Supplement to “Managing Procurement Auction Failure: Bid Requirements or Reserve Prices?”

Jun Ma* Vadim Marmer† Pai Xu‡

Abstract

This supplement provides additional theoretical and empirical results, and also econometric details omitted from the main text. Section S1.1 shows lemmas used in the proof of Proposition 3.2. Section S1.2 gives a result on existence of unequal equilibrium entry probabilities. Section S1.3 provides proofs of the results discussed in Section 8. Section S2.1 examines the observed auction heterogeneity in explaining bid variation. Section S2.2 presents counterfactual results under an alternative copula specification. Section S3.1 provides a detailed introduction to our econometric results and reviews the related econometric and statistical literature. Section S3.2 gives identification results. Section S3.3 describes the generalized method of moments (GMM) estimation procedure used in the main text. Section S3.3.1 establishes the asymptotic normality of the GMM estimator (Proposition S4). Section S3.3.2 presents a bootstrap method for estimating the asymptotic variance and establishes its consistency (Proposition S5). Appendix A presents ancillary results (concentration bounds for the local polynomial density estimators). Appendix B contains the proofs of the main results in Section S3.3.

Contents

This version: February 24, 2026.

*School of Economics, Renmin University of China

†Vancouver School of Economics, University of British Columbia

‡Business School, University of Hong Kong

Notation. For any $m \in \mathbb{N}$, denote $[m] \coloneqq \{1, ..., m\}$. For a finite set A, let $|A|$ denote the number of elements in A. Let $\mathbf{1}_m$ ($\mathbf{0}_m$) denote an m-dimensional vector whose elements are all one (zero). Let $\mathbf{I}_m$ denote the m-dimensional identity matrix. Let $a \wedge b$ and $a \vee b$ be shorthand notations for $\min\{a, b\}$ and $\max\{a, b\}$, respectively.

S1 Additional theoretical results

S1.1 Auxiliary lemmas used in the proof of Proposition 3.2

Lemma S1. *Let k and m be positive integers. For any fixed $\varepsilon \in (0,1)$,*

$$\int_0^p u^k (1-u)^{n-m} \, du = \frac{1}{(n-m)^{k+1}} \left(\int_0^\infty t^k \exp(-t) \, dt + \delta_n(p) \right),$$

where $\delta_n(p) \downarrow 0$ uniformly in $p \in [\varepsilon, 1]$.

Proof of Lemma S1. By the change of variables $t = (n-m)u$,

$$\int_0^p u^k (1-u)^{n-m} du = \frac{1}{(n-m)^{k+1}} \int_0^{(n-m)p} t^k \left(1 - \frac{t}{n-m}\right)^{n-m} dt.$$

By the dominated convergence theorem, the integral on the right-hand side of the above equation converges to $\int_0^\infty t^k \exp(-t) \, dt$ uniformly in $p \in [\varepsilon, 1]$ as $n \uparrow \infty$. ■

Lemma S2. *Under the assumptions in the statement of Proposition 3.2,*

$$\int_{\underline{v}}^{\overline{v}} C_2^2(F(v), p) (1 - C(F(v), p))^{n-2} \, dv = \frac{1}{(n-2)^2} \left(\frac{2C_{21}^2(0,p)}{f(\underline{v}) C_1^3(0,p)} + \delta_n(p) \right),$$

where $\delta_n(p) \downarrow 0$ uniformly in $p \in [\varepsilon, 1]$.

Proof of Lemma S2. By change of variables $u = C(F(v), p)$,

$$\int_{\underline{v}}^{\overline{v}} C_2^2(F(v), p)(1 - C(F(v), p))^{n-2} dv = \int_0^p \frac{C_2^2(C^{-1}(u,p), p)}{f(F^{-1}(C^{-1}(u,p))) C_1(C^{-1}(u,p), p)} (1-u)^{n-2} du,$$

where $C^{-1}(\cdot, p)$ denotes the inverse function of $C(\cdot, p)$. Define

$$\begin{aligned} \phi(u,p) &\coloneqq C_2(C^{-1}(u,p), p), \\ \psi(u,p) &\coloneqq f(F^{-1}(C^{-1}(u,p))) C_1(C^{-1}(u,p), p). \end{aligned} \tag{S1}$$

Since $C_2(0,p) = 0$, we have $\phi(0,p) = 0$. Let $\phi_1(u,p) \coloneqq \partial \phi(u,p)/\partial u$, $\phi_{11}(u,p) \coloneqq \partial^2 \phi(u,p)/\partial u^2$, $C_{11}(u,p) \coloneqq \partial^2 C(u,p)/\partial u^2$, and $C_{211}(u,p) \coloneqq \partial C_{21}(u,p)/\partial u$. By differentiation and the fact that

$\partial C^{-1}(u,p)/\partial u = \left(C_1\left(C^{-1}(u,p),p\right)\right)^{-1}$, we have

$$\begin{aligned}
\phi_1(u,p) &= \frac{C_{21}\left(C^{-1}(u,p),p\right)}{C_1\left(C^{-1}(u,p),p\right)}, \\
\phi_{11}(u,p) &= \frac{C_{211}\left(C^{-1}(u,p),p\right)C_1\left(C^{-1}(u,p),p\right) - C_{21}\left(C^{-1}(u,p),p\right)C_{11}\left(C^{-1}(u,p),p\right)}{C_1^3\left(C^{-1}(u,p),p\right)}.
\end{aligned}$$

By second-order Taylor expansion and the fact that $\inf_{(u,p)\in[0,1]\times[\varepsilon,1]} C_1\left(C^{-1}(u,p),p\right) > 0$,

$$\phi(u,p) = \phi_1(0,p)u + \tilde{\phi}(u,p)u^2, \tag{S2}$$

for some function $\tilde{\phi}(\cdot,\cdot)$ bounded on $[0,1]\times[\varepsilon,1]$. Similarly,

$$\psi(u,p) = \psi(0,p) + \tilde{\psi}(u,p)u, \tag{S3}$$

for some function $\tilde{\psi}(\cdot,\cdot)$ bounded on $[0,1]\times[\varepsilon,1]$. Clearly, we have

$$\phi_1(0,p) = \frac{C_{21}(0,p)}{C_1(0,p)} \text{ and } \psi(0,p) = f(\underline{v})C_1(0,p).$$

By this result, Lemma S1, the fact that $\inf_{(u,p)\in[0,1]\times[\varepsilon,1]} \psi(u,p) > 0$, and $\int_0^\infty t^2 \exp(-t)\,dt = 2$,

$$\begin{aligned}
\int_{\underline{v}}^{\overline{v}} C_2^2(F(v),p)(1-C(F(v),p))^{n-2}dv &= \int_0^p \frac{(\phi_1(0,p)u + \tilde{\phi}(u,p)u^2)^2}{\psi(0,p)+\tilde{\psi}(u,p)u}(1-u)^{n-2}du \\
&= \frac{1}{(n-2)^3}\left(\frac{2C_{21}^2(0,p)}{f(\underline{v})C_1^3(0,p)} + o(1)\right).
\end{aligned} \tag{S4}$$

■

S1.2 Existence of unequal equilibrium entry probabilities

Recall that $R(p,n)$ denotes the expected revenue of the marginal entrant. We show the following result.

Lemma S3. *Suppose that Assumption 3.1 holds, and the copula density function is bounded away from zero on* $[0,1]^2$*. Then, for any fixed* $\varepsilon \in (0,1)$,

$$\inf_{p\in[\varepsilon,1]} \{R(p,n_1) - R(p,n_2)\} > 0, \tag{S5}$$

if $n_1 > n_2$ *are sufficiently large.*

Proof of Lemma S3. It suffices to show

$$\inf_{p\in[\varepsilon,1]} \{R(p,n+1) - R(p,n)\} > 0, \tag{S6}$$

if n is sufficiently large. We decompose

$$\begin{aligned}
R(p,n) &= R_{+}(p,n) - R_{-}(p,n), \text{ where} \\
R_{+}(p,n) &:= \int_{\underline{v}}^{\overline{v}} C_2(F(v),p)(1-C(F(v),p))^{n-1}\,dv, \\
R_{-}(p,n) &:= (1-p)^{n-1}\int_{\underline{v}}^{\overline{v}} C_2(F(v),p)\,dv.
\end{aligned}$$

First, note that

$$\begin{aligned}
&\inf_{p\in[\varepsilon,1]}\{R(p,n+1)-R(p,n)\} \geq \\
&\qquad \inf_{p\in[\varepsilon,1]}\{R_{+}(p,n+1)-R_{+}(p,n)\} - \sup_{p\in[\varepsilon,1]}|R_{-}(p,n+1)| - \sup_{p\in[\varepsilon,1]}|R_{-}(p,n)|. \quad \text{(S7)}
\end{aligned}$$

Since under the "good news" assumption, $p\mapsto(1-p)^{n-1}\int_{\underline{v}}^{\overline{v}} C_2(F(v),p)\,dv$ is decreasing, we have

$$\begin{aligned}
\sup_{p\in[\varepsilon,1]}|R_{-}(p,n)| &= \sup_{p\in[\varepsilon,1]}(1-p)^{n-1}\int_{\underline{v}}^{\overline{v}} C_2(F(v),p)\,dv \\
&\leq (1-\varepsilon)^{n-1}\int_{\underline{v}}^{\overline{v}} C_2(F(v),\varepsilon)\,dv, \qquad \text{(S8)}
\end{aligned}$$

where the right-hand side of the inequality decays exponentially as $n\uparrow\infty$. Similarly,

$$\sup_{p\in[\varepsilon,1]}|R_{-}(p,n+1)| \leq (1-\varepsilon)^{n}\int_{\underline{v}}^{\overline{v}} C_2(F(v),\varepsilon)\,dv, \qquad \text{(S9)}$$

where the right-hand side of the inequality decays exponentially as $n\uparrow\infty$. Now note that

$$\begin{aligned}
& R_{+}(p,n+1)-R_{+}(p,n) \\
= \;& \int_{\underline{v}}^{\overline{v}} C_2(F(v),p)(1-C(F(v),p))^{n}\,dv - \int_{\underline{v}}^{\overline{v}} C_2(F(v),p)(1-C(F(v),p))^{n-1}\,dv \\
= \;& \int_{\underline{v}}^{\overline{v}} C_2(F(v),p)\,C(F(v),p)(1-C(F(v),p))^{n-1}\,dv \\
> \;& 0. \qquad \text{(S10)}
\end{aligned}$$

By change of variables $u=C(F(v),p)$,

$$\int_{\underline{v}}^{\overline{v}} C_2(F(v),p)\,C(F(v),p)(1-C(F(v),p))^{n-1}\,dv = \int_0^p \frac{\phi(u,p)}{\psi(u,p)}u(1-u)^{n-1}\,du,$$

where $\phi(u,p)$ and $\psi(u,p)$ are defined by (S1). Then, by (S2), (S3), and Lemma S1, we have

$$\int_{\underline{v}}^{\overline{v}} C_2(F(v),p)\,C(F(v),p)(1-C(F(v),p))^{n-1}dv$$

$$
\begin{aligned}
&= \int_0^p \frac{\phi_1(0,p)u + \tilde{\phi}(u,p)u^2}{\psi(0,p) + \tilde{\psi}(u,p)u} u(1-u)^{n-1} du \\
&= \frac{1}{(n-1)^3} \left(\frac{2C_{21}(0,p)}{C_1^2(0,p) f(\underline{v})} + o(1) \right).
\end{aligned}
$$

By this result and (S10), we have

$$
(n-1)^3 \inf_{p\in[\varepsilon,1]} \{R_+(p,n+1) - R_+(p,n)\} \geq \inf_{p\in[\varepsilon,1]} \frac{2C_{21}(0,p)}{C_1^2(0,p) f(\underline{v})} - o(1),
$$

where

$$
\inf_{p\in[\varepsilon,1]} \frac{2C_{21}(0,p)}{C_1^2(0,p) f(\underline{v})} \geq \frac{2}{f(\underline{v})} \cdot \left(\inf_{p\in[\varepsilon,1]} C_{21}(0,p) \right) > 0.
$$

Now (S6) follows from this result, (S7), (S8) and (S9). ■

Let n_1 and n_2 be sufficiently large such that (S5) holds. We have $\kappa_{n_1} = R(p_{n_1}, n_1)$, $\kappa_{n_2} = R(p_{n_2}, n_2)$. If the entry costs are constant and $p_{n_1}, p_{n_2} \in [\varepsilon, 1]$, it follows from Lemma S3 that we cannot have $p_{n_1} = p_{n_2}$.

