ESTIMATION OF AUCTIONS WITH INCOMPLETE BIDDING DATA

KONRAD MENZEL[†] AND PAOLO MORGANTI[♯]

Abstract. We consider estimation of independent private auction models when only a subset of the bids for each auction are observed. Even though many objects of interest remain nonparametrically identified from incomplete bidding data, estimation of the distribution of valuations and some of its functionals is shown to be *irregular* in the sense that the inverse of the mapping from the parent distribution to that of the observable bids is not Lipschitz continuous. We derive the optimal rate of convergence for the c.d.f. of valuations depending on the number of bidders and which particular bids are observed for each auction. Furthermore, we propose a trimming procedure that yields an estimator which is asymptotically Gaussian at an adaptive rate. We also discuss implications for other functionals of the parental distribution. In particular it is shown that expected revenue and optimal reserve price are not estimable at the root-n rate in general, but the rates will depend on the relative sizes of the observed and the counterfactual auctions. While most of our results are on second-price auctions, we also demonstrate how our findings apply to first-price and descending-bid formats. Our results also suggests that imposing smoothness restrictions on the underlying valuation distribution may improve large-sample behavior of nonparametric estimators substantially.

JEL Classification: C13, C14, D44
Keywords: Empirical Auctions, Uniform Consistency, Irregular Identification, Lipschitz Continuity

The distribution of bidders' valuations is one of the primary objects of interest in the empirical analysis of independent private values (IPV) auctions. Under the IPV assumption, auction theory makes strong predictions on virtually any question of practical interest based on knowledge about this distribution, which in turn is known to be nonparametrically identified under fairly general conditions.

In this paper, we analyze asymptotic properties of nonparametric estimators for the cumulative distribution function (c.d.f.) of valuations when bidding data is incomplete, that is if only particular bids are observed for each auction in our sample. We find that even though the

---

Date: March 2010, this version: August 2010.
We thank Bryan Graham, Jörg Stoye, Nicola Persico, Stéphane Bonhomme, and seminar participants at the New York Area Econometric Colloquium, NYU Stern, and Northwestern University for comments. Any suggestions are welcome.
[†] corresponding author, NYU, Department of Economics, Email: konrad.menzel@nyu.edu
[♯] ITAM, Department of Business Administration, Email: paolo.morganti@itam.mx.

pointwise rate of convergence for the c.d.f. of valuations is generally the same as the rate for the distribution function of the order statistics corresponding to the observed bids, rates with respect to the uniform metric and the L_2-norm, respectively, are considerably slower and depend on the number of bidders, and which particular bids are observed.

Incomplete 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 highest bid is recorded. In auctions for one single good, the transaction price is linked to the first or second highest order statistic for most formats, whereas e.g. for book building in an auction of r identical units of the good with K bids, the resulting transaction price would depend on the $(K - r)$th order statistic. In the latter case, a nonparametric estimator based only on the transaction price should be expected to do poorly in approximating both the upper and the lower tail of the valuation distribution.

The results of this paper are also relevant for constructing bounds for the distribution of valuations in settings in which complete bidding data may be available, but some of the conditions of the benchmark model for the auction are relaxed. Haile and Tamer (2003) analyze 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 (2008) for a treatment of the statistical problem of constructing this bound. Also in a recent study Aradillas-López, Gandhi, and Quint (2010) propose 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.

Nonparametric Approach. Beyond its practical relevance, the problem of estimating auctions has two features that make it very appealing to theorists and empirical researchers alike: for one, the structure of the problem is very rich and theoretically well-understood, so that given the auction format, the only major unknown is the distribution of bidders' valuations, $F(v)$. If we are willing to treat $F(v)$ as a structural parameter, auction theory makes strong predictions about outcomes for alternative auction formats. The other attractive characteristic of empirical auctions is that this "deep" parameter $F(v)$ is under reasonable conditions nonparametrically identified from bidding data in a very wide range of settings.

Nonparametric identification of features of a model also allows to interpret parametric procedures as a plausible statistical approximation rather than treating the literal specification of the model as prior knowledge, and nonparametric non-identification 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 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.

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) provide 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) derive optimal nonparametric estimators for first-price auctions when all bids are observed, and in this case, the distribution of valuations can be estimated at the usual nonparametric rate.

Description of Results. Intuitively, if we only observe the highest and/or second highest bid in an auction with a large number of bidders, it is difficult to learn about the lower tail of the distribution, and this is reflected by the slope of the inverse mapping that is used to recover $F_V(v)$ from the joint distribution of bids, and which is in many cases not bounded as we approach the lower bound of the support of the distribution of valuations. This problem affects the rate of convergence of nonparametric estimators, which will in general depend on how "close" any of the observed bids are to the highest and lowest bid in each auction. The resulting rates depend on the norm on F_V, where the sup-norm leads to a slower rate than the L_2 norm, and the point-wise rate for the estimator is the same as that for the joint distribution of bids.[1] We also find that the upper bounds on convergence rates improve substantially if we require the valuation distribution to be smooth, however standard nonparametric procedures need in general not attain the faster rates unless these shape restrictions are imposed for estimation.

[1]For more standard results on optimal rates with respect to the sup- and L_p-norm in nonparametric estimation see also Ibragimov and Has'minskii (1981) and Stone (1983)

Since the quantile transformation linking the distribution of observed bids 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 total variation norm on c.d.f.s.. In this sense, the inverse problem of recovering the parent distribution from the joint distribution of observable bids is not ill-posed. However, in order 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. This feature of our problem bears some resemblance with the irregularly identified problems considered by Khan and Tamer (2009).

The slower speed of convergence for the distribution of valuations also affects the rate for estimators of other quantities of practical interest - e.g. the optimal reserve price - which can be calculated from $F_V(v)$. We analyze bounds on the rate of convergence for linear functionals, expected revenue and the optimal reserve price. In particular it is shown that expected revenue and optimal reserve price is 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.

Finally we show how our findings apply to nonparametric estimation of first-price auctions with incomplete bidding data and descending bid auctions, for which by design only the highest bid is observable. Here, the formal analysis is complicated by the fact that in equilibrium, bidders do not directly reveal their true valuations, but adjust their bids by a factor which depends on the underlying valuation distribution. We propose a nonparametric estimator that uses only the highest bid for each auction and give an upper bound for the rate of convergence of any nonparametric estimator.

Outline of Paper. We will now formally state the estimation problem analyzed in this paper. Section 3 derives optimal convergence rates for nonparametric estimators of the valuation distribution, and section 4 derives the asymptotic distribution of a regularized version of that estimator. We then discuss how these findings affect nonparametric estimators for functionals of the distribution of values, and section 6 shows how to extend the main rate result to the case of first-price auctions.

2. 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 a known number K of bidders, and for each auction, the bidders' valuations $(V_1, \ldots, V_K)$ are drawn independently from the distribution $F_0(v) \in \mathcal{F}_0(\mathcal{V})$ with support $\mathcal{V} \subset \mathbb{R}_+$, where $\mathcal{F}_0(\mathcal{V})$ denotes a subset of the set of c.d.f.s on $\mathcal{V}$ (i.e. the set of upper semi-continuous, nondecreasing functions from $\mathcal{V}^s$ to the unit interval which attain the values 0 and 1 in the closure of $\mathcal{V}$).

The focus will be on symmetric independent private values (IPV) second-price auctions with exogenous participation, but we will argue that the qualitative findings also apply to asymmetric auctions and other auction formats. However, the independence assumption is crucial for non-parametric identification (see Athey and Haile (2002)). The formal assumptions on the auction format and equilibrium bids are summarized in the following assumption:

Assumption 2.1. (Second Price Auction) *We observe data from n i.i.d. auctions of a single good with K risk-neutral bidders each, where (i) participation is exogenous, and (ii) the auction satisfies symmetric independent private values (IPV), $V_i \overset{iid}{\sim} F_0(v)$ 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 $f(v)$. (iv) The auction is sealed-bid second-price or a strategically equivalent format, and participants play weakly dominant strategy with bids $B_i \equiv b^*(V_i) = V_i$.*

In order to keep our results general, we will allow the dataset available to the econometrician to be any r-dimensional subvector of the complete vector $(V_{i1}, \dots, V_{iK})$ of bids:

Assumption 2.2. (Observable Bids) *We observe the $k_1 < k_2 < \cdots < k_r$th lowest bids $\mathbf{B}_i = (B_{ik_1}, B_{ik_2}, \dots, B_{ik_r})$.*

For example, if Assumption 2.1 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.

We will now characterize the tail behavior of the p.d.f. of V in terms of its quantiles and define

$$h(\tau; F) := f\left(F^{-1}(\tau)\right)$$

Assumption 2.3. (Tail behavior of $F_0(v)$) *(i) The p.d.f. $f(v)$ of V_i is bounded from above and away from zero in the interior of the support, and the first p derivatives of $f(v)$ are bounded. (ii) There exist constants α_1, α_2 such that for low quantiles τ, the behavior of the p.d.f. of V_i is characterized by*

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

and in the upper tail of the distribution,

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

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

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 second part of Assumption 2.3 holds with $\alpha_1 = 0$, whereas if V follows a log-normal distribution, the statement holds for $\alpha_1 = \frac{9}{5}$ and potentially larger values.

