mathcal{V}$, there exists a finite n (depending on A) such that the perturbations F_{1n} and F_{2n} coincide with F_0 on A and, therefore, do not impose any restrictions on the rate of convergence. □

PROOF OF THEOREM 2.4. To establish this result, we show how to extend the arguments in the proof of Lemma C.1 to a problem of estimating bounds rather than a single distribution function.

We first show that $n^{(p+1)/(k_r+p)}$ is an upper bound on the rate of convergence. Consider a sequence τ_{1n} such that $\lim_n \tau_{1n} n^{-1/(k_r+p)} = \tau_1 \in (0, 1)$, the interior of the unit interval. By the assumptions of the theorem, for every n, there exists a distribution $f_{0n}(v)$ such that Condition 2.2(i)–(iii) hold at $v \equiv v_{1n} := F_{0n}^{-1}(\tau_{1n})$.

We now consider a perturbation of the density by $\xi_{1n}(v)$ defined in (C.2). For the perturbed model $f_{1n}(v) = f_{0n}(v)[1 + \xi_{1n}(v)]$, the log-likelihood ratio between the implied distributions of order statistics $\log\left(\frac{dG(B_{ik_1}, ..., B_{ik_r}; F_{1n})}{dG(B_{ik_1}, ..., B_{ik_r}; F_{0n})}\right)$ is bounded in absolute value by $B\tau_{1n}^p$ for n large enough and a finite universal constant B by Condition 2.2(i) and (ii), and the same arguments as in the proof of Lemma C.1.

Next notice that the proof of Lemma C.1 only requires the polynomial condition on the behavior of the marginal distributions $G_{k_1}(v; F_{0n})$ and $G_{k_r}(v; F_{0n})$ assumed under (i) in the statement of the Condition 2.2(i).

Along the sequence F_{0n}, the estimand of interest is given by

$$H_{0n}(v) = H^*\big(\psi_{k_1}\big(G_{k_1}(v; F_{0n})\big), \ldots, \psi_{k_r}\big(G_{k_r}(v; F_{0n})\big)\big).$$

Let the pseudometric be

$$\varrho(F_{1n}, F_{0n}) := \sup_{v \in \mathcal{V}} \big| H^*\big(\psi_{k_1}\big(G_{k_1}(v; F_{1n})\big), \ldots, \psi_{k_r}\big(G_{k_r}(v; F_{1n})\big)\big)$$

$$- H^*\big(\psi_{k_1}\big(G_{k_1}(v; F_{0n})\big), \ldots, \psi_{k_r}\big(G_{k_r}(v; F_{0n})\big)\big)\big|.$$

Since $F_{1n}(v) = F_{0n}(v) + \int_0^v \xi_{1n}(t)\, dt$, the lower bound on the one-sided derivatives of $H^*(\cdot)$ together with Condition 2.2(ii) imply that we have

$$\varrho(F_{1n}, F_{0n}) \geq \big| H^*\big(\psi_{k_1}\big(G_{k_1}(v_{1n}; F_{1n})\big), \ldots, \psi_{k_r}\big(G_{k_r}(v_{1n}; F_{1n})\big)\big)$$

$$- H^*\big(\psi_{k_1}\big(G_{k_1}(v_{1n}; F_{0n})\big), \ldots, \psi_{k_r}\big(G_{k_r}(v_{1n}; F_{0n})\big)\big)\big|$$

$$\geq \frac{1}{2} b_H \tau_{1n}^{p+1}.$$

Now using the same argument as in the proof of Lemma C.1, it follows that $r_n := n^{(p+1)/(k_r+p)}$ is an upper bound for the rate of convergence for any nonparametric estimator $\widehat{H}_n$.