S1.3 Soft bid requirement

This section derives the expected profit and expected winning bid under the soft bid requirement format introduced in Section 8. By the same arguments as in the proof of Proposition 3.1, the expected profit of a bidder who receives signal s and enters the auction can be written as

$$
\begin{aligned}
\Pi(p,n,\kappa,s;\theta) &= \int_{\underline{v}}^{\overline{v}} C_2(F(v),s;\theta) H(v \mid p,n;\theta) dv - \kappa \\
&= \int_{\underline{v}}^{\overline{v}} C_2(F(v),s;\theta) \left\{ \Lambda^{n-1}(v \mid p;\theta) - \frac{(1-p)^{n-1}}{p} (1 - \Lambda(v \mid p;\theta)) \right\} dv - \kappa \\
&= \int_{\underline{v}}^{\overline{v}} C_2(F(v),s;\theta) \left\{ (1 - C(F(v),p;\theta))^{n-1} - \frac{(1-p)^{n-1}}{p} C(F(v),p;\theta) \right\} dv - \kappa.
\end{aligned}
$$

In the limit of perfectly informative signals, the expected revenue vanishes:

$$
\begin{aligned}
&\lim_{\theta\uparrow\infty} \Pi(p,n,\kappa,p;\theta) \\
&= \int_{\underline{v}}^{\overline{v}} \mathbb{1}(F(v) \geq p) \left\{ (1 - \min\{F(v),p\})^{n-1} - \frac{(1-p)^{n-1}}{p} \min\{F(v),p\} \right\} dv - \kappa \\
&= -\kappa.
\end{aligned}
$$

The probability of having at least two active bidders or one active bidder who submits a bid lower than a random draw from the bid distribution is

$$\begin{aligned} P(p,n) &:= \left(1-(1-p)^n - np(1-p)^{n-1}\right) + \frac{np(1-p)^{n-1}}{2} \\ &= 1-(1-p)^n - \frac{np(1-p)^{n-1}}{2}. \end{aligned}$$

The equilibrium bidding strategy is still given by

$$\beta(v \mid p,n) := v + \int_v^{\overline{v}} \frac{H(u \mid p,n)}{H(v \mid p,n)} du.$$

The expected winning bid conditional on at least two active bidders or one active bidder who submits a bid lower than a random draw from the bid distribution is

$$\begin{aligned} &\frac{1}{P(p,n)} \left\{ \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} j \int_{\underline{v}}^{\overline{v}} \beta(v \mid p,n)(1-F^*(v \mid p))^{j-1} dF^*(v \mid p) \right. \\ &\left. + \frac{np(1-p)^{n-1}}{2} \cdot \left(2\int_{\underline{v}}^{\overline{v}} \beta(v \mid p,n)(1-F^*(v \mid p)) dF^*(v \mid p) \right) \right\} \\ = &\frac{np}{P(p,n)} \int_{\underline{v}}^{\overline{v}} \beta(v \mid p,n) H(v \mid p,n) dF^*(v \mid p), \end{aligned}$$

where

$$2\int_{\underline{v}}^{\overline{v}} \beta(v \mid p,n)(1-F^*(v \mid p)) dF^*(v \mid p) = \int_{\underline{v}}^{\overline{v}} \beta(v \mid p,n) d\left(1-(1-F^*(v \mid p))^2\right)$$

is the expected winning bid when there is only one active bidder. By the same arguments as in the proof of Proposition 3.3,

$$\begin{aligned} &\int_{\underline{v}}^{\overline{v}} \beta(v \mid p,n) H(v \mid p,n) d\Lambda(v \mid p) \\ = &-\underline{v}H(\underline{v} \mid p,n) - \int_{\underline{v}}^{\overline{v}} H(v \mid p,n) dv - \int_{\underline{v}}^{\overline{v}} \Lambda(v \mid p) d(\beta(v \mid p,n) H(v \mid p,n)) \\ = &-\underline{v} - \int_{\underline{v}}^{\overline{v}} \Lambda^{n-1}(v \mid p) dv + \frac{(1-p)^{n-1}}{p}(\overline{v}-\underline{v}) - \frac{(1-p)^{n-1}}{p} \int_{\underline{v}}^{\overline{v}} \Lambda(v \mid p) dv \\ &- \int_{\underline{v}}^{\overline{v}} \Lambda(v \mid p) v H'(v \mid p,n) dv \\ = &-\underline{v} - \int_{\underline{v}}^{\overline{v}} \Lambda^{n-1}(v \mid p) dv + \frac{(1-p)^{n-1}}{p}(\overline{v}-\underline{v}) - \frac{(1-p)^{n-1}}{p} \int_{\underline{v}}^{\overline{v}} \Lambda(v \mid p) dv \\ &-(n-1)\int_{\underline{v}}^{\overline{v}} \Lambda^{n-1}(v \mid p) v\Lambda'(v \mid p) dv - \frac{(1-p)^{n-1}}{p} \int_{\underline{v}}^{\overline{v}} \Lambda(v \mid p) v\Lambda'(v \mid p) dv. \end{aligned}$$

By this result and (A.5),

$$\begin{aligned}
& n \int_{\underline{v}}^{\overline{v}} \beta (v \mid p, n) \, H (v \mid p, n) \, d\Lambda (v \mid p) \\
= \quad & (1-p)^{n-1} \, \overline{v} \left(1 + \left(\frac{n}{2} - 1\right) p\right) - \underline{v} - n \int_{\underline{v}}^{\overline{v}} \Lambda^{n-1} (v \mid p) \left(1 - \frac{n-1}{n} \Lambda (v \mid p)\right) dv.
\end{aligned}$$

Therefore, the expected winning bid conditional on the event defining $P(p, n)$ is

$$\frac{1}{P(n,p)} \left\{ \underline{v} - (1-p)^{n-1} \, \overline{v} \left(1 + \left(\frac{n}{2} - 1\right) p\right) + n \int_{\underline{v}}^{\overline{v}} \Lambda^{n-1} (v \mid p) \left(1 - \frac{n-1}{n} \Lambda (v \mid p)\right) dv \right\}.$$

S2 Additional empirical results

S2.1 Determinants of bid variation

This section examines the role of observed heterogeneity in explaining bid variation. The main conclusion is that the engineer's estimate explains the overwhelming portion of bid variation. Moreover, once bids are standardized by the estimate, very little of the remaining variation is explained by the estimate or the observed project characteristics.

We regress the logarithm of the bid, log(Bid), and the logarithm of the bid standardized by the engineer's estimate, log(Bid/Estimate), on the logarithm of the engineer's estimate and a set of observed project characteristics. The observed characteristics include the logarithm of the number of work days, the logarithm of the acreage of full width mowing, the logarithm of the acreage of other mowing, indicators for the number of items (2, 3, 4, 5, or 7 items; 1 item is the omitted category), and indicators for whether the project is a state highway or an interstate highway (local roads are the omitted category). All specifications also control for the number of potential bidders, since the bidding function depends on n.

Table S1 reports the results. Columns (1) and (2) use log(Bid) as the dependent variable. Column (2) excludes the engineer's estimate but includes the observed project covariates; it already attains a high in-sample fit ($R^2 = 0.675$). This indicates that the observed project characteristics account for a substantial fraction of the cross-contract variation in bids. Adding the engineer's estimate in column (1) increases the fit to $R^2 = 0.935$. The implied partial R^2 for the engineer's estimate, computed relative to column (2), is 0.800. Thus, the engineer's estimate explains the bulk of the variation in log(Bid) left unexplained by the other covariates. The coefficient on log(Estimate) is close to one (1.032), consistent with bids scaling nearly proportionally with the estimate.

The list of significant covariates in column (1) includes the number of work days, the acreage of other mowing, and several item count indicators. However, in our selected sample of homogeneous projects with just one item, these factors become irrelevant.

The inclusion of the engineer's estimate in column (1) renders the acreage of full width mowing

Table S1: Determinants of bid variation

	log(Bid)		log(Bid/Estimate)	
	(1)	(2)	(3)	(4)
Engineer's Estimate (log)	1.032***	–	0.032**	–
	(0.013)		(0.013)	
Number of work days (log)	-0.025***	0.310***	-0.025***	–
	(0.008)	(0.030)	(0.008)	
Acreage of full width mowing (log)	-0.007	0.162***	-0.007	-0.004
	(0.006)	(0.020)	(0.006)	(0.005)
Acreage of other mowing (log)	-0.012***	0.044***	-0.012***	-0.011***
	(0.004)	(0.014)	(0.004)	(0.004)
2 items	0.036	-0.072	0.036	0.035
	(0.030)	(0.094)	(0.030)	(0.030)
3 items	0.070**	0.127	0.070**	0.072**
	(0.034)	(0.114)	(0.034)	(0.034)
4 items	0.025	0.385**	0.025	0.033
	(0.060)	(0.171)	(0.060)	(0.060)
5 items	0.148**	0.310*	0.148**	0.156**
	(0.068)	(0.188)	(0.068)	(0.069)
7 items	0.231**	0.638***	0.231**	0.245***
	(0.091)	(0.107)	(0.091)	(0.091)
State	0.000	0.687***	0.000	0.017
	(0.018)	(0.035)	(0.018)	(0.015)
Interstate	0.016	0.112***	0.016	0.017
	(0.012)	(0.025)	(0.012)	(0.011)
R^2	0.935	0.675	0.062	0.056
Partial R^2 (added variables)	0.800	–	0.006	–

Notes:

1. The dependent variable is the logarithm of the bid in columns (1) and (2), and the logarithm of the bid standardized by the engineer's estimate in columns (3) and (4).
2. Controls include dummies for the number of potential bidders (suppressed).
3. HC3 robust standard errors are in parentheses.
4. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
5. Observations: 1,593 in all specifications.
6. Partial R^2 is computed as $(R^2_{\text{full}} - R^2_{\text{controls}})/(1 - R^2_{\text{controls}})$, using (1) vs (2) and (3) vs (4), where R^2_{controls} is from the restricted model excluding the added regressor(s) and R^2_{full} is from the unrestricted model including them.

insignificant, while the number of work days remains significant. After controlling for the engineer's estimate, the number of work days has a negative coefficient, suggesting that projects with longer durations tend to have lower bids, all else equal, possibly reflecting efficiency gains from spreading work over a longer time horizon. As we discuss below, however, the engineer's estimate and the number of work days have little explanatory power for standardized bids.