For a given parent distribution F, denote the joint c.d.f. of the $(k_1, \dots, k_r)$th order statistics by

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

For example the c.d.f. for the k_1th order statistic can be expressed as

$$G_{k_1}(b_{k_1}; F) := \sum_{m=k_1}^{K} F(b_{k_1})^m [1 - F(b_{k_1})]^{K-m} = \frac{K!}{(k_1-1)!(K-k_1)!} \int_0^{F(b_{k_1})} s^{k_1-1}(1-s)^{K-k_1} dt$$

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

$$G_{k_1,k_2}((b_{k_1}, b_{k_2}); F) = N(k_1, k_2; K) \int_0^{F_1} \int_0^{F_2} s_1^{k_1-1}(s_2 - s_1)^{k_2-k_1-1}(1-s_2)^{K-k_2} ds_2 ds_1 \qquad (2.1)$$

where $N(k_1, k_2; K) = \frac{K!}{k_1!(k_2-k_1-1)!(K-k_2)!}$, and $F_s := F(b_{k_s})$, see e.g. David and Nagaraja (2003). We give an expression for the general case in the appendix.

Example 2.1. *To frame thoughts, suppose that we observe the winning bid B_{Ki} 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 $G^K_{B_K}(v, F) = [F(v)]^K$, and the maximum likelihood estimator for the parent distribution is given by*

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

where $\widehat{G}_{nK}(v) = G_K(v, \mathbb{P}_n) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}\left\{B_i^{(K:K)} \le v\right\}$ is the empirical c.d.f. of B_{Ki}.

In this case, the maximum-likelihood estimator has a closed form and is guaranteed to be non-decreasing in v. Also, from Donsker's Theorem, $\sqrt{n}(\hat{G}^K_{n,K} - F^K) \rightsquigarrow \mathbb{G}_F$, a Brownian bridge, and since in addition $\phi_K(\tau)$ is uniformly continuous on the unit interval, the maximum likelihood estimator $\hat{F}_n(v)$ is uniformly consistent for $F(v)$. However, it is important to notice that for any $K > 1$, the mapping $\phi_K^{-1}(\tau) = \sqrt[K]{\tau}$ is not Lipschitz-continuous in $\tau \in [0, 1]$, which will in general affect the rates of convergence for $\widehat{F}_n(v)$ as a function, and also has implications for the limiting distribution of $\widehat{F}_n(v)$ and other nonparametric estimators.

3. Estimation of the c.d.f. and Optimal Rates

In this section, we are going to give bounds on the rate of convergence of the nonparametric estimator for the parent distribution $F_0(v)$. Following Stone (1980), we say that r_n is an *upper bound to the rate of convergence* of $\hat{F}_n$ under the norm $\|\cdot\|$ if

$$\liminf_n \sup_{F\in\mathcal{F}_0} P_F\left(\|\hat{F}_n - F\| > cr_n\right) > 0 \tag{3.1}$$

for any sequence of estimators $\{\hat{F}_n\}_{n\geq 0}$, and

$$\lim_{c\to 0}\liminf_n \sup_{F\in\mathcal{F}_0} P_F\left(\|\hat{F}_n - F\| > cr_n\right) = 1 \tag{3.2}$$

These bounds are not specific to any given estimator $\hat{F}_n$ in the problem. We will establish these bounds on the rate of convergence by constructing a worst-case scenario in terms of a true distribution $F_0 \in \mathcal{F}$ and a local perturbation that can't 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 determining the sharpest bound on the rate is the same for all problems we are considering.

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

$$\lim_{c\to\infty}\limsup_n \sup_{F\in\mathcal{F}_0} P_F\left(\|\hat{F}_n - F\| > cr_n\right) = 0 \tag{3.3}$$

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 equation (3.3) 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 a large number of distinct, and equally reasonable estimators all of which achieve the optimal rate. However, for us the main purpose of establishing optimality of a rate is to demonstrate that an upper bound on the rate is in fact *sharp*, and therefore a useful measure of the difficulty of the nonparametric estimation problem at hand.

3.1. **Upper Bounds on the Rate of Convergence.** In order to illustrate the main idea behind the rates of convergence, we are first going to have a look at the setting in Example 2.1 for which we have a closed-form expression for the maximum-likelihood estimator.

Example 3.1. ***(Example 2.1, 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, $\widehat{G}_{nK}$. By the delta-rule, we have that the point-wise asymptotic mean-square error of this estimator is given by

$$\begin{aligned} MSE\left(\widehat{F}_n\left(F_0^{-1}(\tau)\right)\right) &:= \mathbb{E}\left[\left(\widehat{G}_{nk}\left(F_0^{-1}(\tau)\right) - G_K\left(F_0^{-1}(\tau); F_0\right)\right)^2 \left(G_K\left(F_0^{-1}(\tau); F_0\right)\right)^{\frac{2}{K}-2}\right] \\ &= \frac{1}{n}(1-\tau^K)\tau^{2-K} \end{aligned}$$

Hence if $K > 2$, the MSE at the τ_nth quantile goes to zero only if $\tau_n n^{\frac{1}{K}} \to \infty$. Hence it must be true that for this estimator $\mathbb{P}_{F_0}\left(\sup_{v\in\mathcal{V}}\left|\hat{F}_n(v) - F_0(v)\right| > cn^{-\frac{1}{K}}\right) > 0$ for any $c > 0$. Theorems 3.1 and 3.1 below imply that this is in fact also the optimal uniform rate for estimating $F(v)$.

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.

For the more general setting of n i.i.d. IPV second-price auction for which we observe a vector $(V_{ik_1}, \dots, V_{ik_r})$ of r different bids, we can establish the following upper bound on the rate of convergence:

Theorem 3.1. *Let $\hat{F}_n$ be an estimator for F_0. Then under Assumptions 2.1-2.3, $r_n = n^{-\lambda}$ is an upper bound on the rate of convergence satisfying (3.1) and (3.2), where*

(a) $\lambda = \min\left\{\frac{p+1}{k_1+p}, \frac{p+1}{K-k_r+1+p}, \frac{1}{2}\right\}$ *for the sup-norm* $\|h\|_\infty := \sup_v |h(v)|$,

(b) $\lambda = \min\left\{\frac{1}{2}, \frac{q(1+p)-\alpha_1+1}{q(k_1+p)}, \frac{q(1+p)-\alpha_2+1}{q(K-k_r+1+p)}\right\}$ *for the $L_q(\mu)$-norm* $\|h\|_q := \left(\int h(v)^q \mu(dv)\right)^{1/q}$, *and*

(c) *restricting the function to a compact subset $A \subset \mathcal{V}$, the rate under the norm $\|\cdot\|_q$ is $r_n = n^{1/2}$ for any $q = 1, 2, \dots$ or $q = \infty$.*

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 2.1, for $p = 0$ the upper bound on the uniform rate of convergence implied by Theorem 3.1 is $r_n = n^{-\frac{1}{K}}$, which is the same as the result of our informal discussion of the problem before.

Example 3.2. *In order to get an intuition for the rate result, suppose we observe the k_1 and k_2th order statistics, where $k_2 \geq k_1$, and we want to distinguish whether the observed data was generated by a distribution F_0 or a distribution F_1 obtained by perturbing the parent distribution F_0 below the τ_1-quantile and above the $1-\tau_2$-quantile, respectively. Then from the expression (2.1), the probability of observing a realization of (V_{ik_1}, V_{ik_2}) that is informative for a test between F_0 and F_1 is given by $\mathbb{P}_{F_0}\left(V_{ik_1} \leq F_0^{-1}(\tau_1)\right) = G_{k_1}\left(F_0^{-1}(\tau_1); F\right) = N(k_1; K) \int_0^{\tau_1} s^{k_1-1}(1-s)^{K-k_1} ds = O\left(\tau_1^{k_1}\right)$, whereas the probability of an observation that is relevant for a test between F_0 and F_2 equals $\mathbb{P}_{F_0}\left(V_{ik_1} \geq F_0^{-1}(1-\tau_2)\right) = N(k_2; K) \int_{1-\tau_2}^{1} s^{k_2-1}(1-s)^{K-k_2} ds = O\left(\tau_2^{K-k_2+1}\right)$. Hence if either $k_1 > 1$ or $k_2 < K$, the probability of an observation that is informative about a small perturbation of the tails of the distribution decreases much faster than the size of that perturbation, and the uniform rate of convergence of any estimator $\widehat{F}_n$ will then be driven by the more problematic tail of the distribution.*

3.2. **Consistent Estimation of the Valuation Distribution.** For clarity of the exposition we will focus on the case in which only the kth lowest bid out of K is observed. We will 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 3.1 can be achieved by an estimator that combines the inverted empirical marginal c.d.f.s of the observed bids in a straightforward manner.

For any $k = 1, \ldots, K$, we will define the mapping

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

which is strictly increasing by inspection. From the facts about order statistics given in Appendix A, the 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

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

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,

$$\widehat{F}_n(v) = \phi_k^{-1}\left(\widehat{G}_{nk}(v)\right) =: \psi_k\left(\widehat{G}_{nk}\right) \tag{3.4}$$

where $\psi_k(\tau) := \phi_k^{-1}(\tau)$ for $\tau \in [0,1]$.

As discussed earlier, the inverse mapping $\psi_k(\cdot)$ is not Lipschitz-continuous, but 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 $\tau^{\frac{1}{k}-1}$ for τ close to zero, and like $(1-\tau)^{\frac{1}{K-k+1}-1}$ for τ close to one which diverge to infinity if $k>1$ or $K-k>2$, respectively.