To establish that $n^{(p+1)/(K-k_1+1+p)}$ is an upper bound on the rate of convergence, we consider a sequence τ_{2n} such that $\lim_n \tau_{1n} n^{-1/(K-k_1+p)} = \tau_2 \in (0, 1)$ and a sequence of distributions $f_{0n}(v)$ satisfying Condition 2.2(i)–(iii) at $v \equiv v_{2n} := F_{0n}^{-1}(1 - \tau_{2n})$. For the perturbed model $f_{2n}(v) = f_{0n}(v)[1 + \xi_{2n}(v)]$, where $\xi_{1n}(v)$ is defined in (C.2), we can use the same steps as before to show that the expectation of the absolute value of the log-likelihood ratio is bounded from above by a finite constant that does not depend on n and that $\varrho(F_{2n}, F_{0n}) \geq \frac{1}{2} b_H \tau_{2n}^{p+1}$.

As in the main result in Theorem 2.3, the bound by the regular rate $n^{1/2}$ follows from smooth perturbations of F_{0n} in the interior of the support.   □

Appendix D: Proof of Theorem 3.1

Fix a value of k. We are now going to establish that for the estimator $\widehat{F}_{nk}$, the conditions of Lemma 3.1 hold. Define $\psi_k(\tau) := \phi_k^{-1}(\tau)$. Then $\psi_k'(\tau) = \frac{1}{\phi_k'(\phi_k^{-1}(\tau))}$ and as shown in the proof of Theorem 2.1, $\psi_k''(\tau) = \frac{\phi_k''(\phi_k^{-1}(\tau))}{[\phi_k'(\phi_k^{-1}(\tau))]^3}$. Also recall that $\psi_k(\tau)$ behaves like $\tau^{1/k}$ for small values of τ and is approximated by $\tau^{1/(K-k+1)}$ for values of τ sufficiently close to 1.

In particular, $\psi_k'(\tau) = O(\tau^{1/k-1})$ for $\tau \to 0$ and $\psi_k'(1 - \tau) = O((1 - \tau)^{1/(K-k)-1})$ so that for a given choice of α_{nk}, $\tau_{1nk} := \tau_{1k}^*(\alpha_n) = O(\alpha_{nk}^{-k/(k-1)}) = O(n^{-(k-1)/(2(2k-1))})$ and $1 - \tau_{2nk} := \tau_{2k}^*(\alpha_n) = O(\alpha_n^{-(K-k+1)/(K-k)}) = O(n^{-(K-k)/(2(2(K-k)+1))})$.

If $K = 1$, then $\psi''_1(\tau) = 0$, in which case the approximation in Theorem 3.1 is trivially true without any need for regularization; therefore, we only consider the case $k \geq 2$ in the remainder of this argument. Since $\frac{\psi''_k(\tau)}{|\psi'_k(\tau)|}$ diverges for $k \geq 2$ as $\tau \to 0$ and for $K - k \geq 2$ as $\tau \to 1$, we can bound the supremum $\sup_{\tau \in [\tau_1, 1/2]} |\frac{\psi''_k(\tau)}{|\psi'_k(\tau)|}|$ of the ratio by a multiple $2\frac{\tau_1^{1/k-2}}{\tau_1^{1/k-1}} = \tau_1^{-1}$ for τ_1 sufficiently small. A similar argument applies to the upper tail of the distribution.

In the following discussion, we can, without loss of generality, restrict our attention to the case in which V_k is uniformly distributed, that is, $F_0(\tau) = \tau$ for every $\tau \in [0, 1]$. Note that by assumption, $\lim_{n \to \infty} \frac{r_n}{\tau_{jn}} = c < \infty$, potentially zero, for $j = 1, 2$. Then along a sequence $h_n \to h$ of functions $h_n : [0, 1] \to \mathbb{R}$ converging to $h(\tau)$ with respect to the sup-norm, where $\sup_{\tau \in [0,1]} |h(\tau)| := \|h\|_\infty < \infty$ and $\tau + r_n^{-1} h_n(\tau)$ is a proper c.d.f. for n large enough, we have, by a mean-value expansion in $h(\tau)$,