Columns (3) and (4) report the analogous regressions using standardized bids, log(Bid/Estimate). The standardized specifications have very low explanatory power: $R^2 = 0.062$ in column (3) and $R^2 = 0.056$ in column (4). The partial R^2 associated with adding the engineer's estimate and the number of work days to the control set is only 0.006, implying that once bids are normalized by the estimate, the estimate and work days contribute negligibly to explaining the remaining variation. Because the standardized outcome removes scale variation by construction, the residual dispersion in standardized bids remains largely unexplained by the included covariates.

Overall, the results suggest three complementary patterns. First, observed contract characteristics explain a large share of the level variation in bids. Second, the engineer's estimate is an especially powerful summary statistic for bid levels, explaining most of the remaining variation in log(Bid) beyond the other covariates. Third, very little of the variation in standardized bids is explained by the engineer's estimate or the number of work days, as reflected in the small partial R^2 values.

S2.2 Counterfactual results under an alternative copula specification

In this section, we repeat the counterfactual analysis of Section 7 using the Joe copula assumption instead of the Frank copula assumption. Under the Joe copula assumption, the estimated Spearman's rank correlation is 0.60 (versus 0.68 under the Frank copula), and the estimated weighted average entry cost is 0.056, or 5.6% of the engineer's estimate (versus 0.046, or 4.6%). As in Section 7.1, we set the reserve price to $r = 1$ and initially keep the entry costs at their estimated levels for each n. The results are reported in Table S2.

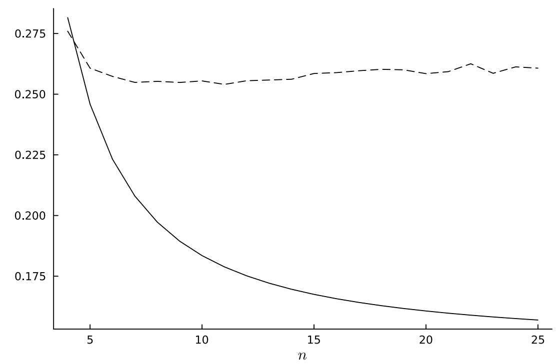
Table S2: Counterfactual entry and bidding probabilities, auction failure probabilities, and expected winning bids under the two formats for different numbers of potential bidders n at the estimated entry costs κ_n (Joe copula)

n	**Bid requirement**			**Reserve price**			
	prob. entry	prob. failure	expect. win. bid	prob. entry	prob. bidding	prob. failure	expect. win. bid
9	0.208	0.412	0.960	0.219	0.149	0.234	0.916
10	0.196	0.388	0.953	0.204	0.139	0.223	0.912
12	0.154	0.428	0.952	0.170	0.118	0.223	0.910
13	0.265	0.104	0.884	0.242	0.164	0.098	0.867
14	0.270	0.075	0.868	0.248	0.168	0.077	0.854

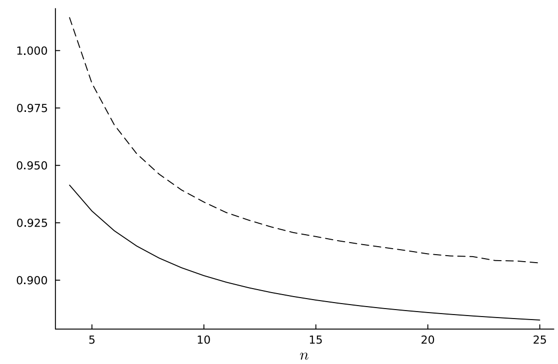
The qualitative patterns are the same as under the Frank copula assumption. The auction failure probability is substantially higher under the bid requirement format for n between 9 and 12, with the gap equal to 17.8%, 16.5%, and 20.5% for $n = 9, 10, 12$, respectively. For $n = 13, 14$, the failure probabilities under the two formats are nearly identical, with gaps of only 0.6% and -0.2%. Compared to the Frank copula results (22.1%, 20.3%, 28.1%, 0.5%, and -0.4%), the gaps are somewhat narrower under the Joe copula, consistent with the lower estimated signal informativeness. Entry probabilities are uniformly higher under the Joe specification, reflecting its lower estimated dependence between signals and private costs. The expected winning bid is lower under the reserve price format by 1.4%–4.4% of the engineer's estimate, similar to the 2.0%–3.7% range under the Frank copula.

Next, as in Section 7.1, we set the entry cost to its estimated weighted average $\kappa = 0.056$ and extend the range of potential bidders to 4–25. The results are shown in Figure S1. Under the Joe copula assumption, the auction failure probability in the bid requirement format is remarkably flat, staying in a narrow range of 0.255–0.275 for $n \geq 7$. This contrasts with the results under the Frank copula assumption, where the failure probability under the bid requirement format declines from approximately 0.55 ($n = 4$) to 0.23 ($n = 25$). Under the reserve price format, the failure probability declines from approximately 0.28 ($n = 4$) to 0.15 ($n = 25$). The gap between the two formats is close to zero at $n = 4$ and widens to approximately 11% at $n = 25$. The reserve price format also yields a lower expected winning bid throughout the range of n, with the gap approximately 4% of the engineer's estimate. Under the Frank copula assumption, this gap varies more widely, from 1.4% to 9.2%.

We next examine the role of signal informativeness. We focus on auctions with $n = 10$ potential



(a) Auction failure probability



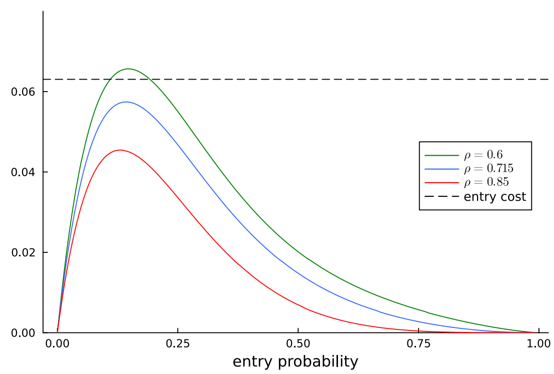
(b) Expected winning bid

Figure S1: Counterfactual auction failure probabilities and expected winning bids for different numbers of potential bidders n and estimated weighted average entry cost $\kappa = 0.056$ under the bid requirement (dashed line) and reserve price (solid line) formats; the reserve price $r = 1$ (Joe copula)

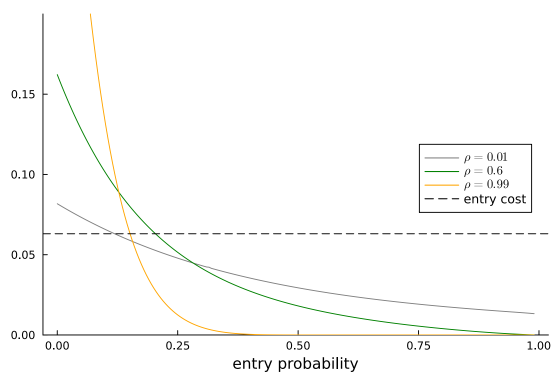
bidders and set the entry cost to 0.063. Figure S2 plots the marginal entrant's expected revenue from entry for different entry probabilities and levels of signal informativeness. Under the bid requirement format, at $\rho = 0.60$, the revenue curve is just above the entry cost 0.063. At $\rho = 0.715$, the revenue curve lies entirely below the entry cost, implying zero entry. The threshold at which entry stops under the bid requirement format is therefore between approximately 0.63 and 0.65, lower than the threshold under the Frank copula assumption. Under the reserve price format, the revenue curves cross the entry cost line for all levels of ρ, so entry remains positive regardless of signal informativeness.

Figure S3 reports the auction failure probabilities and expected winning bids under the two formats as functions of signal informativeness ρ, for $n = 10$. Under the bid requirement format, there is no entry for ρ above approximately 0.65, and the auction fails with probability one, compared to a threshold of 0.715 under the Frank copula assumption. Under the reserve price format, the auction failure probability starts at 0.62 for $\rho = 0$ and reaches its minimum of approximately 0.13 at $\rho \approx 0.85$, remaining strictly below one for all ρ. The two formats yield equal failure probabilities at $\rho \approx 0.50$; for higher signal informativeness, the reserve price format is superior. The expected winning bid is also lower under the reserve price format at all levels of ρ.

In summary, we get the same qualitative conclusions as those from the analysis under the Frank copula assumption. The reserve price format dominates the bid requirement format in terms of both auction failure probability and expected winning bid at the estimated level of signal informativeness.

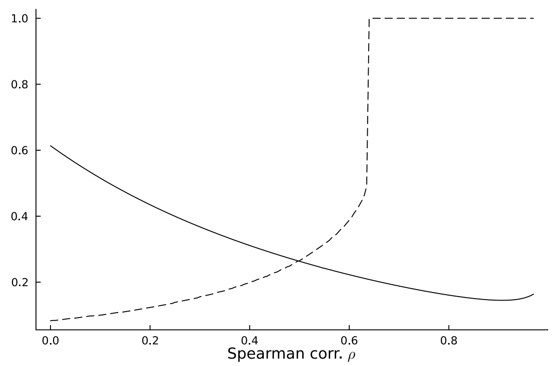


(a) Bid requirement

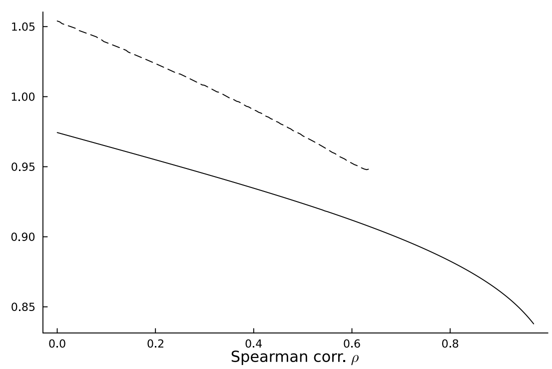


(b) Reserve price

Figure S2: Entry cost and marginal entrant's expected revenue from entry under the bid requirement and reserve price formats in auctions with $n = 10$ potential bidders for different entry probabilities and levels of signal informativeness as measured by Spearman's rank correlation ρ (Joe copula)



(a) Auction failure probability



(b) Expected winning bid

Figure S3: Probabilities of auction failure and expected winning bids under the bid requirement (dashed line) and reserve price (solid line) formats in auctions with $n = 10$ potential bidders for different levels of signal informativeness as measured by Spearman's rank correlation ρ (Joe copula)

S3 Econometric details

S3.1 Introduction and related literature

Li and Zheng (2009) uses variants of the models of Levin and Smith (1994) and Samuelson (1985), with a focus on identifying the entry effect. Marmer, Shneyerov, and Xu (2013) considers the fully nonparametric endogenous entry model, for which the models of Levin and Smith (1994) and Samuelson (1985) are viewed as limit cases, and identifies the selection effect. This paper studies the semiparametric model discussed in the conclusion of Gentry and Li (2014) and develops a complete theory for identification and estimation that, to the best of our knowledge, has not been studied in the literature.[1] We derive testable sufficient conditions that ensure semiparametric local and global identification of the copula parameter, in the sense of Lewbel (2019), and present them in Proposition S2.