We will 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 will 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 3.1. *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, (ii) there exists finite constants $A_s > 0$ and $\delta_1, \dots, \delta_S$ such that for all $s = 1, \dots, S$, $\lim_{\tau\to\tau_s^*}\left|\frac{\psi(\tau)-\psi(\tau_s^*)}{|\tau-\tau_s^*|^{\delta_s}}\right| = A_s$. Also assume that $\psi'(\tau)$ is monotone and doesn't 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

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

the same constants as in the statement of Condition 3.1.

Theorem 3.2. *Suppose that Condition 3.1 holds and let $\delta := \min\{\delta_1, \dots, \delta_S\}$. Then the rate $r_n = n^{-\delta}$ is achievable for an estimator of the function $\psi(G_0(\tau))$ with respect to the sup-norm.*

See the appendix for a proof. We can now use this result to establish a uniform rate of consistency for the estimator in (3.4):

Proposition 3.1. *Suppose Assumptions 2.1 and 2.2 hold with $r = 1$ and $k_1 = k$. Then the estimator $\hat{F}_n$ in equation (3.4) achieves the rate $r_n = n^{-\lambda}$ with $\lambda = \min\left\{\frac{1}{k_1}, \frac{1}{K-k_r+1}, \frac{1}{2}\right\}$ under the sup-norm. In particular, if $p = 0$ this convergence rate is optimal in the sense of Stone (1980).*

Proof: By Assumption 2.1, 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 3.1 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 $A_1 = \frac{1}{N(k,K)^{\frac{1}{k}}}$ and $A_2 = \frac{1}{N(k,K)^{\frac{1}{K-k+1}}}$. Hence we can apply Theorem 3.2 with $\delta = \min\left\{\frac{1}{k}, \frac{1}{K-k+1}\right\}$, so that the estimator in (3.4) 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 3.1 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, it is straightforward to extend the argument and establish achievability of the bound on the rate established in Theorem 3.1 by considering a procedure which combines the estimators obtained from inverting the marginal distribution of each bid separately.

4. Asymptotic Linearity

Even though, as argued before, the estimation problem is not ill-posed, its linearization is in the cases in which the second derivative of the nonparametric likelihood vanishes near the boundaries of the support of valuations. Since Gaussian asymptotics rely on a linear approximation, there will be a need for regularization unless we restrict our attention to linear functionals of the valuation distribution that are

Example 4.1 ((Example 2.1, continued))**.** *From Donsker's theorem The empirical c.d.f.,* $\sqrt{n}(\widehat{G}_n(v) - G(v)) \stackrel{d}{\to} N(0, G(v)(1-G(v)))$ *uniformly in* v*. The quantile transformation* $\psi_K(\tau) := \sqrt[K]{\tau}$ *is strictly monotone, uniformly continuous in* $\tau \in [0,1]$*, but the corresponding functional mapping* $\psi_K(G)$ *is not Hadamard-differentiable. From the delta rule,*

$$\sqrt{n}(\widehat{F}_n(v) - F(v)) \stackrel{d}{\to} N\left(0, G(v)(1-G(v))|\psi_K'(G(v))|^2\right)$$

pointwise *in* $v \in (\underline{v}, \overline{v})$*. However, this convergence is generally not uniform in* v*, and 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* $\hat{G}_n(v)$ *decreases linearly in* $G(v) \to 0$ *mitigates, but does not resolve this problem as long as* $K \geq 3$*.*

In order 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 satisfying the following condition:

Condition 4.1. *(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 $\tilde{\psi}_n$ there exist continuous linear maps $\tilde{\psi}_n' : \mathbb{B} \to \Gamma$ such that*

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

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 an appropriate way of smoothing 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 4.1 to obtain the same conclusion, which is stated in the following Lemma:

Lemma 4.1. *Suppose Condition 4.1 holds for a sequence of mappings $\tilde{\psi}_n$. Then for every $v \in [\underline{v}, \overline{v}]$,*

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

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 which leads to an asymptotically Gaussian estimator $\hat{F}_n$ of F_0. For the case of one observed bid corresponding to the kth order statistic, the estimator maximizing (**??**) is given by $\psi_k(\hat{G}_{k,n}(v))$, where $\hat{G}_{k,n}$ is the empirical c.d.f. of the kth lowest bid, and for every $k = 1, \dots, K$, $\psi_k(\tau) := \phi_k^{-1}(\tau)$ 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 (A.1). 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 that 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 $\tilde{\psi}_k(\tau; \alpha_n)$, where $\sup_{\tau\in[0,1]} |\tilde{\psi}_k'(\tau, \alpha_{nk})| \le \alpha_{nk}$. More specifically, we define

$$\tau_{1k}^*(\alpha) := \min \{\tau \in [0,1] : \psi_k'(\tau) \le \alpha\}$$

and

$$\tau_{2k}^*(\alpha) := \min \{\tau \in [0,1] : \psi_k'(1-\tau) \le \alpha\}$$

for any $\alpha \ge 1$, and propose the modification

$$\tilde{\psi}_k(\tau, \alpha) := \begin{cases} \alpha\tau & \text{if } \tau \le \tau_{1k}^*(\alpha) \\ 1 - \alpha\tau & \text{if } \tau \ge 1 - \tau_{2k}^*(\alpha) \\ \alpha\tau + w_k(\tau, \alpha)(\psi_k(\tau) - \alpha\tau) & \text{otherwise} \end{cases}$$

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

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

$$\widehat{F}_{nk}(v) := \tilde{\psi}_k\left(\widehat{G}_{nk}(v),\alpha_{nk}\right) \tag{4.1}$$

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

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

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, and furthermore it can be seen easily that this estimator is guaranteed to be nondecreasing.

In order to characterize the distribution of the joint estimator, define

$$\hat{S}_n(v) := \frac{1}{n}\sum_{i=1}^{n}\sum_{s,t=1}^{r}\left[\psi'_{k_s}\left(\hat{G}_{k_s}(v)\right)\psi'_{k_t}\left(\hat{G}_{k_t}(v)\right)\right]^{-1}\left(\mathbb{1}\left\{V_{ik_s}\le v\right\}-\hat{G}_{k_s}(v)\right)\left(\mathbb{1}\left\{V_{ik_t}\le v\right\}-\hat{G}_{k_t}(v)\right)$$

We can now give rates for the bound α_{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 4.1. *Suppose that for the regularized estimator in equation (4.1), α_{nk} satisfies $\limsup_n \alpha_{nk} n^{-\lambda} = 0$ for all $k=k_1,\dots,k_r$, where $\lambda = \frac{3k^*-2}{4k^*-2}\frac{k^*-1}{k^*}$ and $k^* := \max\{k, K-k+1\}$. Then the estimator $\widehat{F}_{nk}(\tau)$ satisfies*

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

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 $v_1 \le v_2$. Furthermore, we have that the estimator $\widehat{F}_n$

$$\sqrt{n}\hat{S}_n(v)^{-\frac{1}{2}}\left(\widehat{F}_n(v)-F_0(v)\right) \xrightarrow{d} N(0,1)$$

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

Note in particular that the rate on α_{nk} implies that the "pasting points", $\tau^*_{1k}(\alpha_{nk})$ and $\tau^*_{2k}(\alpha_{nk})$, converge to zero more slowly than the rates needed to achieve the optimal rate of convergence for the estimator derived in Section 3. Also, the rate on α_{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.

The results in Theorem 4.1 can also be used 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, many functionals of F_0 that are of economic interest are generally not linear, and we will leave the distribution theory for those cases for future research.

5. Functionals of the Valuation Distribution

In empirical research on auctions, the distribution of valuations is 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 are going to 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 $\tilde{K}$.

5.1. **Linear Functionals.** Consider linear functionals of the valuation distribution

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

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

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

Assumption 5.1. *(i) There are $\underline{\tau} \in [0,1]$ and $\bar{\tau} \in [0,1]$ such that $\omega(\tau; F)$ does not change sign on $[0, \underline{\tau}]$ or $[\bar{\tau}, 1]$. (ii) Furthermore, there exist constants β_1, β_2, such that for all $F \in \mathcal{F}_0$ the behavior of $\omega(\tau; F)$ is described by*

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

We can also state this condition in terms of primitive assumptions on the p.d.f.: by the chain rule, $\omega'(\tau; F) = \frac{d}{d\tau} w\left(F^{-1}(\tau)\right) = \frac{w'\left(F^{-1}(\tau)\right)}{h(\tau; F)}$, so that β_1 depends implicitly on the tail behavior of $h(\tau; F)$ given in Assumption 2.3.

Proposition 5.1. *Suppose Assumptions 2.1, 2.2, and 5.1 hold. Then $r_n = n^{-\max\left\{\frac{1}{2}, \frac{2+p+\beta_1}{k_1+p}, \frac{2+p+\beta_2}{K-k_r+1+p}\right\}}$ is an upper bound to the rate of convergence for estimating the linear functional $T(F) = \int_0^\infty w(v)F(dv) = \int_0^1 F^{-1}(s)\omega(s; F)ds$.*

For example, suppose that we observe the transaction price of n i.i.d. second-price auctions with K bidders, and that we are interested in estimating the expectation of V_i, $w(v) \equiv v$ for all values of v. Hence, $\omega'(\tau; F) = \frac{1}{h(\tau;F)}$. Hence if the support of V_i is bounded and the p.d.f. $f(v)$ is bounded away from zero, $\beta_1 = \beta_2 = 0$. Then if in addition $K \geq 5$, by Proposition 5.1, a nonparametric estimator for the expectation of V_i can at best achieve the rate $r_n = n^{-\frac{2}{K-1}}$.

On the other hand, if we observe all K bids for each auctions, as e.g. 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.