$$R_n(h_n) := \sup_{\tau \in [0,1]} \left| r_n \left(|\tilde{\psi}'_k(\tau)| \right)^{-1} \left(\tilde{\psi}_k \left(\tau + r_n^{-1} h_n(\tau) \right) - \tilde{\psi}_k(\tau) - r_n^{-1} \psi'(\tau) h_n(\tau) \right) \right|$$

$$= \sup_{\tau \in [0,1]} \left| r_n \left(|\tilde{\psi}'_k(\tau)| \right)^{-1} \tilde{\psi}''_k \left(\tau + r_n^{-1} \bar{h}_n(\tau) \right) r_n^{-1} h_n(\tau) \right|$$

$$= \sup_{\tau \in [r_n^{-1} h_n(0), 1 + r_n^{-1} h_n(1)]} \left| \left(|\tilde{\psi}'_k \left(\tau - r_n^{-1} \bar{h}_n(\tau) \right) | \right)^{-1} \tilde{\psi}''_k(\tau) h \left(\tau - r_n^{-1} \bar{h}_n(\tau) \right) \right|$$

for n large enough, where $\bar{h}_n(\tau)$ takes a value between zero and $h_n(\tau)$ for every value of τ.

Noting that for $\tau < \tau_{1nk}$, $\tilde{\psi}''_k(\tau) = 0$, and given our previous discussion of the tail behavior of the derivatives of $\tilde{\psi}_k(\tau)$, we can now bound

$$R_n(h_n) \leq \sup_{\tau \in [\tau_{1n}, 1/2]} \left| \frac{(\tau - r_n^{-1} \bar{h}_n(\tau))^{1-1/k}}{\tau^{2-1/k}} h \left(\tau - r_n^{-1} \bar{h}_n(\tau) \right) \right|$$

$$+ \sup_{\tau \in [1/2, \tau_{2n}]} \left| \frac{(1 - \tau - r_n^{-1} \bar{h}_n(\tau))^{1-1/(K-k+1)}}{(1-\tau)^{2-1/(K-k+1)}} h \left(\tau - r_n^{-1} \bar{h}_n(\tau) \right) \right|$$

$$\leq \sup_{\tau \in [\tau_{1n}, 1/2]} \left| \frac{(\tau + r_n^{-1} |\bar{h}_n(\tau)|)^{1-1/k}}{\tau^{2-1/k}} \right| \left| h \left(\tau - r_n^{-1} \bar{h}_n(\tau) \right) \right|$$

$$+ \sup_{\tau \in [1/2, \tau_{2n}]} \left| \frac{(1 - \tau + r_n^{-1} |\bar{h}_n(\tau)|)^{1-1/(K-k+1)}}{(1-\tau)^{2-1/(K-k+1)}} \right| \left| h \left(\tau - r_n^{-1} \bar{h}_n(\tau) \right) \right|$$

$$\leq 2 \left(r_n^{1-1/k} \tau_{1nk}^{1/k-2} + r_n^{1-1/(K-k+1)} (1 - \tau_{2nk})^{1/(K-k+1)-2} \right)$$

$$\times \sup_{\tau \in [0,1]} |h_n(\tau)|^{(2K-1)/K}$$

for n large enough, since by assumption, $\sup_{\tau \in [0,1]} |h_n(\tau)| \leq 2 \sup_{\tau \in [0,1]} |h(\tau)| + 1 < \infty$, say. Here, $r_n = n^{-1/2}$, so that by our assumptions on τ_{1nk} and τ_{2nk}, this expression goes

to zero for any limiting function $h(\tau)$ that satisfies $\sup_{\tau \in [0,1]} |h(\tau)| < \infty$. The same argument applies to linear combinations of estimators for different values of k, so that the regularization scheme in Theorem 3.1 satisfies Condition 3.1. Hence, it follows from Lemma 3.1 that the proposed normalized estimator satisfies the uniform approximation posited in Theorem 3.1. □