Chen, Gentry, Li, and Lu (2026) consider the first-price auction model with both endogenous entry and risk aversion under a parametric assumption on the copula function for the signal and the private cost. They show that the utility function and the distribution of private values are nonparametrically identified conditional on the copula parameter, which can be set identified. In contrast, we study both global and local point identification of the copula parameter under risk neutrality and a parametric copula assumption. Chen, Gentry, Li, and Lu (2026) also show that the utility function is nonparametrically identified when there is sufficient variation in the observed instruments and the number of potential bidders, and they establish parametric identification of the utility function. In either case, the distribution of private costs conditional on entry is nonparametrically identified. An adaptation of our identification results provides testable sufficient conditions that ensure (local or global) identification of the copula parameter in the semiparametric model with endogenous entry, risk aversion, and a parametric copula assumption.

We propose a practical generalized method of moments (GMM) estimator and develop its first-order asymptotic theory; several intermediate results are of independent interest. Our semiparametric GMM estimation uses the empirical CDF of pseudo values constructed from a nonparametrically estimated inverse bidding strategy. Guerre, Perrigne, and Vuong (2000) proposes kernel density estimation using these pseudo values, and Ma, Marmer, and Shneyerov (2019) derives the first-order asymptotic theory of the pseudo-value-based density estimator.[2] The asymptotic properties of the pseudo-value-based cumulative distribution function (CDF) estimator do not appear to have been studied in the literature. The econometric theory derived in this paper complements Ma, Marmer, and Shneyerov (2019) by providing the first-order asymptotic properties of the pseudo-value-based CDF estimator; our proof technique also differs from that of Ma, Marmer, and Shneyerov (2019).

[1] Gentry and Li (2014) considers high-bid auctions with selective entry. The econometric part of our paper considers procurement (low-bid) auctions under the 2-bid requirement. However, our results extend straightforwardly to other scenarios, including high-bid auctions (with or without binding reserve prices) and procurement auctions with binding reserve prices.

[2] See Marmer and Shneyerov (2012) for an alternative estimation method that does not use the estimated inverse bidding strategy or the pseudo values.

Recently, Zincenko (2024) study estimation and inference of the seller's expected revenue in high-bid first-price auctions and derive the asymptotic linearization of a pseudo-value-based CDF estimator. Their approach, however, uses a kernel-smoothed bid CDF estimator in the inverse bidding strategy for constructing the pseudo values, whereas Guerre, Perrigne, and Vuong (2000) and Ma, Marmer, and Shneyerov (2019) use the empirical CDF of bids without kernel smoothing. We follow the latter approach, since the empirical CDF incurs no smoothing bias, and show that our estimator admits the desired asymptotic linearization via a proof entirely different from that of Zincenko (2024).

Estimating the inverse bidding strategy requires plugging in a nonparametric estimator of the compactly supported bid density. The standard kernel density estimator suffers from boundary bias, which can be addressed by either trimming (Guerre, Perrigne, and Vuong, 2000) or boundary bias correction (see, e.g., Hickman and Hubbard, 2015 or Ma, Marmer, Shneyerov, and Xu, 2021). Following Ma, Marmer, Shneyerov, and Xu (2021), we use the boundary-adaptive local linear density estimator (Lejeune and Sarda, 1992; Jones, 1993), which has the advantage of not requiring additional tuning parameters; see also Cheng, Fan, and Marron (1997) and Chen and Huang (2007). The uniform convergence of this estimator over the entire support does not appear to have been established in the literature.

We derive concentration bounds for the more general local polynomial (LP) density estimators (Bickel and Doksum, 2015, Chapter 11.3), as well as their first and second derivatives, extending the classical concentration results for kernel density estimators (Giné and Guillou, 2002). These results yield the uniform rate of convergence of the LP density estimator over the entire support and are useful in other semiparametric estimation problems that involve compactly supported density functions as nuisance parameters.

The optimal weight matrix of the semiparametric GMM estimator depends on the asymptotic variance of the pseudo-value-based CDF estimator, which takes a complicated form. To avoid direct estimation of this variance, we use a bootstrap approach. We show that our bootstrap variance estimator is consistent for the asymptotic variance of the pseudo-value-based CDF estimator.[3] The proof hinges on the concentration bounds for the local linear density estimator derived in Appendix A.

S3.2 Semiparametric identification results

This section discusses the identification of the model primitives on which our estimation approach is based. The following proposition shows that the equilibrium entry probability $p_n = p(n, \kappa; \theta_0)$ is identified from the observed entry frequencies via equation (5.1).

[3] The bootstrap variance estimator may overestimate the asymptotic variance asymptotically, even when the consistency of the bootstrap distribution can be established. See Hahn and Liao (2021) for examples and general theory. Our consistency result rules out this possibility in our specific context. Ma, Marmer, and Shneyerov (2019) shows consistency of the bootstrap distribution for the pseudo-value-based density estimator, but their result does not directly imply the consistency of our bootstrap variance estimator.

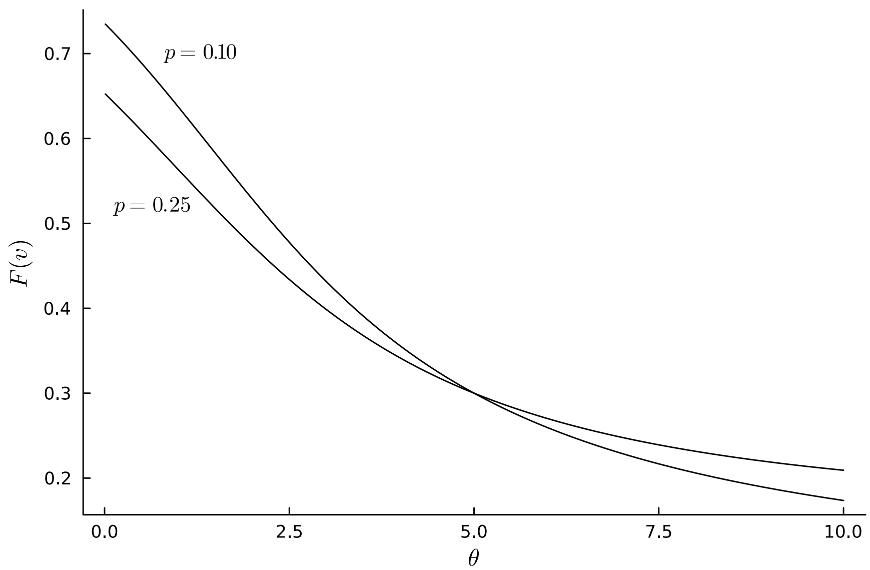


Figure S4: Pairs of candidate values $(\theta, F(v))$ that satisfy the restriction $F^*(v \mid p) = C(F(v), p; \theta)/p$ for entry probabilities $p = 0.10$ and $p = 0.25$. The true values are $\theta = 5.0$ and $F(v) = 0.3$. Each line plots $\theta \mapsto Q(F^*(v \mid p), p; \theta)$, where $Q(\cdot, p; \theta)$ is the inverse function of $C(\cdot, p; \theta)/p$.

Proposition S1. *Suppose that Assumption 5.1 holds. Under the bid requirement format, the equilibrium entry probability p_n is nonparametrically identified for all numbers of potential bidders $n \geq 3$.*

Proof of Proposition S1. Define

$$\varphi_n(p) := \frac{p\left(1 - (1-p)^{n-1}\right)}{1 - (1-p)^n - np\,(1-p)^{n-1}}.$$

We have $\lim_{p \downarrow 0} \varphi_n(p) = 2/n$ and $\varphi_n(1) = 1$. For $n \geq 3$, the function $\varphi_n(p)$ is continuously differentiable on $(0, 1]$ with a derivative $\varphi_n'(p) > 0$ and $\lim_{p \downarrow 0} \varphi_n'(p) = (n-2)/3n$. Hence, the inverse function $\varphi_n^{-1}(\cdot)$ is well-defined, and the equilibrium entry probability p_n is identified for $n \geq 3$. ■

The identification of the distribution of private costs conditional on entry, $F^*(\cdot \mid p_n)$, follows from Proposition 5.1. We can therefore treat it as known for the purpose of identifying the signal informativeness (i.e., copula) parameter θ and the marginal CDF of private costs $F(\cdot)$.

The copula parameter θ is globally identified if for any $\tilde{\theta} \in \Theta$ and a CDF $\tilde{F}$ supported on $[\underline{v}, \bar{v}]$, the condition $F^*(v \mid p_n) = C(\tilde{F}(v), p_n; \tilde{\theta})/p_n$ for all $v \in [\underline{v}, \bar{v}]$ and $n \in \mathcal{N}$ implies $\tilde{\theta} = \theta_0$. Note that by the monotonicity of the copula function, it also follows that $\tilde{F}(v) = F(v)$. We say that θ is locally identified if the property is valid for all $\tilde{\theta}$ in an open neighborhood around θ_0. These definitions follow those in Lewbel (2019).

The restrictions in (5.4) imply that for all $v \in [\underline{v}, \bar{v}]$, the collection of curves

$$\{\theta \mapsto Q(F^*(v \mid p_n), p_n; \theta) : n \in \mathcal{N}\}$$

must intersect at θ_0 (the true copula parameter). Global identification requires that no other point in Θ satisfies this condition. This is guaranteed if, for some v and some $n_1 \neq n_2$, the curves $\theta \mapsto Q(F^*(v \mid p_{n_1}), p_{n_1}; \theta)$ and $\theta \mapsto Q(F^*(v \mid p_{n_2}), p_{n_2}; \theta)$ have a single crossing; the unique intersection point then constructively identifies θ_0. Figure S4 illustrates: the two plotted curves correspond to $\theta \mapsto Q(F^*(v \mid p), p; \theta)$ for two different values of p and a particular v, and they cross uniquely at $\theta = 5.0$.

Denote $C_\theta(u, v; \theta) \coloneqq \partial C(u, v; \theta)/\partial\theta$ and $C_1(u, v; \theta) \coloneqq \partial C(u, v; \theta)/\partial u$. The following proposition provides sufficient conditions for the global and local identification of the copula parameter. Since all expressions involved can be estimated from the data, the conditions are testable.