5.2. **Expected Revenue.** Next we are going to perform the following thought experiment: suppose we observe the kth highest bid from n repeated sealed bid second-price auctions of K bidders with independent private values, and based on this data we want to predict expected revenue, i.e. the expectation of the second highest bid, for an auction of the same format with $\tilde{K}$ bidders. Clearly if $\tilde{K} = K$ and $k = K - 1$, i.e. we observe the second-highest bid for the observed auctions, 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 e.g. 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 kth highest out of $\tilde{K}$ bids based on observations of the $k_1, \ldots, k_r$th highest bids out of K bidders:

Proposition 5.2. *Suppose Assumptions 2.1-2.3 hold. Then* $r_n = n^{-\max\left\{\frac{1}{2}, \frac{k(1+p)+1-\alpha_1}{k_1+p}, \frac{(\tilde{K}-k)(1+p)+1-\alpha_2}{K-k_r+1+p}\right\}}$ *is an upper bound to the rate of convergence for estimating the expectation of the kth 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 doesn't rule out the possibility that expected revenue can be estimated at 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.

5.3. **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^* maximizing 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

$$\pi(p;F) \;=\; v_0F^n(p) + n\int_p^\infty \left(vf(v) - (1-F(v))\right)F^{n-1}(v)dv$$

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 3.3, the estimator $\hat{F}_n$ proposed in section 3 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 $\pi(p,F)$,

$$\begin{aligned} 0 = &\; \frac{d}{dp}\pi(p;F) = n(v_0-p)F^{n-1}(p)f(p) + n(1-F(p))F^{n-1}(p) \\ \Leftrightarrow \quad p = &\; v_0 + \frac{1-F(p)}{f(p)} \end{aligned}$$

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 $\frac{1}{4}\tau^{p+1}$. Since by Lemma B.1 in the appendix, 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 $\tau_{1n} := \bar{\tau}n^{-\min\left\{\frac{1}{2},\frac{1}{K-2+p}\right\}}$. It is also immediate that the rate cannot be faster than root-n, whereas the possibility of perturbations on the upper tail does not impose further restrictions on the rate.

We can now state this observation as a formal result:

Proposition 5.3. *Suppose Assumptions 2.1-2.3 hold and that the seller's valuation is $v_0 \le \underline{v}$. Then without further restrictions on $\mathcal{F}_0$, $r_n = n^{-\min\left\{\frac{1}{2},\frac{p+1}{k_1+p}\right\}}$ 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 $r_n = n^{-\frac{1}{2}}$.

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, e.g. 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.

Shape restrictions on the seller's surplus function can also be helpful to obtain faster rates for estimators of the optimal reserve price: Suppose now that $\pi(p, F)$ is concave in p for all $F \in \mathcal{F}_0$, and that 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 $n^{1/2}$ 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 doesn't work for estimators that do not impose concavity on $\widehat{\pi_n(p;F)}$ in a given sample.

6. First-Price and Descending Bid Auctions

So far, all our results were 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 descending sequence of prices, and the object is won by the first bidder willing to accept current price. In particular the remaining $K - 1$ potential buyers do not reveal their type, so that if bidding strategies are strictly monotone, only the bid corresponding 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 are going to 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

$$b(v;F) := v - \frac{1}{K-1} \frac{F(v) b'(v;F)}{f(v)} \tag{6.1}$$

We will now replace the model for the second-price sealed-bid auction from Assumption 2.1 with a new assumption

Assumption 6.1. (First-Price Auction) *Assumption 2.1 (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 satisfying (6.1).*

Now denote $g(v;F) := \frac{f\left(b^{-1}(v;F)\right)}{b'(v;F)}$, the marginal distribution of bids with the corresponding c.d.f. $G(v;F) = F\left(b^{-1}(v;F)\right)$, so that we can rewrite equation (6.1) 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 (3.1) and (3.2):

Proposition 6.1. *Let $\hat{F}_n$ be an estimator for F_0. Then under Assumptions 2.2,2.3, and 6.1, $r_n = n^{-\lambda}$ 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 can be strengthened to $r_n^* = \max\left\{\left(\frac{n}{\log n}\right)^{-\frac{p}{2p+1}}, n^{-\frac{1+p}{k_1+p}}, n^{-\frac{p}{K-k_r+1+p}}\right\}$ 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 imputation for the shedding factor in (6.2) only affects the overall bound of 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 quasi-valuations $b^{-1}(B_{iK};F)$ of the highest bidders. We now give a brief explanation 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 c.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.

$g_K(v;F)$ and the c.d.f. $G_K^K(v;F)$ of the *observed* bid

$$v_i := b^{-1}(b_i;F) = b_i + \frac{1}{K-1}\frac{G(b_i;F)}{g(b_i;F)} = b_i + \frac{1}{K-1}\frac{G_K^K(b_i;F)}{g_K^K(b_i;F)} \tag{6.2}$$

Hence we can 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 quasi-valuations.

7. 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 there may be a lot to be gained from combining different bids or data from auctions of different sizes.

Alternatively, since the slow rates are driven entirely by the difficulty in estimating the tails of the distribution of valuations, 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 in 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. While we do not derive convergence rates for estimators imposing these shape restrictions, we conjecture that the bounds on the rate under smoothness restrictions $p \geq 1$ derived in section 3 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 $(k_1,\dots,k_r)$th order statistics is given by[2]

$$g^K_{k_1,\dots,k_r}(\mathbf{v};F) = N(k_1,\dots,k_r;K)[F(v_{k_1})]^{k_1-1}f(v_{k_1})[F(v_{k_2})-F(v_{k_1})]^{k_2-k_1-1}f(v_{k_2})\dots[1-F(v_{k_r})]^{K-k_r}f(v_{k_r})$$

[2]see e.g. David and Nagaraja (2003), p.12

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

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

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

Appendix B. Proofs for Upper Bounds on the Rate of Convergence

We will use the notation $\lesssim$ for "smaller than up to a universal constant." Choose some $\tau_0\in(0,1)$ and define $B:=\frac{\sup_{v\in\mathcal{V}} f_V(v)}{\sup_{v\in\mathcal{V}} f_V(v)}$ which is finite by Assumption 2.3. Also let $\alpha_n=an^{-1/2}$ for some positive $a<\min\{\tau_0,1-\tau_0\}$, $\tau_{1n}=\tau_0 n^{-\frac{1}{k_1+p}}$ and $\tau_{2n}=\tau_0 n^{-\frac{1}{K-k_r+p}}$ be sequences of numbers between zero and one that converge to zero.

Let $\psi(t)$ be a nonnegative function with support $\left[0,\frac{1}{2}\right]$ with $\int_0^{\frac{1}{2}}\psi(v)\,dv=\frac{1}{2}$, $\sup_{t\in\mathbb{R}}\psi(t)<\frac{1}{B}$, and whose first p derivatives are bounded uniformly in t. In order to obtain an upper bound to the rate of convergence, we will consider perturbations of the true p.d.f. that are of the form

$$f_{jn}(v):=f_0(v)\,[1+\psi_{jn}(v)] \tag{B.1}$$

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

$$\begin{aligned} \psi_{1n}(v): &= \tau_{1n}^p\left(\psi\left(F_0^{-1}\left(\tau_{1n}^{-1}\left(\frac{\tau_{1n}}{2}-F_0(v)\right)\right)\right)-\beta_{1n}\psi\left(F_0^{-1}\left(\tau_{1n}^{-1}\left(F_0(v)-\frac{\tau_{1n}}{2}\right)\right)\right)\right) \\ \psi_{2n}(v): &= \tau_{2n}^p\left(\psi\left(F_0^{-1}\left(\tau_{2n}^{-1}\left(F_0(v)-1+\frac{\tau_{2n}}{2}\right)\right)\right)-\beta_{2n}\psi\left(F_0^{-1}\left(\tau_{2n}^{-1}\left(1-\frac{\tau_{2n}}{2}-F_0(v)\right)\right)\right)\right) \\ \psi_{3n}(v): &= \alpha_n\left(\psi\left(F_0(v)-\tau_0\right)-\beta_{3n}\psi\left(\tau_0-F_0(v)\right)\right) \end{aligned}$$

and the sequences β_{jn} are chosen in a way such that $\int\psi_{jn}(v)f_0(v)dv=0$. Note that $\frac{1}{B}\le\beta_{jn}\le B$ for all n so that $f_{jn}(v)=f_0(v)\,[1+\psi_{jn}(v)]$ is 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 non-negative 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 semi-metric 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.[3]

Lemma B.1. *Consider perturbations F_{1n} and F_{2n} of the c.d.f. $F_0(v)$ that are of the form as in equation (B.1). Suppose that for constants $\gamma_1,\gamma_2,\gamma_3>0$, $\varrho(F_{1n},F_0)\gtrsim\tau_n^{\gamma_1}$, $\varrho(F_{2n},F_0)\gtrsim\tau_n^{\gamma_2}$, and $\varrho(F_{3n},F_0)\gtrsim\alpha_n^{\gamma_3}$ for all*

[3] A semi-metric 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) $\varrho(x,z)\le\varrho(x,y)+\varrho(y,z)$.