APPENDIX E: RATES FOR FUNCTIONALS OF THE DISTRIBUTION OF VALUATIONS

PROOF OF PROPOSITION 4.1. Using integration by parts, we can rewrite the functional $T(F)$ at F as

$$T(F) = \int_0^\infty w(v) F(dv) = \left[w(v) F(v) \right]_0^\infty - \int_0^\infty w'(v) F(v) \, dv.$$

Since by Assumption 4.1, $\lim_{\tau \to 0} \omega(\tau; F)$ is bounded uniformly in F and, furthermore, $\lim_{\tau \to 1} \omega(\tau; F) = 0$ for all F, the first term is equal to zero.

From the definition of $\psi(\cdot)$ and the lower bound on the density $f_0(v)$ from Assumption 1.4, there exist $\eta_1, \eta_2 \in (0, \frac{1}{2})$ and $\kappa > 0$ such that $\int_{\inf \mathcal{V}}^{F_0^{-1}(\tilde{\tau})} \psi_{jn}(v) f_0(v) \, dv > \kappa \tilde{\tau}^{1+p}$ for all $\tilde{\tau} \in [\eta_1 \tau_{1n}, \eta_2 \tau_{1n}]$. Also by construction of the perturbation, $F_{1n}(v) \geq F_0(v)$ for all v, so that by Assumption 4.1, the integrand does not change sign on the interval $[0, \tau_{1n}]$ for n large enough. Hence,

$$\begin{aligned}
\left| T(F_{1n}) - T(F_0) \right| &= \left| \int_{\inf \mathcal{V}}^{F_0^{-1}(\tau_{1n})} w'(v) \left(F_{1n}(v) - F_0(v) \right) dv \right| \\
&= \left| \int_0^{\tau_{1n}} \omega'(s; F_0) \left(F_{1n}(F_0^{-1}(s)) - s \right) ds \right| \\
&\geq \left| \int_{\eta_1 \tau_{1n}}^{\eta_2 \tau_{1n}} \omega'(s; F_0) \kappa s^{1+p} \, ds \right| \\
&\gtrsim \tau_{1n}^{2+p+\beta_1}
\end{aligned}$$

for n sufficiently large. Similarly, $|T(F_{2n}) - T(F_0)| \gtrsim \tau_{2n}^{2+p+\beta_2}$ and $|T(F_{3n}) - T(F_0)| \gtrsim \alpha_n$, so that by Lemma C.1, conditions (2.1) and (2.2) hold with

$$r_n = n^{\min\{1/2, (2+p+\beta_1)/(k_1+p), (2+p+\beta_2)/(K-k_r+1+p)\}}.$$ □

PROOF OF PROPOSITION 4.2. Note that since, by assumption, $V > 0$ with probability 1, then

$$\mathbb{E}_F[V_{k:\tilde{K}}] = \int_0^\infty \left[1 - G_k^{\tilde{K}}(v; F) \right] dv.$$

By the same argument as in the proof of Proposition 4.1, we can find $\eta_1, \eta_2 \in (0, \frac{1}{2})$ and $\kappa > 0$ such that $\int_{\inf \mathcal{V}}^{F_0^{-1}(\tilde{\tau})} \psi_{jn}(v) f_0(v) \, dv > \kappa \tilde{\tau}^{1+p}$ for all $\tilde{\tau} \in [\eta_1 \tau_{1n}, \eta_2 \tau_{1n}]$. Since $F_{1n}(v) - $