Proposition S2. *Suppose that Assumptions 4.1 and 5.1 hold, and the support $\mathcal{N}$ of the distribution of the number of potential bidders contains two or more elements, with the smallest element greater than two. (a) The copula parameter θ_0 is globally identified if for some $n_1, n_2 \in \mathcal{N}$ and $v \in [\underline{v}, \bar{v}]$,*

$$\min_{\theta \in \Theta} \frac{\partial}{\partial \theta}\Big(Q\left(F^*\left(v \mid p_{n_1}\right), p_{n_1}; \theta\right) - Q\left(F^*\left(v \mid p_{n_2}\right), p_{n_2}; \theta\right)\Big) > 0. \tag{S11}$$

(b) The copula parameter θ_0 is locally identified if for some $n_1, n_2 \in \mathcal{N}$ and $u \in [0, 1]$,

$$\frac{C_\theta(u, p_{n_2}; \theta_0)}{C_1(u, p_{n_2}; \theta_0)} \neq \frac{C_\theta(u, p_{n_1}; \theta_0)}{C_1(u, p_{n_1}; \theta_0)}. \tag{S12}$$

(c) A sufficient condition for (S12) can be stated in terms of the cross-derivative of the copula function: for some distinct $n_1, n_2 \in \mathcal{N}$ and $u \in [0, 1]$, $p_{n_1} \neq p_{n_2}$ and

$$\max_{p \in [p_{n_1} \wedge p_{n_2}, p_{n_1} \vee p_{n_2}]} \partial^2 C(u, p; \theta_0)/\partial p \partial\theta < 0. \tag{S13}$$

Proof of Proposition S2. By Propositions S1 and 5.1, p_n and $F^*(\cdot \mid p_n)$ are identified for all $n \in \mathcal{N}$. After partialling out $F(\cdot)$, equation (5.4) imposes the following restriction on the copula parameter: $\Delta_{n_1,n_2}(v; \theta_0) = 0$, for all $v \in [\underline{v}, \bar{v}]$ and all distinct $n_1, n_2 \in \mathcal{N}$, where

$$\Delta_{n_1,n_2}(v; \theta) \coloneqq Q(F^*(v \mid p_{n_1}), p_{n_1}; \theta) - Q(F^*(v \mid p_{n_2}), p_{n_2}; \theta).$$

A sufficient condition for the uniqueness of the solution for θ is that the function $\theta \mapsto \Delta_{n_1,n_2}(v; \theta)$ is strictly increasing for some $v \in [\underline{v}, \bar{v}]$ and some distinct $n_1, n_2 \in \mathcal{N}$, which is satisfied if (S11) holds.

For Part (b), θ_0 is locally identified if $\partial\Delta_{n_1,n_2}(v; \theta_0)/\partial\theta > 0$ for some $v \in [\underline{v}, \bar{v}]$ and some distinct

$n_1, n_2 \in \mathcal{N}$. By the implicit function theorem,

$$\frac{\partial Q(y,v;\theta)}{\partial\theta} = -\frac{C_\theta(Q(y,v;\theta),v;\theta)}{C_1(Q(y,v;\theta),v;\theta)}.$$

By Assumption 5.1(vi), $Q(F^*(v \mid p_n), p_n; \theta_0) = F(v)$ for all v and $n \in \mathcal{N}$. We have

$$\frac{\partial\Delta_{n_1,n_2}(v;\theta_0)}{\partial\theta} = \frac{C_\theta(F(v),p_{n_2};\theta_0)}{C_1(F(v),p_{n_2};\theta_0)} - \frac{C_\theta(F(v),p_{n_1};\theta_0)}{C_1(F(v),p_{n_1};\theta_0)} > 0,$$

where the inequality holds by (S12).

For Part (c), pick $n_1, n_2 \in \mathcal{N}$ such that $p_{n_1} > p_{n_2}$. Then we have

$$\begin{aligned}
\frac{C_\theta(F(v),p_{n_2};\theta_0)}{C_1(F(v),p_{n_2};\theta_0)} - \frac{C_\theta(F(v),p_{n_1};\theta_0)}{C_1(F(v),p_{n_1};\theta_0)} &\geq \frac{C_\theta(F(v),p_{n_2};\theta_0) - C_\theta(F(v),p_{n_1};\theta_0)}{C_1(F(v),p_{n_1};\theta_0)} \\
&\geq \frac{-\int_{p_{n_2}}^{p_{n_1}} C_{2\theta}(F(v),p;\theta_0)dp}{C_1(F(v),p_{n_1};\theta_0)} \\
&> 0,
\end{aligned}$$

where the first inequality holds by Assumption 4.1(i) and $p_{n_1} > p_{n_2}$, and because $C_1(F(v),p;\theta_0)$ equals $F_{S|V}(p \mid v)$, the conditional CDF of the signal ranks S given the private costs V; the last inequality holds by (S13). ■

Both global and local identification conditions require $p_{n_1} \neq p_{n_2}$ for some pair of numbers of potential bidders $n_1, n_2 \in \mathcal{N}$. The local conditions in equations (S12) and (S13) are easier to verify than the global condition in (S11), since the former only require estimates of the entry probabilities p_n, $n \in \mathcal{N}$. Since the true value θ_0 is unknown in practice, the local conditions must be verified for all values of θ (Lewbel, 2019).

S3.3 Generalized method of moments estimation

Let

$$\begin{aligned}
q_n &\coloneqq \mathrm{E}\left[\frac{N_l^*}{N_l} \mid N_l = n\right] \\
&= \frac{\mathrm{E}\left[\mathbb{1}(N_l = n)\frac{N_l^*}{N_l}\right]}{\mathrm{E}\left[\mathbb{1}(N_l = n)\right]},
\end{aligned}$$

where the second equality follows from the LIE, and then we have $q_n = \varphi_n(p_n)$. The sample analogue of q_n is thus given by

$$\widehat{q}_n \coloneqq \frac{1}{n\,|\{l : N_l = n\}|}\sum_{l:N_l=n} N_l^*. \tag{S14}$$

Let $\widehat{p}_n \coloneqq \varphi_n^{-1}(\widehat{q}_n)$. A consistent estimator of $G(b \mid n)$ is

$$\widehat{G}(b \mid n) \coloneqq \frac{\sum_{l:N_l=n} \sum_{i=1}^{N_l^*} \mathbb{1}(B_{il} \leq b)}{\sum_{l:N_l=n} N_l^*}. \tag{S15}$$

Let $\overline{b}_n \coloneqq \beta(\overline{v} \mid p_n, n)$ and $\underline{b}_n \coloneqq \beta(\underline{v} \mid p_n, n)$ denote the boundary points. Let

$$\begin{aligned} \widehat{\overline{b}}_n &\coloneqq \max\{B_{il} : i = 1, ..., N_l^*, N_l = n\}, \\ \widehat{\underline{b}}_n &\coloneqq \min\{B_{il} : i = 1, ..., N_l^*, N_l = n\} \end{aligned}$$

be the estimated boundary points. The local-linear-type boundary adaptive kernel density estimator of $g(b \mid n)$ is

$$\widehat{g}(b \mid n) \coloneqq \frac{\sum_{l:N_l=n} \sum_{i=1}^{N_l^*} \frac{1}{h}\mathcal{K}_1\left(B_{il}, b \mid h, \widehat{\underline{b}}_n, \widehat{\overline{b}}_n\right)}{\sum_{l:N_l=n} N_l^*}, \tag{S16}$$

where the function $\mathcal{K}_1$ is defined in Appendix A. Then we construct the nonparametric estimator $\widehat{\xi}(\cdot \mid n)$ of the inverse bidding function $\xi(\cdot \mid p_n, n)$.

In practical estimation of the copula parameter, we consider finitely many grid points $v_1 < \cdots < v_J$ in $(\underline{v}, \overline{v})$. Then we have a finite set of restrictions

$$F^*(v_j \mid p_n) = C(F(v_j), p_n; \theta_0)/p_n, \text{ for } (j, n) \in \{1, ..., J\} \times \mathcal{N}, \tag{S17}$$

for the $J+1$ parameters $(\theta_0, F(v_1), ..., F(v_J))$. Let $M \coloneqq |\mathcal{N}|$ and let $n_1 < n_2 < \cdots < n_M$ be the elements of $\mathcal{N}$. Write the equations as

$$Q(F^*(v_j \mid p_{n_k}), p_{n_k}; \theta_0) = F(v_j), \text{ for } (j, k) \in [J] \times [M]. \tag{S18}$$

For $\boldsymbol{x} = (x_1, ..., x_M)^\top$ and $\boldsymbol{y} = (y_1, ..., y_M)^\top$, denote

$$\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{y}; \theta) \coloneqq \begin{pmatrix} Q(x_1, y_1; \theta) \\ \vdots \\ Q(x_M, y_M; \theta) \end{pmatrix}.$$

Denote $\boldsymbol{F}_0 \coloneqq (F(v_1), ..., F(v_J))^\top$, $\boldsymbol{p}_0 \coloneqq (p_{n_1}, ..., p_{n_M})^\top$, $\boldsymbol{F}_j^* \coloneqq (F^*(v_j \mid p_{n_1}), ..., F^*(v_j \mid p_{n_M}))^\top$ for $j \in [J]$, and $\boldsymbol{F}^* \coloneqq \left(\boldsymbol{F}_1^{*\top}, ..., \boldsymbol{F}_J^{*\top}\right)^\top$. For $\boldsymbol{z} \coloneqq \left(\boldsymbol{x}_1^\top, ..., \boldsymbol{x}_J^\top\right)^\top \in \mathbb{R}^{JM}$, denote

$$\boldsymbol{\Psi}(\boldsymbol{z}, \boldsymbol{y}, \theta) \coloneqq \begin{pmatrix} \boldsymbol{Q}(\boldsymbol{x}_1, \boldsymbol{y}; \theta) \\ \vdots \\ \boldsymbol{Q}(\boldsymbol{x}_J, \boldsymbol{y}; \theta) \end{pmatrix}. \tag{S19}$$

Equation (S18) can be written in vector form as

$$\boldsymbol{\Psi}(\boldsymbol{F}^*, \boldsymbol{p}_0, \theta_0) = (\mathbf{I}_J \otimes \mathbf{1}_M)\boldsymbol{F}_0. \tag{S20}$$

Let $\widehat{\boldsymbol{p}} := (\widehat{p}_{n_1}, ..., \widehat{p}_{n_M})^\top$, $\widehat{\boldsymbol{F}}_j^* := \left(\widehat{F}^*(v_j \mid p_{n_1}), ..., \widehat{F}^*(v_j \mid p_{n_M})\right)^\top$ for $j \in [J]$, and $\widehat{\boldsymbol{F}}^* := \left(\widehat{\boldsymbol{F}}_1^{*\top}, ..., \widehat{\boldsymbol{F}}_J^{*\top}\right)^\top$. Let $\mathbf{W} \in \mathbb{R}^{MJ\times MJ}$ be a weight matrix and $\boldsymbol{F} := (F_1, ..., F_J)^\top$. Let

$$\mathbb{F} := \left\{(F_1, ..., F_J)^\top \in [0,1]^J : F_1 \leq F_2 \leq \cdots \leq F_J\right\}.$$

The GMM criterion function is

$$\widehat{D}(\theta, \boldsymbol{F}; \mathbf{W}) := \left(\boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) - (\mathbf{I}_J \otimes \mathbf{1}_M)\boldsymbol{F}\right)^\top \mathbf{W}\left(\boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) - (\mathbf{I}_J \otimes \mathbf{1}_M)\boldsymbol{F}\right).$$

The GMM estimator is defined by

$$\left(\widehat{\theta}(\mathbf{W}), \widehat{\boldsymbol{F}}(\mathbf{W})\right) := \underset{(\theta,\boldsymbol{F})\in\Theta\times\mathbb{F}}{\arg\min}\ \widehat{D}(\theta, \boldsymbol{F}; \mathbf{W}). \tag{S21}$$

To solve the minimization problem in (S21), we can easily partial out $\boldsymbol{F}$ given fixed θ. This requires solving a quadratic program under linear inequality constraints. Let

$$\widehat{\boldsymbol{F}}(\theta; \mathbf{W}) := \underset{\boldsymbol{F}\in\mathbb{F}}{\arg\min}\ \widehat{D}(\theta, \boldsymbol{F}; \mathbf{W}). \tag{S22}$$

Then, after partialling out $\boldsymbol{F}$, we can calculate the GMM estimator

$$\widehat{\theta}(\mathbf{W}) = \underset{\theta\in\Theta}{\arg\min}\ \left(\boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) - (\mathbf{I}_J \otimes \mathbf{1}_M)\widehat{\boldsymbol{F}}(\theta; \mathbf{W})\right)^\top \mathbf{W}\left(\boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) - (\mathbf{I}_J \otimes \mathbf{1}_M)\widehat{\boldsymbol{F}}(\theta; \mathbf{W})\right)$$

by one-dimensional grid search.