$\tau_0 \in (0,1)$ and $\delta < \tau_0$. Then under Assumptions 2.1 and 2.2

$$\limsup_n P\left(\varrho(\hat{F}_n, F_0) > cn^{-\max\left\{\frac{\gamma_3}{2}, \frac{\gamma_1}{k_1+p}, \frac{\gamma_2}{K-k_r+p}\right\}}\right) > 0$$

and

$$\lim_{c\to 0}\limsup_n P\left(\varrho(\hat{F}_n, F_0) > cn^{-\max\left\{\frac{\gamma_3}{2}, \frac{\gamma_1}{k_1+p}, \frac{\gamma_2}{K-k_r+p}\right\}}\right) = 1$$

PROOF: Consider local alternatives of the form $f_n(v) = f_{1n}(v)$ as defined in equation (B.1). The c.d.f. $F_{1n}(v)$ corresponding to $f_{1n}(v)$ is given by

$$F_{1n}(v) := \int_0^v f_0(s)(1+\psi_{1n}(s))ds$$

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

In order to construct the likelihood ratio, note that

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

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

$$1 - B\tau_{1n}^p \le \frac{F_{1n}(v_{k_t}) - F_{1n}(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})} \le 1 + B\tau_{1n}^p \tag{B.2}$$

In order to avoid an additional case distinction, we will define $k_0 := 0$, $k_{r+1} := K+1$, $V_{i0} := \inf \mathcal{V}$, and $V_{ir+1} := \sup \mathcal{V}$. 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 $F_0(V_{i0}) = 0$ and $F_0(V_{ir+1}) = 1$.

Now by Assumptions 2.1 and 2.2, the likelihood ratio for an observation $(V_{k_1}, \dots, V_{k_r})$ is given by the Radon-Nykodym derivative

$$\begin{aligned} L_n(V_{ik_1}, \dots, V_{ik_r}) &= \frac{dG(V_{ik_1}, \dots, V_{ik_r}; F_n)}{dG(V_{ik_1}, \dots, V_{ik_r}; F_0)} \\ &= \left(\frac{F_n(V_{ik_1})}{F_0(V_{ik_1})}\right)^{k_1-1} \left(\frac{F_n(V_{ik_2}) - F_n(V_{ik_1})}{F_0(V_{ik_2}) - F_0(V_{ik_1})}\right)^{k_2-k_1-1} \cdots \left(\frac{1-F_n(V_{ik_r})}{1-F_0(V_{ik_r})}\right)^{K-k_r} \prod_{s=1}^r \frac{f_n(V_{ik_s})}{f_0(V_{ik_s})} \\ &= \prod_{s=0}^r \left(\frac{F_n(V_{ik_{s+1}}) - F_n(V_{ik_s})}{F_0(V_{ik_{s+1}}) - F_0(V_{ik_s})}\right)^{k_{s+1}-k_s-1} \prod_{s=1}^r \frac{f_n(V_{ik_s})}{f_0(V_{ik_s})} \end{aligned}$$

Now define the random variables $\chi_{is1} := \mathbb{1}\left\{F_0(V_{ik_s}) < \frac{\tau_{1n}}{2}\right\}$ and $\chi_{is2} := \mathbb{1}\left\{\frac{\tau_{1n}}{2} \le F_0(V_{ik_s}) < \tau_n\right\}$ for $s = 1, \dots, r$, and set $\chi_{i01} := \chi_{i11}$ and $\chi_{i02} = \chi_{ir+11} = \chi_{ir+12} = 0$. Taking logs, we obtain

$$\begin{aligned} l_n(V_{ik_1}, \dots, V_{ik_r}) &:= \log\left(L_n(V_{ik_1}, \dots, V_{ik_r})\right) \\ &= \sum_{s=1}^{r} \Big\{ \chi_{is1} \log\left(1 + \tau_{1n}^p \psi\left(\tau_{1n}^{-1} F_0^{-1}\left(\frac{\tau_{1n}}{2} - F_0(v)\right)\right)\right) \\ &\quad + \chi_{is2} \log\left(1 - \beta_{1n}\tau_{1n}^p \psi\left(\tau_{1n}^{-1} F_0^{-1}\left(F_0(v) - \frac{\tau_{1n}}{2}\right)\right)\right)\Big\} \\ &\quad + \sum_{s=0}^{r} (k_{s+1} - k_s - 1)(\chi_{is1} + \chi_{is2}) \log\left(\frac{F_n(v_{k_t}) - F_n(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})}\right) \\ &= \sum_{s=1}^{r} \Big\{ \chi_{is1} \tau_{1n}^p \psi\left(\tau_{1n}^{-1} F_0^{-1}\left(\frac{\tau_{1n}}{2} - F_0(v)\right)\right) - \chi_{is2}\beta_{1n}\tau_{1n}^p \psi\left(\tau_{1n}^{-1} F_0^{-1}\left(F_0(v) - \frac{\tau_{1n}}{2}\right)\right)\Big\} \\ &\quad + \sum_{s=0}^{r} (k_{s+1} - k_s - 1)(\chi_{is1} + \chi_{is2}) \log\left(\frac{F_n(v_{k_t}) - F_n(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})}\right) + o_P\left(\tau_{1n}^p \max\{1, \beta_{1n}\}\right) \end{aligned}$$

from a Taylor expansion of the log around 1.

From equation (B.2), we can see that if $\chi_{is1} = \chi_{is2} = 0$ for all $s = 1, \dots, r$, $l_n(V_{ik_1}, \dots, V_{ik_r}) = 0$, so that the ith observation only contributes to the likelihood ratio if $V_{i1} < \tau_{1n}$. Also by inspection, we can bound $\left|\log\left(\frac{F_n(v_{k_t}) - F_n(v_{k_s})}{F_0(v_{k_t}) - F_0(v_{k_s})}\right)\right| \le B\tau_{1n}^p$ for $v_{k_t} > v_{k_s}$ and any $\tau_n \in [0, 1]$. Hence, for any realization of $(V_{ik_1}, \dots, V_{ik_r})$, $|l_n(V_{ik_1}, \dots, V_{ik_r})| \le (K+1)B\tau_{1n}^p$.

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

$$\begin{aligned} \mathbb{E}_{F_0}\left|\sum_{i=1}^{n} l_n(V_{ik_1}, \dots, V_{ik_r})\right| &= \int_{(\mathcal{V}^r)^n} \left|\sum_{i=1}^{n} l_n(v_{ik_1}, \dots, v_{ik_r})\right| \otimes_{i=1}^{n} dG(v_{ik_1}, \dots, v_{ik_r}; F_0) \\ &\le n(K+1)B\tau_{1n}^p \mathbb{E}_{F_0}\left[\sum_{s=0}^{r} (\chi_{is1} + \chi_{is2})(k_{s+1} - k_s)\right] \\ &\le 2n(K+1)B\tau_{1n}^p \sum_{s=1}^{r} k_s \tau_{1n}^{k_s} \qquad \text{(B.3)} \end{aligned}$$

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 $\limsup_n \tau_{1n} n^{\frac{1}{k_1+p}} < \infty$, we have

$$\limsup_n \mathbb{E}_{F_0}|\log(L_n)| < C \qquad \text{(B.4)}$$

for a positive constant $C < \infty$. Similarly, for a perturbation of the type $f_n(v) = f_{2n}(v)$ we need $\limsup_n \tau_{2n} n^{\frac{1}{K-k_r+p}} < \infty$ for (B.4) to hold.

Using (B.4), 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 are now going to re-state his argument: suppose the rate τ_n was not an upper bound on the rate of convergence. Then there would be a statistical procedure to decide between $f_n(v)$ and $f_0(v)$ such that the limsup of the probability of a statistical error is equal to zero.

In particular if we put prior probability $\frac{1}{2}$ on each f_n and f_0, the posterior probability of f_n sis

$$\pi\left(f = f_n|\{\mathbf{V}_1,\dots,\mathbf{V}_n\}\right) = \frac{L_n}{1+L_n}$$

Given the constant in equation (B.4), choose $\varepsilon = (1+\exp(C/2))^{-1} > 0$. Then by (B.4) and the Markov Inequality

$$\begin{aligned}\mathbb{P}\left(\varepsilon < \pi\left(f = f_n|\{\mathbf{V}_1,\dots,\mathbf{V}_n\}\right) < 1-\varepsilon\right) &= \mathbb{P}\left(\varepsilon < \frac{L_n}{1+L_n} < 1-\varepsilon\right) \\ &= \mathbb{P}\left(\frac{1}{1+\exp(C/2)} < \frac{L_n}{1+L_n} < \frac{\exp(C/2)}{1+\exp(C/2)}\right) \\ &= \mathbb{P}\left(\exp(-C/2) < L_n < \exp(C/2)\right) \\ &\geq \mathbb{P}\left(|\log L_n| < C/2\right) \geq 1 - \frac{\mathbb{E}|\log L_n|}{C}\end{aligned} \tag{B.5}$$

Hence, taking limits

$$\liminf_n \mathbb{P}_{F_0}\left(\varepsilon < \pi\left(f = f_n|\{\mathbf{V}_1,\dots,\mathbf{V}_n\}\right) < 1-\varepsilon\right) \geq 1 - \limsup_n \frac{\mathbb{E}|\log L_n|}{C} > \frac{1}{2}$$

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

$$\begin{aligned}P\left(\varrho(\hat{F}_n, F_0) > cn^{-\frac{\gamma_1}{k_1+p}}\right) &\geq \frac{1}{2}\mathbb{P}_{F_0}\left(\frac{1}{2}\varrho(\hat{F}_n, F_0) > cn^{-\frac{\gamma_1}{k_1+p}}\right) + \frac{1}{2}\mathbb{P}_{F_{1n}}\left(\frac{1}{2}\varrho(\hat{F}_n, F_0) > cn^{-\frac{\gamma_1}{k_1}}\right) \\ &\geq \frac{1}{2}\mathbb{P}_{F_0}\left(\frac{1}{2}\varrho(F_n, F_0) > c\tau_{1n}^{\gamma_1}\right) + \frac{1}{2}\mathbb{P}_{F_n}\left(\frac{1}{2}\varrho(F_n, F_0) > c\tau_{1n}^{\gamma_1}\right) \\ &\geq \frac{1}{2}\mathbb{P}_{F_0}(\delta_n(\hat{F}_n) = F_{1n}) + \frac{1}{2}\mathbb{P}_{F_n}(\delta_n(\hat{F}_n) = F_{1n}) > \frac{\varepsilon}{4}\end{aligned} \tag{B.6}$$