$F_0(v) = 0$ for all $v \geq F_0^{-1}(\tau_{1n})$, we can write

$$\mathbb{E}_{F_0}[V_{k:\tilde{K}}] - \mathbb{E}_{F_{1n}}[V_{k:\tilde{K}}]$$

$$= \int_0^{F_0^{-1}(\tau_n)} \big[G_k^{\tilde{K}}(v; F_n) - G_k^{\tilde{K}}(v; F_0) \big] \, dv$$

$$= \frac{\tilde{K}!}{k!(\tilde{K}-k)!}$$

$$\times \int_0^{F_0^{-1}(\tau_n)} \big[F_{1n}(v)^k \big(1 - F_{1n}(v)\big)^{\tilde{K}-k} - F_0(v)^k \big(1 - F_0(v)\big)^{\tilde{K}-k} \big] \, dv$$

$$= \frac{\tilde{K}!}{k!(\tilde{K}-k)!}$$

$$\times \int_0^{\tau_{1n}} \big[\big(F_{1n}\big(F_0^{-1}(s)\big)\big)^k \big(1 - F_{1n}\big(F_0^{-1}(s)\big)\big)^{\tilde{K}-k} - s^k (1-s)^{\tilde{K}-k} \big] h(s; F_0)^{-1} \, ds$$

$$\geq \frac{\tilde{K}!}{k!(\tilde{K}-k)!}$$

$$\times \int_{\eta_1 \tau_{1n}}^{\eta_2 \tau_{1n}} \big[(s + \kappa s^{1+p}) k \big(1 - s - \kappa s^{1+p}\big)^{\tilde{K}-k} - s^k (1-s)^{\tilde{K}-k} \big] h(s; F_0)^{-1} \, ds$$

$$\gtrsim \int_{\eta_1 \tau_{1n}}^{\eta_2 \tau_{1n}} s^{k(1+p)} h(s; F_0)^{-1} \, ds$$

$$\gtrsim \tau_{1n}^{k(1+p)+1-\alpha_1}$$

for n sufficiently large, since the integrand is always nonnegative. Similarly,

$$\big| \mathbb{E}_{F_{2n}}[V_{k:\tilde{K}}] - \mathbb{E}_{F_0}[V_{k:\tilde{K}}] \big| \gtrsim \tau_{2n}^{(\tilde{K}-k)(1+p)+1-\alpha_2}$$

and

$$\big| \mathbb{E}_{F_{3n}}[V_{k:\tilde{K}}] - \mathbb{E}_{F_0}[V_{k:\tilde{K}}] \big| \gtrsim \alpha_n$$

so that the conclusion follows from Lemma C.1. $\qquad\square$

PROOF OF PROPOSITION 5.1. In a first step we apply a modification of Lemma C.1 to the distribution $G_0(v)$ of a random bid $B_{\widehat{K}}$, where the index $\widehat{K}$ is drawn at random from a uniform distribution over $\{1, 2, \ldots, K\}$, the set of all bidders.

Let $\eta_n := \eta n^{-1/(2p+1)}$, and let τ_0 and $\psi(\cdot)$ be as defined in Appendix B. In an analogous fashion as before, we define the perturbations of the distribution of a random bid $g_{jn}(v) := g_0(v)[1 + \psi_{jn}(v)]$ for $j = 1, 2, 3$, where

$$\psi_{1n}(v) := \tau_{1n}^p \bigg(\beta_{1n} \psi \bigg(G_0^{-1} \bigg[\tau_{1n}^{-1} \bigg(\frac{\tau_{1n}}{2} - G_0(v) \bigg) \bigg] \bigg)$$

$$- \psi \bigg(G_0^{-1} \bigg[\tau_{1n}^{-1} \bigg(G_0(v) - \frac{\tau_{1n}}{2} \bigg) \bigg] \bigg) \bigg),$$

$$\psi_{2n}(v) := \tau_{2n}^p \left(\beta_{2n} \psi \left(G_0^{-1} \left[\tau_{2n}^{-1} \left(G_0(v) - 1 + \frac{\tau_{2n}}{2} \right) \right] \right) \right.$$
$$\left. - \psi \left(G_0^{-1} \left[\tau_{2n}^{-1} \left(1 - \frac{\tau_{2n}}{2} - G_0(v) \right) \right] \right) \right),$$

$$\psi_{3n}(v) := \eta_n^p \left\{ \psi(G_0^{-1}[\eta_n^{-1}(\tau_0 - G_0(v))]) - \beta_{3n} \psi(G_0^{-1}[\eta_n^{-1}(\tau_0 - G_0(v))]) \right\},$$

where for all j, the sequence β_{jn} is bounded between $\frac{1}{B}$ and B, and is chosen in a way such that $g_{jn}(v)$ is a density. Also let $G_{jn}(v)$ be the corresponding cumulative distribution functions.