We now consider a useful special case. The first-order conditions corresponding to the unconstrained quadratic programming problem $\min_{\boldsymbol{F}\in\mathbb{R}^J} \widehat{D}(\theta, \boldsymbol{F}; \mathbf{W})$ are

$$-2(\mathbf{I}_J \otimes \mathbf{1}_M)^\top \mathbf{W}\boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) + 2(\mathbf{I}_J \otimes \mathbf{1}_M)^\top \mathbf{W}(\mathbf{I}_J \otimes \mathbf{1}_M)\boldsymbol{F} = 0.$$

In the case of $\mathbf{W} = \mathbf{I}_{MJ}$, since $(\mathbf{I}_J \otimes \mathbf{1}_M)^\top (\mathbf{I}_J \otimes \mathbf{1}_M) = M \cdot \mathbf{I}_J$ and

$$(\mathbf{I}_J \otimes \mathbf{1}_M)^\top \boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) = \begin{bmatrix} \mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_1^*, \widehat{\boldsymbol{p}}; \theta\right) \\ \vdots \\ \mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_J^*, \widehat{\boldsymbol{p}}; \theta\right) \end{bmatrix},$$

the minimizer corresponding to the unconstrained problem $\min_{\boldsymbol{F}\in\mathbb{R}^J} \widehat{D}(\theta, \boldsymbol{F}; \mathbf{I}_{MJ})$ is given by

$$\left((\mathbf{I}_J \otimes \mathbf{1}_M)^\top (\mathbf{I}_J \otimes \mathbf{1}_M)\right)^{-1} (\mathbf{I}_J \otimes \mathbf{1}_M)^\top \boldsymbol{\Psi}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \theta\right) = \begin{pmatrix} \frac{\mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_1^*, \widehat{\boldsymbol{p}}; \theta\right)}{M} \\ \vdots \\ \frac{\mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_J^*, \widehat{\boldsymbol{p}}; \theta\right)}{M} \end{pmatrix}. \tag{S23}$$

Since $\mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_1^*, \widehat{\boldsymbol{p}}; \theta\right) \leq \cdots \leq \mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_J^*, \widehat{\boldsymbol{p}}; \theta\right)$, the unconstrained minimizer satisfies the linear inequality constraints in (S22). Therefore, $\widehat{\boldsymbol{F}}(\theta; \mathbf{I}_{MJ})$ has a closed-form solution

$$\widehat{\boldsymbol{F}}(\theta; \mathbf{I}_{MJ}) = \begin{pmatrix} \frac{\mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_1^*, \widehat{\boldsymbol{p}}; \theta\right)}{M} \\ \vdots \\ \frac{\mathbf{1}_M^\top \boldsymbol{Q}\left(\widehat{\boldsymbol{F}}_J^*, \widehat{\boldsymbol{p}}; \theta\right)}{M} \end{pmatrix},$$

since it must coincide with the unconstrained minimizer.

S3.3.1 Asymptotic normality

We first establish the asymptotic normality of the pseudo-value-based CDF estimator. In addition to Assumptions 3.1 and 5.1 in the main text, we impose the following assumption on the data-generating process. Recall that $f = F'$ denotes the probability density function (PDF) of the marginal distribution of private costs and $C_1(x, y) \coloneqq \partial C(x, y) / \partial x$.

Assumption S1. *(a) f is twice continuously differentiable and bounded away from zero on $[\underline{v}, \overline{v}]$. (b) $C_1(\cdot, y)$ is bounded away from zero for all $y \in (0, 1)$.*

Since

$$f^*(v \mid p_n) \coloneqq \frac{\partial F^*(v \mid p_n)}{\partial v} = \frac{C_1(F(v), p_n)}{p_n} \cdot f(v),$$

Assumption S1(b) guarantees that $f^*(\cdot \mid p_n)$ also satisfies the conditions in Assumption S1(a). If $C(\cdot, \cdot)$ is an Archimedean copula with a twice differentiable strict generator $\varphi : [0, 1] \to [0, \infty]$ such that $\varphi'(u) < 0$ and $\varphi''(u) \geq 0$ for $u \in (0, 1)$, $\varphi(1) = 0$, and $\varphi(0) = \infty$ (see Nelsen, 2006, Chapter 4.1 for details on Archimedean copulas), then the one-sided derivative $\varphi'(1)$ exists and satisfies $\varphi'(1) \leq 0$.[4] Furthermore, $\partial^2 C(x, y) / \partial x^2 \leq 0$, which implies that $C_1(\cdot, y)$ is non-increasing. Moreover,

$$\lim_{t \uparrow 1} C_1(t, y) = \frac{\varphi'(1)}{\varphi'(y)}.$$

Thus, Assumption S1(b) holds if $\varphi'(1) < 0$.

We also require the following mild condition on the kernel function $K(\cdot)$ used in (S16).

Assumption S2. *The kernel function $K(\cdot)$ is symmetric, compactly supported on $[-1, 1]$, and twice continuously differentiable on $\mathbb{R}$.*

Proposition S3 establishes the asymptotic normality of the estimator. Let $\beta'(\cdot \mid p_n, n)$ denote the derivative of $\beta(\cdot \mid p_n, n)$, and let $g''(\cdot \mid n)$ denote the second derivative of $g(\cdot \mid n)$.

[4] Since $\varphi''(u) \geq 0$ for $u \in (0, 1)$, φ' is non-decreasing on $(0, 1)$. It follows that $\lim_{u \uparrow 1} \varphi'(u) = \sup\{\varphi'(u) : u \in (0, 1)\}$ and $\lim_{u \uparrow 1} \varphi'(u) \leq 0$. Therefore, by the mean value theorem, $\varphi'(1)$ exists and equals $\lim_{u \uparrow 1} \varphi'(u)$.

Proposition S3. *Suppose that Assumptions 3.1 and 5.1 in the main text and Assumptions S1 and S2 hold, and that the bandwidth h is chosen to be proportional to $L^{-\gamma}$ with $1/5 \leq \gamma < 1/3$. Then the following results hold.*

(a) The estimator $\widehat{F}^*(v \mid p_n)$ is asymptotically normal:

$$\sqrt{Lh}\left(\widehat{F}^*(v \mid p_n) - F^*(v \mid p_n) - \Xi(v \mid n)\left(\int K(u)\,u^2 du\right)h^2\right) \to_d \mathrm{N}\left(0, \Sigma(v \mid n)\int K^2(u)\,du\right),$$

where

$$\begin{aligned}
\Xi(v \mid n) &:= -\frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))\,\beta'(v \mid p_n, n)\,g''(\beta(v \mid p_n, n) \mid n)}{2(n-1)\,g(\beta(v \mid p_n, n) \mid n)}, \\
\Sigma(v \mid n) &:= \frac{\eta_n^2(p_n, G(\beta(v \mid p_n, n) \mid n))\,(\beta'(v \mid p_n, n))^2}{(n-1)^2\,g(\beta(v \mid p_n, n) \mid n)}.
\end{aligned}$$

(b) Let $\boldsymbol{\Sigma}_j := (\Sigma(v_j \mid n_1), ..., \Sigma(v_j \mid n_M))^\top$ and $\boldsymbol{\Xi}_j := (\Xi(v_j \mid n_1), ..., \Xi(v_j \mid n_M))^\top$. The joint asymptotic distribution is given by:

$$\sqrt{Lh}\left(\widehat{\boldsymbol{F}}^* - \boldsymbol{F}^* - \boldsymbol{\Xi}\left(\int K(u)\,u^2 du\right)h^2\right) \to_d \mathrm{N}\left(0, \mathbf{\Sigma}\int K^2(u)\,du\right),$$

where $\mathbf{\Sigma}$ is a diagonal matrix with $\left(\boldsymbol{\Sigma}_1^\top, \boldsymbol{\Sigma}_2^\top, ..., \boldsymbol{\Sigma}_J^\top\right)^\top$ being the diagonal elements and $\boldsymbol{\Xi} := \left(\boldsymbol{\Xi}_1^\top, \boldsymbol{\Xi}_2^\top, ..., \boldsymbol{\Xi}_J^\top\right)^\top$.

Let

$$\begin{aligned}
\boldsymbol{\Psi}_0(\boldsymbol{z}, \boldsymbol{y}, \theta) &:= \frac{\partial \boldsymbol{\Psi}(\boldsymbol{z}, \boldsymbol{y}, \theta)}{\partial \theta}, \\
\boldsymbol{\Psi}_1(\boldsymbol{z}, \boldsymbol{y}, \theta) &:= \frac{\partial \boldsymbol{\Psi}(\boldsymbol{z}, \boldsymbol{y}, \theta)}{\partial \boldsymbol{z}^\top}
\end{aligned}$$

denote the partial derivatives. The asymptotic theory of our GMM estimator parallels that of the local GMM estimator in Lewbel (2007). Denote $\boldsymbol{\vartheta} := \left(\theta, \boldsymbol{F}^\top\right)^\top$, $\boldsymbol{\vartheta}_0 := \left(\theta_0, \boldsymbol{F}_0^\top\right)^\top$, $\widehat{\boldsymbol{\vartheta}}(\mathbf{W}) := \left(\widehat{\theta}(\mathbf{W}), \widehat{\boldsymbol{F}}(\mathbf{W})^\top\right)^\top$ and write

$$\boldsymbol{\Upsilon}(\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{\vartheta}) := \boldsymbol{\Psi}(\boldsymbol{z}, \boldsymbol{y}, \theta) - (\mathbf{I}_J \otimes \mathbf{1}_M)\,\boldsymbol{F}.$$

Equation (S20) can then be written compactly as

$$\boldsymbol{\Upsilon}(\boldsymbol{F}^*, \boldsymbol{p}_0, \boldsymbol{\vartheta}_0) = \mathbf{0}_{JM},$$

and the estimator $\widehat{\boldsymbol{\vartheta}}(\mathbf{W})$ can be represented as

$$\widehat{\boldsymbol{\vartheta}}(\mathbf{W}) = \underset{\boldsymbol{\vartheta} \in \Theta \times \mathbb{F}}{\arg\min}\, \widehat{D}(\boldsymbol{\vartheta}; \mathbf{W}),$$

where

$$\widehat{D}(\boldsymbol{\vartheta};\mathbf{W}) \coloneqq \boldsymbol{\Upsilon}^{\top}\left(\widehat{\boldsymbol{F}}^{*},\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}\right)\mathbf{W}\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^{*},\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}\right).$$