Applying the same argument to the perturbation F_{2n}, we also obtain

$$\liminf_n P_{F_0}\left(\varrho(\hat{F}_n, F_0) > cn^{-\frac{\gamma_2}{K-k_r+p}}\right) > \frac{\varepsilon}{4} \tag{B.7}$$

Finally, consider a perturbation of $F_0(v)$ in the interior of $\mathcal{V}$ that is of the form F_{3n}. Note that the corresponding statistical experiment $L(\mathbf{V}, \alpha)$ is differentiable with respect to α in quadratic mean, so that by a mean-value expansion around $\alpha_0 = 0$, under F_0 the log-likelihood satisfies

$$\begin{aligned}\sum_{i=1}^n l_n(V_{ik_1},\dots,V_{ik_r};\alpha_n) &= \log L_{3n}(V_{ik_1},\dots,V_{ik_r};\alpha_n) \\ &= 0 + \alpha_n \sum_{i=1}^n \frac{\partial}{\partial\alpha}\log L_{3n}(V_{ik_1},\dots,V_{ik_r};0) + \frac{\alpha_n^2}{2}\frac{\partial^2}{\partial\alpha^2}\log L_{3n}(V_{ik_1},\dots,V_{ik_r};\bar{\alpha}_n) \\ &= O_P\left(\sqrt{n}\alpha_n\right) + n\alpha_n^2 O_P(1)\end{aligned} \tag{B.8}$$

where $\bar{\alpha}_n \in [0, \alpha_n]$ for all n. The score identity can also be verified in a tedious calculation which will be omitted.

Now choose a sequence α_n such that $\limsup_n \sqrt{n}\alpha_n < \infty$. Then the log-likelihood ratio for an observation $(V_{ik_1}, \dots, V_{ik_r})$ satisfies

$$\limsup_n \mathbb{E}_{F_0} \left| \sum_{i=1}^n l_{3n}(V_{ik_1}, \dots, V_{ik_r}) \right| = \limsup_n \mathbb{E}_{F_0} \left| O_P\left(\sqrt{n}\alpha_n\right) + n\alpha_n^2 O_P(1) \right| < \infty$$

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

$$\liminf_n P_{F_0}\left(\varrho(\hat{F}_n, F_0) > cn^{-\frac{\gamma_3}{2}}\right) > \frac{\varepsilon}{4} \tag{B.9}$$

Taken together, (B.6), (B.7), and (B.9) establish the first assertion of the Lemma.

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

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

for some $M > 1$ and $m = 0, \dots, M-1$. Again following the reasoning in Stone (1980) and adapting the arguments leading to (B.3) and (B.5), we can show that the overall error probability of any procedure $\delta_{nM} : \mathcal{F}_0 \to \{F_{n0,M}, \dots, F_{nM,M}\}$ of classifying F into the M points based on $\hat{F}_n$ is at least $1 - \frac{2}{M}$ can be bounded from below by

$$\liminf_n P_{F_0}\left(\varrho(\hat{F}_n, F_0) > cn^{-\frac{\gamma_1}{k_1+p}}\right) \geq \liminf_n P_{F_0}(\delta_{nM}(\hat{F}_n) \neq F_0) > 1 - \frac{2}{M}$$

Since for any $M > 1$ we can pick $c > 0$ small enough such that for large n, $\varrho(F_{n1,M}, F_0) > cn^{-\frac{\gamma_1}{k_1+p}}$, 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 $\square$

B.1. **Proof of Theorem 3.1.** : Without loss of generality, consider the case $K - k_r \leq k_1$. Then by Lemma B.1,

For part (a), note that by Assumption 2.3 for the local alternatives defined in (B.1), $\sup_{v\in\mathcal{V}} |F_{1n}(v) - F_0(v)| \geq \frac{\tau_{1n}^{p+1} \inf_{v\in\mathcal{V}} f_V(v)}{2}$, $\sup_{v\in\mathcal{V}} |F_{2n}(v) - F_0(v)| \geq \frac{\tau_{2n}^{p+1} \inf_{v\in\mathcal{V}} f_V(v)}{B}$, 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 B.1 equations (3.1) and (3.2) hold with $r_n = cn^{-\max\left\{\frac{1}{2}, \frac{p+1}{k_1+p}, \frac{p+1}{K-k_r+p}\right\}}$.

Next we will establish part (b). From the definition of $\psi(\cdot)$ and the lower bound on the density $f_0(v)$ from Assumption 2.3, there exist $\eta_1, \eta_2 \in \left(0, \frac{1}{2}\right)$ 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

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

for small values of τ_{1n} using the rates imposed in Assumption 2.3. Hence,

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

We can apply an analogous argument to the perturbations F_{2n} and F_{3n} so that by Lemma B.1, conditions (3.1) and (3.2) hold with $r_n = cn^{-\max\left\{\frac{1}{2}, \frac{q(1+p)-\alpha_1+1}{q(k_1+p)}, \frac{q(1+p)-\alpha_2+1}{q(K-k_r+1+p)}\right\}}$.

Part (c) follows immediately from part (b) noticing that 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 $\square$

Appendix C. Proof of Consistency Results

C.1. **Proof of Theorem 3.2:** Denote

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

Since $\psi(\tau)$ is differentiable, at every $v \in \mathcal{V}$ a mean-value expansion gives

$$\psi\left(\widehat{G}_n(v)\right) - \psi\left(G_0(v)\right) = \psi'(\overline{G}_n(v))\left(\widehat{G}_n(v) - G_0(v)\right) \tag{C.1}$$

where $\overline{G}_n(v)$ is an intermediate value between $\widehat{G}_n(v)$ and $G_0(v)$. Note that 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 $s = 1, \dots, S$ with positive probability.

For a given choice of $\eta > 1$, we will therefore partition the sample space by defining the event

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

Also denote the event

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

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

In order 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

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

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

We can bound the norm of the envelope function by

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

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 $[l, u] := \{f : l(v) \leq f(v) \leq u(v) \text{ for all } v \in \mathcal{V}\}$ with $\|u - l\| < \varepsilon$ with respect to a norm $\|\cdot\|$ needed to cover $\mathcal{F}$. Also let the entropy integral

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

for any $\delta > 0$.

For the class $\mathcal{F}_{ns}$, we can construct ε-brackets of the form $[l, u]$ with $l(v) := \frac{\mathbf{1}\{v \le v_U\}}{\min_{v\in\{v_L,v_U\}} |G_0(v)-\tau_s^*|}$ and $u(v) := \frac{\mathbf{1}\{v \le v_L\}}{\min_{v\in\{v_L,v_U\}} |G_0(v)-\tau_s^*|}$ where $v_L < v_U$ satisfies $|\psi'(v_L)|^2(G_0(v_U) - G_0(v_L)) < \varepsilon^2$. We will 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 $\|f\|_{P,2}$ 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 $v_{U1} \ge \varepsilon G_0(v_{L1})\|F_{1n}\|_{P,2} = \varepsilon n^{\frac{1}{2}-\frac{1}{2}} = \varepsilon$. Hence the upper bound for the next higher bracket does not decrease in n. Hence we can bound the bracketing number by $N_{[]}(\varepsilon\|F_{sn}\|_{P,2}, \mathcal{F}_{sn}, \|\cdot\|_{P,2}) \le 1 + \frac{1}{\varepsilon}$, so that

$$J_{[]}(1, \mathcal{F}_{sn}, L_2(P)) = \int_0^1 \sqrt{1 + N_{[]}(\varepsilon\|F_{sn}\|_{P,2}, \mathcal{F}_{sn}, \|\cdot\|_{P,2})} d\varepsilon \le \int_0^1 \sqrt{1 + \log(3) - \log(\varepsilon)} d\varepsilon < \infty$$

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

Using Theorem 2.14.2 in van der Vaart and Wellner (1996), we can now bound

$$\mathbb{E}^* \left| \sup_{v\in\mathcal{V}_n(\eta)} \frac{\widehat{G}_n(v) - G_0(v)}{G_0(v) - \tau_s^*} \right| \lesssim n^{-1/2} J_{[]}(1, \mathcal{F}_{sn}, L_2(P)) \|F_{sn}\|_{P,2} \tag{C.3}$$

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

$$\begin{aligned}
P\left(\mathcal{A}_n^C(\eta)\right) &= 1 - P\left(\sup_{v\in\mathcal{V}_n(\eta)} \frac{|G_0(v) - \tau_s^*|}{\left|\widehat{G}_n(v) - \tau_s^*\right|} \le \eta \text{ for all } s = 1, \ldots, S \right) \\
&\le \sum_{s=1}^S P\left(\inf_{v\in\mathcal{V}_n(\eta)} \frac{\left|\widehat{G}_n(v) - G_0(v) + G_0(v) - \tau_s^*\right|}{|G_0(v) - \tau_s^*|} \le \frac{1}{\eta} \right) \\
&\le \sum_{s=1}^S \left\{ P\left(\inf_{v\in\mathcal{V}_n(\eta)} \frac{\left|\widehat{G}_n(v) - G_0(v)\right|}{|G_0(v) - \tau_s^*|} \le \frac{1}{\eta} - 1 \right) \right. \\
&\quad \left. + \ P\left(\operatorname{sign}(\widehat{G}_n(v) - \tau_s^*) \ne \operatorname{sign}(G_0(v) - \tau_s^*) \text{ for some } v \in \mathcal{V}_n(\eta) \right) \right\} \\
&\le 2\sum_{s=1}^S P\left(\inf_{v\in\mathcal{V}_n(\eta)} \frac{\left|\widehat{G}_n(v) - G_0(v)\right|}{|G_0(v) - \tau_s^*|} \le -\frac{\eta - 1}{\eta} \right) \\
&\lesssim \frac{2SJ_{[]}(1, \mathcal{F}_{sn}, L_2(P))}{(\eta - 1)}
\end{aligned} \tag{C.4}$$

where the last step uses Markov's inequality together with (C.3) and (C.2). This bound on the probability can be made arbitrarily small by choosing a sufficiently large value of $\eta > 1$.