Following the same arguments as in the proof of Lemma C.1, the expectation of the absolute value of the log-likelihood ratio for the sequence of deviations G_{3n} is of the order $n\eta_n^{2p+1}$ and, therefore, is bounded as $n \to \infty$ so that the error probability of any classification procedure to distinguish between G_0 and G_{3n} is bounded away from zero. For the deviations G_{1n} and G_{2n}, the argument is identical to the original version of the lemma. Hence, if $\varrho(G_0, G_{3n}) \geq \eta_n^{\gamma_3}$, the conclusion of Lemma C.1 can be modified to

$$\limsup_n P(\varrho(\widehat{G}_n, G_0) > cn^{-\min\{\gamma_3/(2p+1), \gamma_1/(k_1+p), \gamma_2(K-k_r+p)\}}) > 0$$

and

$$\lim_{c \to 0} \limsup_n P(\varrho(\widehat{G}_n, G_0) > cn^{-\min\{\gamma_3/(2p+1), \gamma_1/(k_1+p), \gamma_2(K-k_r+p)\}}) = 1.$$

Therefore, it suffices to show that $\varrho(G_1, G_2) := \sup_{v \in \mathcal{V}} |F(v; G_1) - F(v; G_2)|$ satisfies the conditions of this modification of Lemma C.1 with $\gamma_1 = p + 1$, $\gamma_2 = p$, and $\gamma_3 = p$:

Fix $b \in \mathcal{V}$ and let $\tau := G_0(b)$. Since the bidding functions are strictly monotone in valuations, $F(b^{-1}(b; F)) = \tau$, that is, the ordering of quantiles is preserved. For perturbation F_{jn}, note that using (5.2) and a mean-value expansion, we can write the valuation implied by bid $b = G_0^{-1}(\tau)$ as

$$b^{-1}(b; F_{1n}) = b + \frac{1}{K-1} \frac{G_{1n}(b)}{g_{1n}(b)}$$
$$= b + \frac{1}{K-1} \frac{G_0(b) + \int_{\inf \mathcal{V}}^b \psi_{jn}(s) g_0(s)\, ds}{g_0(b)[1 + \psi_{jn}(b)]}$$
$$= b + \frac{1}{K-1} \frac{G_0(b)}{g_0(b)} \tag{E.1}$$
$$+ \frac{1}{K-1} \left\{ \frac{\int_{\inf \mathcal{V}}^b \psi_{jn}(s) g_0(s)\, ds}{g_0(b)} - \frac{G_0(b)}{\bar{g}_n(b)^2} \psi_{jn}(b) \right\},$$

where $\bar{g}_n(b)$ is an intermediate value between $g_0(b)$ and $g_{jn}(b)$. Also, from the bounds on the density function $f(v)$ in Assumption 1.4, the bidding function $b(v; F)$ and its inverse

are Lipschitz continuous for F_0 and F_{jn}. Using the expansion (E.1), we can write