We say that θ_0 is globally identified from the finite set of restrictions if the system of equations

$$\boldsymbol{\Upsilon}(\boldsymbol{F}^{*},\boldsymbol{p}_0,\boldsymbol{\vartheta}) = \mathbf{0}_{JM} \tag{S24}$$

has a unique solution at $\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0$. Similarly, θ_0 is globally identified from the finite set of restrictions under the stronger condition that for some $k,l\in[M]$ and $j\in[J]$,

$$\min_{\theta\in\Theta}\frac{\partial}{\partial\theta}\left(Q\left(F^{*}\left(v_j \mid p_{n_k}\right),p_{n_k};\theta\right)-Q\left(F^{*}\left(v_j \mid p_{n_l}\right),p_{n_l};\theta\right)\right)>0. \tag{S25}$$

Likewise, θ_0 is locally identified from the finite set of restrictions if there exists an open neighborhood around θ_0 such that for any $\theta\neq\theta_0$ in the neighborhood and $(F_1,...,F_J)\in[0,1]^J$ satisfying $F_1\leq F_2\leq\cdots\leq F_J$, $(\theta,F_1,...,F_J)$ is not a solution to (S24). The condition in Proposition S2(b) guarantees that there exists some $v\in[\underline{v},\overline{v}]$ such that θ_0 is locally identified from the restrictions

$$F^{*}(v \mid p_n) = C(F(v),p_n;\theta_0)/p_n,\ n\in\mathcal{N}.$$

The following mild assumption on the parametric copula family $\{C(\cdot,\cdot;\theta):\theta\in\Theta\}$ is satisfied by most commonly used parametric copula families (e.g., Gaussian, Ali-Mikhail-Haq, Clayton, Frank, Gumbel, and Joe). Let $C_1(x,y;\theta)\coloneqq\partial C(x,y;\theta)/\partial x$, $C_2(x,y;\theta)\coloneqq\partial C(x,y;\theta)/\partial y$ and $C_\theta(x,y;\theta)\coloneqq\partial C(x,y;\theta)/\partial\theta$.

Assumption S3. *For all $\epsilon\in(0,1/2)$, $C_1(x,y;\theta)>0$ for all $(x,y,\theta)\in(0,1)\times[\epsilon,1-\epsilon]\times\Theta$.*

Under Assumption S3, $C(\cdot,y;\theta)/y$ is strictly increasing on $[0,1]$, and therefore, $Q(\cdot,y;\theta)$ is also strictly increasing on $[0,1]$. Then,

$$0<Q(\epsilon,y;\theta)\leq Q(x,y;\theta)\leq Q(1-\epsilon,y;\theta)<1, \tag{S26}$$

for all $(x,y,\theta)\in[\epsilon,1-\epsilon]^2\times\Theta$. Since we can write

$$\frac{C(Q(x,y;\theta),y;\theta)}{y}-x=0,$$

by (S26) and the implicit function theorem, $Q(\cdot,\cdot;\cdot)$ is continuously differentiable on $[\epsilon,1-\epsilon]^2\times\Theta$.

Let $\boldsymbol{\Psi}_1\coloneqq\boldsymbol{\Psi}_1(\boldsymbol{F}^{*},\boldsymbol{p}_0,\theta_0)$, $\boldsymbol{\Psi}_0\coloneqq\boldsymbol{\Psi}_0(\boldsymbol{F}^{*},\boldsymbol{p}_0,\theta_0)$, and $\boldsymbol{\Omega}_0\coloneqq\boldsymbol{\Psi}_1\boldsymbol{\Sigma}\boldsymbol{\Psi}_1$. Clearly, $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Omega}_0$ are both diagonal matrices. Let

$$\boldsymbol{\Pi}_0\coloneqq\frac{\partial\boldsymbol{\Upsilon}(\boldsymbol{F}^{*},\boldsymbol{p}_0,\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}^{\top}}=\left[\begin{array}{cc}\boldsymbol{\Psi}_0 & -(\mathbf{I}_J\otimes\mathbf{1}_M)\end{array}\right].$$

The second equality shows that $\mathbf{\Pi}_0$ has full rank under condition (S25). Denote

$$\begin{aligned}\boldsymbol{\Xi}_{\vartheta}(\mathbf{W}) &:= -\left(\mathbf{\Pi}_0^\top \mathbf{W}\mathbf{\Pi}_0\right)^{-1}\mathbf{\Pi}_0^\top \mathbf{W}\mathbf{\Psi}_1\boldsymbol{\Xi}, \\ \mathbf{\Sigma}_{\vartheta}(\mathbf{W}) &:= \left(\mathbf{\Pi}_0^\top \mathbf{W}\mathbf{\Pi}_0\right)^{-1}\mathbf{\Pi}_0^\top \mathbf{W}\mathbf{\Omega}_0\mathbf{W}\mathbf{\Pi}_0\left(\mathbf{\Pi}_0^\top \mathbf{W}\mathbf{\Pi}_0\right)^{-1}.\end{aligned}$$

Using Proposition S3 and standard arguments for M-estimators (e.g., Hansen, 2022, Chapter 22), we establish the consistency and asymptotic normality of the GMM estimator.

Proposition S4. *Suppose that the conditions in Proposition S3 and Assumption 4.1 in the main text are satisfied. Suppose further that $\widehat{\mathbf{W}} \to_p \mathbf{W}_0$ for some deterministic positive definite matrix $\mathbf{W}_0$, that $\boldsymbol{\vartheta}_0$ is in the interior of $\Theta \times \mathbb{F}$, that $\mathbf{\Pi}_0$ has full rank, that (S24) has a unique solution at $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$, and that Assumption S3 holds. Then the following results hold.*

(a) The GMM estimator is consistent: $\hat{\boldsymbol{\vartheta}}\left(\widehat{\mathbf{W}}\right) \to_p \boldsymbol{\vartheta}_0$.

(b) The GMM estimator is asymptotically normal:

$$\sqrt{Lh}\left(\hat{\boldsymbol{\vartheta}}\left(\widehat{\mathbf{W}}\right) - \boldsymbol{\vartheta}_0 - \boldsymbol{\Xi}_{\vartheta}(\mathbf{W}_0)\left(\left(\int K(u)\,u^2 du\right)h^2\right)\right) \to_d \mathrm{N}\left(0, \mathbf{\Sigma}_{\vartheta}(\mathbf{W}_0)\int K^2(u)\,du\right).$$

Standard calculations (e.g., Hansen, 2022, Theorem 13.5) imply that $\mathbf{\Sigma}_{\vartheta}(\mathbf{W}_0) - \mathbf{\Sigma}_{\vartheta}\left(\mathbf{\Omega}_0^{-1}\right)$ is positive semidefinite, with the optimal weight matrix given by $\mathbf{\Omega}_0^{-1}$.

S3.3.2 Bootstrap estimation of the optimal weight matrix

Estimating the optimal weight matrix $\mathbf{\Omega}_0^{-1}$ requires estimating $\Sigma(v \mid n)$, which takes a complicated form and depends on the derivative $\beta'(v \mid p_n, n)$. We propose the following nonparametric bootstrap estimator of $\Sigma(v \mid n)$. Let

$$\left\{\left(B_{1l}^{\dagger}, ..., B_{N_l^{*\dagger} l}^{\dagger}, N_l^{*\dagger}, N_l^{\dagger}\right) : l \in [L]\right\}$$

be the nonparametric bootstrap sample drawn with replacement from

$$\left\{\left(B_{1l}, ..., B_{N_l^* l}, N_l^*, N_l\right) : l \in [L]\right\}. \tag{S27}$$

Let $\mathrm{E}_{\dagger}[\cdot]$, $\mathrm{Var}_{\dagger}[\cdot]$, and $\Pr_{\dagger}[\cdot]$ denote the conditional expectation, variance, and probability given (S27).

Let

$$\hat{q}_n^{\dagger} := \frac{1}{n\left|\left\{l : N_l^{\dagger} = n\right\}\right|}\sum_{l : N_l^{\dagger} = n} N_l^{*\dagger}$$

be the bootstrap analogue of $\widehat{q}_n$ and let

$$\begin{aligned}\widehat{G}_{\dagger}(b,n) &:= \frac{1}{L}\sum_{l:N_l^{\dagger}=n}\sum_{i=1}^{N_l^{*\dagger}}\mathbb{1}\left(B_{il}^{\dagger}\le b\right),\\ \widehat{g}_{\dagger}(b,n) &:= \frac{1}{L}\sum_{l:N_l^{\dagger}=n}\sum_{i=1}^{N_l^{*\dagger}}\frac{1}{h}\mathcal{K}_1\left(B_{il}^{\dagger},b\mid h,\widehat{\underline{b}}_n,\widehat{\overline{b}}_n\right).\end{aligned}$$

By construction, $\mathrm{E}_{\dagger}\left[\widehat{G}_{\dagger}(b,n)\right]=\widehat{G}(b,n)$. Let

$$\begin{aligned}\widetilde{r}_n^{\dagger} &:= \frac{1}{L}\sum_{l:N_l^{\dagger}=n}N_l^{*\dagger},\\ \widehat{r}_n &:= \frac{1}{L}\sum_{l:N_l=n}N_l^{*},\\ \widehat{\sigma}_{r,n}^2 &:= \frac{1}{L}\sum_{l:N_l=n}^{L}\left(N_l^{*}\right)^2-\widehat{r}_n^2,\\ \widehat{r}_n^{\dagger} &:= \widetilde{r}_n^{\dagger}\vee\left(\widehat{r}_n-\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log(L)}{L}}\right),\end{aligned}$$

$\widehat{G}_{\dagger}(b\mid n):=\widehat{G}_{\dagger}(b,n)/\widehat{r}_n^{\dagger}$ and $\widehat{g}_{\dagger}(b\mid n):=\widehat{g}_{\dagger}(b,n)/\widehat{r}_n^{\dagger}$. Then let $\widehat{p}_n^{\dagger}:=\varphi_n^{-1}\left(\widehat{q}_n^{\dagger}\right)$ and let

$$\widehat{\xi}_{\dagger}(b\mid n)=b-\frac{\eta_n\left(\widehat{p}_n^{\dagger},\widehat{G}_{\dagger}(b\mid n)\right)}{(n-1)\,\widehat{g}_{\dagger}(b\mid n)}$$

be the bootstrap analogue of $\widehat{\xi}(b\mid n)$ and let $\widehat{V}_{il}^{\dagger}:=\widehat{\xi}_{\dagger}\left(B_{il}^{\dagger}\mid N_l^{\dagger}\right)$. Then let

$$\widehat{F}_{\dagger}^{*}(v,n):=\frac{1}{L}\sum_{l:N_l^{\dagger}=n}\sum_{i=1}^{N_l^{*\dagger}}\mathbb{1}\left(\widehat{V}_{il}^{\dagger}\le v\right) \tag{S28}$$

be the bootstrap analogue of $\widehat{F}^{*}(v,n)$. Let $\widehat{F}_{\dagger}^{*}(v\mid p_n):=\widehat{F}_{\dagger}^{*}(v,n)/\widehat{r}_n^{\dagger}$.