Next, we will bound the probability of $\mathcal{B}_n(\eta)$. conditional on $\mathcal{A}_n(\eta)$. First note that by monotonicity of $\psi'(\tau)$ between the critical points τ_s^*, $s = 1, \dots, S$ we can bound

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

conditional on $\mathcal{A}_n(\eta)$ for all $v \in \mathcal{V}_n(\eta)$.

Now define the class of functions

$$\mathcal{H}_n(\eta) := \left\{ \psi'(\eta^{-1}G_0(t))\mathbb{1}\{v \leq t\} \middle| t \in \mathcal{V}_n(\eta)\right\}$$

with envelope function $H_n(\eta) := |\psi'(\eta^{-1}G_0(v))|$.

Noting that for any exponent $\delta_s > 0$ in Condition 3.1, $|\psi'(\tau)|$ is dominated by $\frac{1}{\tau - \tau_s^*}$ for values of τ close to τ^*, we can use the same reasoning as before in order to establish that the bracketing integral $J_{[]}(1, \mathcal{H}_{ns}, L_2(P))$ is bounded. In order to bound the norm of the corresponding envelope functions, let without loss of generality $\delta_s < 1$. Then for n sufficiently large, by Condition 3.1 we can bound

$$\int_{\tau_s^*+\eta n^{-1}}^{\frac{\tau_s^*+\tau_{s+1}^*}{2}} |\psi'(s)|^2 ds \leq 2A_s \int_{\tau_s^*+\eta n^{-1}}^{\frac{\tau_s^*+\tau_{s+1}^*}{2}} |s - \tau_s^*|^{2\delta_s - 2} ds \leq 2A_s \left(n^{-1}\eta\right)^{2\delta_s - 1}$$

We can now use (C.1), (C.5), and Theorem 2.14.2 in van der Vaart and Wellner (1996) to bound

$$\begin{aligned}
\mathbb{E}\left[\sup_{v\in\mathcal{V}_n(\eta)} \left|\psi\left(\widehat{G}_n(v)\right) - \psi\left(G_0(v)\right)\right| \middle| \mathcal{A}_n(\eta)\right] &\leq \mathbb{E} \sup_{v\in\mathcal{V}_n(\eta)} \left|\psi'(\eta^{-1}G_0(v))\right| \left|\widehat{G}_n(v) - G_0(v)\right| \\
&\leq n^{-1/2} J_{[]}(1, \mathcal{H}_{ns}, L_2(P)) \|H_n(\eta)\|_{P,2} \\
&\leq 2A_s J_{[]}(1, \mathcal{F}_{sn}, L_2(P)) \eta^{\delta_s - \frac{1}{2}} n^{-\delta_s}
\end{aligned}$$

for n large enough.

Hence, using Markov's Inequality together with the law of total probability and (C.4),

$$\begin{aligned}
\mathbb{P}\left(\sup_{v\in\mathcal{V}_n(\eta)} \left|\psi\left(\widehat{G}_n(v)\right) - \psi\left(G_0(v)\right)\right| > cr_n\right) &\leq \mathbb{P}(\mathcal{B}_n(\eta)|\mathcal{A}_n(\eta)) + \mathbb{P}(\mathcal{A}_n^C(\eta)) \\
&\leq \frac{2A_s J_{[]}(1, \mathcal{H}_{sn}, L_2(P))}{c\eta^{1-\delta_s}} + \frac{2SJ_{[]}(1, \mathcal{F}_{sn}, L_2(P))}{(\eta - 1)}
\end{aligned}$$

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

Furthermore, from Condition 3.1 it follows that

$$\sup_{v\in\mathcal{V}\setminus\mathcal{V}_n(\eta)} \min_{\tau\in\{\tau_1^*,\dots,\tau_S^*\}} |\tau - G_0(v)| \leq (\eta n)^{-\delta} \tag{C.6}$$

Since by Condition 3.1, $\psi(\tau)$ is monotone on the intervals $\left[\tau_s^* - \eta n^{-1}, \tau_s^*\right]$ and $\left[\tau_s^*, \tau_s^* + \eta n^{-1}\right]$ for every η and n large enough, (C.6) and (C.6) together imply that conditional on the event $\mathcal{A}_n(\eta) \cap \mathcal{B}_n(\eta)$,

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

which completes the proof $\square$

Appendix D. Proof of Theorem 4.1

Fix a value of k. We are now going to establish that for the estimator $\hat{F}_{nk}$ the conditions of Lemma 4.1 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 3.1, $\psi_k''(\tau) = \frac{\phi_k''(\phi_k^{-1}(\tau))}{\left[\phi_k'(\phi_k^{-1}(\tau))\right]^3}$. Also recall that $\psi_k(\tau)$ behaves like $\tau^{\frac{1}{k}}$ for small values of τ and is approximated by $\tau^{\frac{1}{K-k+1}}$ for values of τ sufficiently close to 1.

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

If $K=1$, $\psi_1''(\tau) = 0$, in which case the approximation in Theorem 4.1 is trivially true without any need for regularization, and we will therefore only consider the case $k \geq 2$ in the remainder of this argument. Since $\frac{\psi_k''(\tau)}{\left[\psi_k'(\tau)\right]^2}$ diverges for $k \geq 2$ as $\tau \to 0$ and for $K-k \geq 2$ as $\tau \to 1$, we can bound the supremum $\sup_{\tau \in \left[\tau_1, \frac{1}{2}\right]} \left|\frac{\psi_k''(\tau)}{\left[\psi_k'(\tau)\right]^2}\right|$ of the ratio by a multiple $2\frac{\tau_1^{\frac{1}{k}-2}}{\tau_1^{\frac{2}{k}-2}} = \tau^{-\frac{1}{k}}$ for τ_1 sufficiently small. A similar argument applies to the upper tail of the distribution.

In the following we can, without loss of generality, restrict our attention to the case in which V_i is uniformly distributed, i.e. $F_0(\tau) = \tau$ 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 $\tau + r_n h_n(\tau)$ is a proper c.d.f. for n large enough, we have by a mean-value expansion in $h(\tau)$

$$\begin{aligned}
R_n(h_n) &:= \sup_{\tau\in[0,1]} \left| r_n^{-1} \left(|\tilde{\psi}_k'(\tau)|\right)^{-1} \left(\tilde{\psi}_k(\tau + r_n h_n(\tau)) - \tilde{\psi}_k(\tau)\right) - \left(|\psi'(\tau)|\right)^{-1} \psi'(\tau) h_n(\tau) \right| \\
&= \sup_{\tau\in[0,1]} \left| r_n \left(|\tilde{\psi}_k'(\tau)|\right)^{-2} \tilde{\psi}_k''(\tau + r_n \bar{h}_n(\tau)) h(\tau)^2 \right| \\
&= \sup_{\tau\in[r_n h_n(0), 1 + r_n h_n(1)]} \left| r_n \left(|\tilde{\psi}_k'(\tau - r_n \bar{h}_n(\tau))|\right)^{-2} \tilde{\psi}_k''(\tau) h(\tau - r_n \bar{h}_n(\tau))^2 \right|
\end{aligned}$$

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

$$\begin{aligned}
R_n(h_n) &\leq \sup_{\tau\in\left[\tau_{1n}, \frac{1}{2}\right]} \left| \frac{(\tau + r_n \bar{h}_n(\tau))^{2-\frac{2}{k}}}{\tau^{2-\frac{1}{k}}} r_n h(\tau - r_n \bar{h}_n(\tau))^2 \right| + \sup_{\tau\in\left[\frac{1}{2}, \tau_{2n}\right]} \left| \frac{(\tau + r_n \bar{h}_n(\tau))^{2-\frac{2}{K-k+1}}}{\tau^{2-\frac{1}{K-k+1}}} r_n h(\tau - r_n \bar{h}_n(\tau))^2 \right| \\
&\leq \sup_{\tau\in\left[\tau_{1n}, \frac{1}{2}\right]} \left| \frac{(\tau + r_n |h_n(\tau)|)^{2-\frac{2}{k}}}{\tau^{2-\frac{1}{k}}} r_n h(\tau - r_n \bar{h}_n(\tau))^2 \right| + \sup_{\tau\in\left[\frac{1}{2}, \tau_{2n}\right]} \left| \frac{(\tau - r_n |h_n(\tau)|)^{2-\frac{2}{K-k+1}}}{\tau^{2-\frac{1}{K-k+1}}} r_n h(\tau - r_n \bar{h}_n(\tau))^2 \right| \\
&\leq 2\left(r_n^{2-\frac{2}{k}} \tau_{1nk}^{\frac{1}{k}-2} + r_n^{2-\frac{2}{K-k+1}} \tau_{2nk}^{\frac{1}{K-k+1}-2} \right) r_n \sup_{\tau\in[0,1]} |h_n(\tau)|^{4\frac{K-1}{K}}
\end{aligned}$$