$$
\begin{aligned}
F_{jn}&\bigl(b^{-1}(b; F_{jn})\bigr) - F_0\bigl(b^{-1}(b; F_0)\bigr) \\
&= F_0\bigl(b^{-1}(b; F_0) + \bigl[b^{-1}(b; F_{jn}) - b^{-1}(b; F_0)\bigr]\bigr) \\
&\quad - F_0\bigl(b^{-1}(b; F_0)\bigr) + F_{jn}\bigl(b^{-1}(b; F_{jn})\bigr) - F_0\bigl(b^{-1}(b; F_{jn})\bigr) \qquad\text{(E.2)} \\
&= \int_0^{b^{-1}(b; F_{jn})} \psi_{jn}(s) g_0(s)\, ds \\
&\quad + \frac{f_0(\bar v_n)}{K-1}\left(\frac{\int_{\inf \mathcal V}^{b} \psi_{jn}(s) g_0(s)\, ds}{g_0(b)} - \frac{G_0(b)}{\bar g_n(b)^2}\psi_{jn}(b)\right).
\end{aligned}
$$

Now consider the perturbation F_{1n} and note that for any value $\tau \geq \frac{\tau_{1n}}{2}$, all three terms in the approximation error in (E.2) are nonnegative, so that

$$
\begin{aligned}
\sup_{b\in\mathcal V}&\bigl|F_{1n}\bigl(b^{-1}(b; F_{1n})\bigr) - F_0\bigl(b^{-1}(b; F_0)\bigr)\bigr| \\
&\geq \sup_{\tau\in[\tau_{1n}/2,\tau_{1n}]}\Biggl\{\left|\int_0^{b^{-1}(G_0^{-1}(\tau); F_{1n})} \psi_{1n}(s) g_0(s)\, ds\right| \\
&\quad + \frac{\inf\limits_{v\in\mathcal V} f_0(v) f_0}{K-1}\left(\left|\frac{\int_{\inf\mathcal V}^{G_0^{-1}(\tau)} \psi_{1n}(s) g_0(s)\, ds}{g_0(b)}\right| + \left|\frac{\tau\psi_{1n}(G_0^{-1}(\tau))}{\sup\limits_{v\in\mathcal V} g_0(v)^2}\right|\right)\Biggr\} \\
&\gtrsim \tau_{1n}^{1+p} + \tau_{1n}^{1+p} + \tau_{1n}^{1+p}
\end{aligned}
$$

so that $\gamma_1 = p + 1$. For the local alternatives F_{2n}, the argument is analogous, except that the third term of the approximation $\left|\frac{\tau\psi_{2n}(G_0^{-1}(\tau))}{\sup_{v\in\mathcal V} g_0(v)^2}\right|$ is of the order of τ_{2n}^p for $\tau \in [1 - \tau_{2n}, 1 - \frac{\tau_{2n}}{2}]$, which gives us $\gamma_2 = p$.

Similarly, we have for the perturbation F_{3n} that

$$
\begin{aligned}
\sup_{v\in\mathcal V}&\bigl|F_{3n}(v) - F_0(v)\bigr| \\
&\geq \sup_{\tau\in[\tau_0,\tau_0+\eta_n^{-1}]}\Biggl\{\left|\int_0^{b^{-1}(G_0^{-1}(\tau); F_{3n})} \psi_{3n}(s) g_0(s)\, ds\right| \\
&\quad + \frac{\inf\limits_{v\in\mathcal V} f_0(v)}{K-1}\left(\left|\frac{\int_{\inf\mathcal V}^{G_0^{-1}(\tau)} \psi_{3n}(s) g_0(s)\, ds}{g_0(b)}\right| + \left|\frac{G_0(G_0^{-1}(\tau))\psi_{3n}(G_0^{-1}(\tau))}{\sup\limits_{v\in\mathcal V} g_0(v)^2}\right|\right)\Biggr\} \\
&\gtrsim \eta_n^{1+p} + \eta_n^{1+p} + \eta_n^{p}
\end{aligned}
$$

so that $\gamma_3 = p$, which establishes the claim. $\qquad\square$

REFERENCES

Aït-Sahalia, Y. and J. Duarte (2003), "Nonparametric option pricing under shape restrictions." *Journal of Econometrics*, 116, 9–47. [354]