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, $r_n = n^{-\frac{1}{2}}$, 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 4.1 satisfies Condition 4.1. Hence it follows from Lemma 4.1 that the proposed normalized estimator satisfies the uniform approximation posited in Theorem 4.1 $\square$

Appendix E. Rates for Functionals of the Distribution of Valuations

E.1. **Proof of Proposition 5.1.** Using integration by parts, we can rewrite the functional $T(F)$ at F as

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

Since by Assumption 5.1, $\lim_{\tau\to 0}\omega(\tau;F)$ 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 2.3, there exist $\eta_1,\eta_2\in\left(0,\frac{1}{2}\right)$ 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 5.1, the integrand does not change sign on the interval $[0,\tau_{1n}]$ for n large enough. Hence,

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

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 B.1, , conditions (3.1) and (3.2) hold with $r_n=n^{-\max\left\{\frac{1}{2},\frac{2+p+\beta_1}{k_1+p},\frac{2+p+\beta_2}{K-k_r+1+p}\right\}}$ $\square$

E.2. **Proof of Proposition 5.2.** Note that since by assumption $V>0$ with probability 1,

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

By the same argument as in the proof of Proposition 5.1, we can find $\eta_1,\eta_2\in\left(0,\frac{1}{2}\right)$ 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

$$\begin{aligned}
\mathbb{E}_{F_0}\left[V_{k:\tilde{K}}\right]-\mathbb{E}_{F_{1n}}\left[V_{k:\tilde{K}}\right] &= \int_0^{F_0^{-1}(\tau_n)}\left[G_k^{\tilde{K}}(v;F_n)-G_k^{\tilde{K}}(v;F_0)\right]dv \\
&= \frac{\tilde{K}!}{k!(\tilde{K}-k)!}\int_0^{F_0^{-1}(\tau_n)}\left[F_{1n}(v)^k(1-F_{1n}(v))^{\tilde{K}-k}-F_0(v)^k(1-F_0(v))^{\tilde{K}-k}\right]dv \\
&= \frac{\tilde{K}!}{k!(\tilde{K}-k)!}\int_0^{\tau_{1n}}\left[\left(F_{1n}\left(F_0^{-1}(s)\right)\right)^k\left(1-F_{1n}\left(F_0^{-1}(s)\right)\right)^{\tilde{K}-k}-s^k(1-s)^{\tilde{K}-k}\right]h(s;F_0)^{-1}ds \\
&\geq \frac{\tilde{K}!}{k!(\tilde{K}-k)!}\int_{\eta_1\tau_{1n}}^{\eta_2\tau_{1n}}\left[\left(s+\kappa s^{1+p}\right)k\left(1-s-\kappa s^{1+p}\right)^{\tilde{K}-k}-s^k(1-s)^{\tilde{K}-k}\right]h(s;F_0)^{-1}ds \\
&\gtrsim \int_{\eta_1\tau_{1n}}^{\eta_2\tau_{1n}} s^{k(1+p)}h(s;F_0)^{-1}ds \gtrsim \tau_{1n}^{k(1+p)+1-\alpha_1}
\end{aligned}$$

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

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

and

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

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

E.3. **Proof of Proposition 6.1.** We will in a first step apply a modification of Lemma B.1 to the distribution $G_0(v)$ of a random bid $B_{\hat{K}}$, where the index $\hat{K}$ is drawn at random from a uniform distribution over $\{1,2,\dots,K\}$, the set of all bidders.

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

$$\begin{aligned}
\psi_{1n}(v): &= \tau_{1n}^p\left(\beta_{1n}\psi\left(G_0^{-1}\left[\tau_{1n}^{-1}\left(\frac{\tau_{1n}}{2}-G_0(v)\right)\right]\right)-\psi\left(G_0^{-1}\left[\tau_{1n}^{-1}\left(G_0(v)-\frac{\tau_{1n}}{2}\right)\right]\right)\right) \\
\psi_{2n}(v): &= \tau_{2n}^p\left(\beta_{2n}\psi\left(G_0^{-1}\left[\tau_{2n}^{-1}\left(G_0(v)-1+\frac{\tau_{2n}}{2}\right)\right]\right)-\psi\left(G_0^{-1}\left[\tau_{2n}^{-1}\left(1-\frac{\tau_{2n}}{2}-G_0(v)\right)\right]\right)\right) \\
\psi_{3n}(v): &= \eta_n^p\left\{\psi\left(G_0^{-1}\left[\eta_n^{-1}\left(\tau_0-G_0(v)\right)\right]\right)-\beta_{3n}\psi\left(G_0^{-1}\left[\eta_n^{-1}\left(\tau_0-G_0(v)\right)\right]\right)\right\}
\end{aligned}$$

where for all j the sequence β_{jn} is bounded between $\frac{1}{B}$ and B and 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 B.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 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 B.1 can be modified to

$$\limsup_n P\left(\varrho(\hat{G}_n,G_0)>cn^{-\max\left\{\frac{\gamma_3}{2p+1},\frac{\gamma_1}{k_1+p},\frac{\gamma_2}{K-k_r+p}\right\}}\right)>0$$

and

$$\lim_{c\to 0}\limsup_n P\left(\varrho(\hat{G}_n,G_0)>cn^{-\max\left\{\frac{\gamma_3}{2p+1},\frac{\gamma_1}{k_1+p},\frac{\gamma_2}{K-k_r+p}\right\}}\right)=1$$

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 B.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, $F(b^{-1}(b;F))=\tau$, i.e. the ordering of quantiles is preserved. For perturbation F_{jn}, note that using (6.2) and a mean value expansion, we can write the valuation implied by bid $b=G_0^{-1}(\tau)$ as

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

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 2.3, the bidding function $b(v;F)$ and its inverse are Lipschitz continuous for F_0 and F_{jn}. Using the expansion (E.1),

$$\begin{aligned}F_{jn}\left(b^{-1}\left(b;F_{jn}\right)\right)-F_0\left(b^{-1}\left(b;F_0\right)\right) &= F_{jn}\left(b^{-1}\left(b;F_{jn}\right)\right)-F_0\left(b^{-1}\left(b;F_{jn}\right)\right)\\ &\quad +F_0\left(b^{-1}\left(b;F_0\right)+\left[b^{-1}\left(b;F_{jn}\right)-b^{-1}\left(b;F_0\right)\right]\right)-F_0\left(b^{-1}\left(b;F_0\right)\right)\\ &= \int_0^{b^{-1}(b;F_{jn})}\psi_{jn}(s)g_0(s)ds+\frac{f_0\left(\bar{v}_n\right)}{K-1}\left(\frac{\int_{\inf\mathcal{V}}^b\psi_{jn}(s)g_0(s)ds}{g_0(b)}-\frac{G_0(b)}{\bar{g}_n(b)^2}\psi_{jn}(b)\right)\end{aligned}\tag{E.2}$$

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

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

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

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

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

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

References

Aït-Sahalia, Y., and J. Duarte (2003): "Nonparametric Option Pricing under Shape Restrictions," *Journal of Econometrics*, 116, 9–47.

Aradillas-López, A., A. Gandhi, and D. Quint (2010): "Identification and Testing in Ascending Auctions with Unobserved Heterogeneity," working paper, University of Wisconsin.

Athey, S., and P. Haile (2002): "Identification of Standard Auction Models," *Econometrica*, pp. 2107–2140.

——— (2007): "Nonparametric Approaches to Auctions," in *J. Heckman and E. Leamer,eds., Handbook of Econometrics Vol.6A*, p. 3847.

Chernozhukov, V., S. Lee, and A. Rosen (2008): "Inference on Intersection Bounds," working paper, MIT and UCL.

David, H., and H. Nagaraja (2003): *Order Statistics, 3rd ed.* Wiley and Sons.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525.

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111(1), 1–51.

Ibragimov, I., and R. Has'minskii (1981): *Statistical Estimation - Asymptotic Theory.* Springer, New York.

Khan, S., and E. Tamer (2009): "Irregular Identification, Support Conditions, and Inverse Weight in Estimation," .

Komarova, T. (2009): "Nonparametric Identification in Asymmetric Second-Price Auctions: A New Approach," working paper, LSE.

Mammen, E., and C. Thomas-Agnan (1999): "Smoothing Splines and Shape Restrictions," *Scandinavian Journal of Statistics*, 26, 239–252.

Riley, J., and W. Samuelson (1981): "Optimal Auctions," *American Economic Review*, 71(3), 381–392.

Rockafellar, R. (1972): *Convex Analysis.* Princeton University Press.

Roehrig, C. (1988): "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56(2), 433–447.

Stone, C. (1980): "Optimal Rates of Convergence for Nonparametric Estimators," *Annals of Statistics*, pp. 1348–1360.

——— (1983): "Optimal Global Rates of Convergence for Nonparametric Regression," *Annals of Statistics*, pp. 1040–1063.

van der Vaart, A. (1998): *Asymptotic Statistics.* Cambridge University Press, Cambridge.

van der Vaart, A., and J. Wellner (1996): *Weak Convergence and Empirical Processes.* Springer, New York.