Andrews, D. and M. Schafgans (1998), "Semiparametric estimation of the intercept of a sample selection model." *Review of Economic Studies*, 65, 487–517. [332]

Aradillas-López, A., A. Gandhi, and D. Quint (2013), "Identification and inference in ascending auctions with correlated private values." *Econometrica*, 81, 489–534. [330, 337, 345]

Armstrong, T. (2011), "Bounds in auctions with unobserved heterogeneity," Working paper, Stanford University. [330]

Athey, S. and P. Haile (2002), "Identification of standard auction models." *Econometrica*, 70, 2107–2140. [331, 332]

Athey, S. and P. Haile (2007), "Nonparametric approaches to auctions." In *Handbook of Econometrics*, Vol. 6A (J. Heckman and E. Leamer, eds.), 3847–3965, North-Holland, Amsterdam. [329, 331]

Berry, S., O. Linton, and A. Pakes (2004), "Limit theorems for estimating the parameters of differentiated product demand systems." *Review of Economics Studies*, 71, 613–654. [332]

Chernozhukov, V., S. Lee, and A. Rosen (forthcoming), "Inference on intersection bounds." *Econometrica*. [330, 349]

David, H. and H. Nagaraja (2003), *Order Statistics*, third edition. Wiley, Hoboken. [335, 355]

Guerre, E., I. Perrigne, and Q. Vuong (2000), "Optimal nonparametric estimation of first-price auctions." *Econometrica*, 68 (3), 525–574. [331, 350, 354, 355]

Haile, P. and E. Tamer (2003), "Inference with an incomplete model of English auctions." *Journal of Political Economy*, 111 (1), 1–51. [330, 331, 337, 341, 343, 345, 352]

Horowitz, J. (1992), "A smoothed maximum score estimator for the binary response model." *Econometrica*, 60, 505–531. [332]

Imbens, G. (2004), "Nonparametric estimation of average treatment effects under exogeneity: A review." *Review of Economics and Statistics*, 86, 4–29. [332]

Khan, S. and E. Tamer (2010), "Irregular identification, support conditions, and inverse weight in estimation." *Econometrica*, 78 (6), 2021–2042. [332]

Komarova, T. (2009), "Nonparametric identification in asymmetric second-price auctions: A new approach." Working paper, LSE. [331, 332]

Laffont, J., H. Ossard, and Q. Vuong (1995), "Econometrics of first-price auctions." *Econometrica*, 63 (4), 953–980. [331]

Lewbel, A. (2000), "Semiparametric qualitative response model estimation with unknown heteroskedasticity or instrumental variables." *Journal of Econometrics*, 97, 145–177. [332]

Mammen, E. and C. Thomas-Agnan (1999), "Smoothing splines and shape restrictions." *Scandinavian Journal of Statistics*, 26, 239–252. [354]

Manski, C. (1975), "Maximum score estimation of the stochastic utility model of choice." *Journal of Econometrics*, 3, 205–228. [332]

Riley, J. and W. Samuelson (1981), "Optimal auctions." *American Economic Review*, 71 (3), 381–392. [351]

Rockafellar, R. (1972), *Convex Analysis*. Princeton University Press, Princeton. [352]

Roehrig, C. (1988), "Conditions for identification in nonparametric and parametric models." *Econometrica*, 56 (2), 433–447. [331]

Stone, C. (1980), "Optimal rates of convergence for nonparametric estimators." *The Annals of Statistics*, 8, 1348–1360. [331, 338, 344, 364, 366]

Stone, C. (1983), "Optimal global rates of convergence for nonparametric regression." *The Annals of Statistics*, 10, 1040–1063. [353]

van der Vaart, A. (1998), *Asymptotic Statistics*. Cambridge University Press, Cambridge. [346]

van der Vaart, A. and J. Wellner (1996), *Weak Convergence and Empirical Processes*. Springer, New York. [352, 356, 357, 358]