Proposition S5. *Suppose that the assumptions in the statement of Proposition S4 are satisfied. Then, for all $(j,k)\in[J]\times[M]$,*

$$\frac{\mathrm{Var}_{\dagger}\left[\widehat{F}_{\dagger}^{*}\left(v_j\mid p_{n_k}\right)\right]}{\left(\Sigma\left(v_j\mid n_k\right)\int K^2(u)\,du\right)/(Lh)}\to_p 1.$$

Let $\widetilde{\theta}:=\widehat{\theta}\left(\mathbf{I}_{MJ}\right)$ be the preliminary estimator. Let $\widehat{\boldsymbol{\Psi}}_1:=\boldsymbol{\Psi}_1\left(\widehat{\boldsymbol{F}}^{*},\widehat{\boldsymbol{p}},\widetilde{\theta}\right)$ and let $\widehat{\boldsymbol{\Sigma}}$ be the diagonal matrix with $\left(\widehat{\boldsymbol{\Sigma}}_1^\top, \widehat{\boldsymbol{\Sigma}}_2^\top, ..., \widehat{\boldsymbol{\Sigma}}_J^\top\right)^\top \in \mathbb{R}^{JM}$ being the diagonal elements, where

$$\widehat{\boldsymbol{\Sigma}}_j \coloneqq (Lh)\left(\mathrm{Var}_\dagger\left[\widehat{F}_\dagger^*\left(v_j \mid p_{n_1}\right)\right], ..., \mathrm{Var}_\dagger\left[\widehat{F}_\dagger^*\left(v_j \mid p_{n_M}\right)\right]\right)^\top.$$

The estimated optimal weight matrix is given by $\left(\widehat{\mathbf{\Psi}}_1 \widehat{\mathbf{\Sigma}} \widehat{\mathbf{\Psi}}_1\right)^{-1}$.

Appendix

Notation. For any function $f : A \to \mathbb{R}$, let $\|f\|_A \coloneqq \sup_{x\in A} |f(x)|$. For a univariate function f, denote $f^{(j)}(x) \coloneqq (d/dx)^j f(x)$. For a bivariate function f, denote $D_1 f(x,y) \coloneqq \partial f(x,y)/\partial x$ and $D_2 f(x,y) \coloneqq \partial f(x,y)/\partial y$. For a symmetric matrix $\mathbf{A}$, let $\mathrm{mineig}(\mathbf{A})$ denote its smallest eigenvalue. Let $[a \pm b]$ be shorthand notation for the interval $[a-b, a+b]$. "$a \eqqcolon b$" is understood as "b is defined by a.". We write $a \lesssim b$ if $a \leq C \cdot b$ for some positive constant C that does not depend on the sample size L. Let $\|\boldsymbol{x}\|$ denote the Euclidean norm of a real vector $\boldsymbol{x}$. For a matrix $\mathbf{A}$, $\|\mathbf{A}\|$ is understood as the operator norm of $\mathbf{A}$. "Law of iterated expectations" is abbreviated as "LIE". "First-order conditions" is abbreviated as "FOCs". "With probability approaching one" is abbreviated as "wpa1".

Let $\mathfrak{F}$ denote a class of $\mathbb{R}$-valued functions defined on a compact set S in a finite-dimensional Euclidean space. Let $\mathfrak{F}$ be equipped with a norm $\|\cdot\|$. We say that a finite subset $\mathfrak{F}^\circ$ of $\mathfrak{F}$ is an ε-net if the union of the closed $\|\cdot\|$-balls of radius ε centered at points in $\mathfrak{F}^\circ$ covers $\mathfrak{F}$. $N(\varepsilon, \mathfrak{F}, \|\cdot\|) \coloneqq \inf\{|\mathfrak{F}^\circ| : \mathfrak{F}^\circ \text{ is an } \varepsilon\text{-net of } \mathfrak{F}\}$ is called the ε-covering number. A function $F : S \to \mathbb{R}_+$ is an envelope of $\mathfrak{F}$ if $\sup_{f\in\mathfrak{F}} |f| \leq F$. We say that $\mathfrak{F}$ is a (uniform) Vapnik–Chervonenkis-type (VC-type) class with respect to the envelope F (see, e.g., Giné and Nickl 2015, Definition 3.6.10) if there exist some positive constants (A, V) that are independent of the sample size such that for all $\varepsilon \in (0,1]$,

$$\sup_{Q\in\mathcal{Q}} N\left(\varepsilon \|F\|_{Q,2}, \mathfrak{F}, \|\cdot\|_{Q,2}\right) \leq \left(\frac{A}{\varepsilon}\right)^V, \tag{S29}$$

where $\mathcal{Q}$ denotes the collection of all finitely discrete probability measures on S and $\|f\|_{Q,2} \coloneqq \sqrt{\int f^2 dQ}$.

Appendix A Concentration analysis of the local polynomial density estimator

Let $X_1, ..., X_n$ be an i.i.d. sample where X_i has a bounded PDF f supported on $\mathcal{X} \coloneqq [\underline{x}, \overline{x}]$. For some $p \in \mathbb{N}$, let $\boldsymbol{r}_p(t) \coloneqq (1, t, ..., t^p)^\top$, $\boldsymbol{K}_p(t) \coloneqq \boldsymbol{r}_p(t) K(t)$, $\mathbf{R}_p(t) \coloneqq \boldsymbol{r}_p(t) \boldsymbol{r}_p^\top(t)$ and

$$\begin{aligned}
\mathbf{W}_p(x \mid h, \underline{x}, \overline{x}) &\coloneqq \int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}} \mathbf{R}_p(t) K(t)\, dt, \\
\mathcal{K}_p(X_i, x \mid h, \underline{x}, \overline{x}) &\coloneqq \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p\left(\frac{X_i - x}{h}\right),
\end{aligned}$$

where $\mathbf{e}_1 \coloneqq (1, 0, ..., 0)^\top \in \mathbb{R}^{p+1}$. When h is sufficiently small,

$$\sup_{x \in \mathcal{X}} \left\| \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \right\| \leq \left(\text{mineig}\left(\int_0^1 \mathbf{R}_p(t) K(t)\, dt \right) \right)^{-1} =: \varpi_p. \tag{S30}$$

The local polynomial density estimator of f is

$$\widehat{f}(x) \coloneqq \frac{1}{nh} \sum_{i=1}^n \mathcal{K}_p(X_i, x \mid h, \underline{x}, \overline{x}). \tag{S31}$$

In practice, the boundary points $\underline{x}$ and $\overline{x}$ are unknown. However, validity of the first-order asymptotic results derived in this section is unaffected if we replace the unknown $\underline{x}$ and $\overline{x}$ with their super-consistent estimators $\min\{X_1, X_2, ..., X_n\}$ and $\max\{X_1, X_2, ..., X_n\}$. Let $\widehat{f}'$ and $\widehat{f}''$ be the first and second derivatives of $\widehat{f}$.

Denote

$$\boldsymbol{K}_p'(t) \coloneqq \frac{d\boldsymbol{K}_p(t)}{dt} = \frac{d\boldsymbol{r}_p(t)}{dt} K(t) + \boldsymbol{r}_p(t) K'(t)$$

and $\boldsymbol{K}_p''(t) \coloneqq d^2 \boldsymbol{K}_p(t) / dt^2$. By differentiation,

$$\frac{d}{dx} \mathbf{W}_p(x \mid h, \underline{x}, \overline{x}) = \frac{1}{h} \mathbf{D}_p(x \mid h, \underline{x}, \overline{x}), \tag{S32}$$

where

$$\mathbf{D}_p(x \mid h, \underline{x}, \overline{x}) \coloneqq -\mathbf{R}_p\left(\frac{\overline{x} - x}{h}\right) K\left(\frac{\overline{x} - x}{h}\right) + \mathbf{R}_p\left(\frac{\underline{x} - x}{h}\right) K\left(\frac{\underline{x} - x}{h}\right)$$

and $\sup_{x \in \mathcal{X}} \|\mathbf{D}_p(x \mid h, \underline{x}, \overline{x})\| < \infty$, since $K(\cdot)$ is compactly supported on $[-1, 1]$. By this result, the product rule and

$$\frac{d}{dx} \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) = -\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \frac{d}{dx} \mathbf{W}_p(x \mid h, \underline{x}, \overline{x}) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}), \tag{S33}$$

we have

$$\frac{\partial}{\partial x} \mathcal{K}_p(X_i, x \mid h, \underline{x}, \overline{x}) = \frac{1}{h} \dot{\mathcal{K}}_p(X_i, x \mid h, \underline{x}, \overline{x}),$$

where

$$\begin{aligned}\dot{\mathcal{K}}_p(X_i, x \mid h, \underline{x}, \overline{x}) \quad = \quad & -\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p' \left(\frac{X_i - x}{h} \right) \\ & -\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \mathbf{D}_p(x \mid h, \underline{x}, \overline{x}) \\ & \quad \times \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p \left(\frac{X_i - x}{h} \right). \end{aligned} \tag{S34}$$

Then we have

$$\widehat{f}'(x) = \frac{1}{nh^2} \sum_{i=1}^n \dot{\mathcal{K}}_p(X_i, x \mid h, \underline{x}, \overline{x}). \tag{S35}$$

For some matrix $\mathbf{S}_p(x \mid h, \underline{x}, \overline{x})$,

$$\frac{d}{dx} \mathbf{D}_p(x \mid h, \underline{x}, \overline{x}) = \frac{1}{h} \cdot \mathbf{S}_p(x \mid h, \underline{x}, \overline{x})$$

and $\sup_{x \in \mathcal{X}} \|\mathbf{S}_p(x \mid h, \underline{x}, \overline{x})\| < \infty$. By (S32), (S33) and the product rule,

$$\frac{\partial^2}{\partial x^2} \mathcal{K}_p(X_i, x \mid h, \underline{x}, \overline{x}) = \frac{1}{h^2} \ddot{\mathcal{K}}_p(X_i, x \mid h, \underline{x}, \overline{x}),$$

where

$$\begin{aligned}\ddot{\mathcal{K}}_p(X_i, x \mid h, \underline{x}, \overline{x}) \coloneqq\ & \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p'' \left(\frac{X_i - x}{h} \right) \\ & + 2 \cdot \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \mathbf{D}_p(x \mid h, \underline{x}, \overline{x}) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p' \left(\frac{X_i - x}{h} \right) \\ & + 2 \cdot \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \mathbf{D}_p(x \mid h, \underline{x}, \overline{x}) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \\ & \quad \times \mathbf{D}_p(x \mid h, \underline{x}, \overline{x}) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p \left(\frac{X_i - x}{h} \right) \\ & - \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \mathbf{S}_p(x \mid h, \underline{x}, \overline{x}) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p \left(\frac{X_i - x}{h} \right). \end{aligned}$$

Then we have

$$\widehat{f}''(x) = \frac{1}{nh^3} \sum_{i=1}^n \ddot{\mathcal{K}}_p(X_i, x \mid h, \underline{x}, \overline{x}).$$

Let $\overline{f}(x) \coloneqq \mathrm{E}\left[\widehat{f}(x)\right]$, $\overline{f}'(x) \coloneqq \mathrm{E}\left[\widehat{f}'(x)\right]$ and $\overline{f}''(x) \coloneqq \mathrm{E}\left[\widehat{f}''(x)\right]$. We have the following results on the bias $\overline{f}(x) - f(x)$ and its first and second derivatives.

Proposition A.1. *Suppose that f is $(p+1)$-times continuously differentiable on $[\underline{x}, \overline{x}]$, that $h \downarrow 0$ as $n \uparrow \infty$, and that the kernel function $K(\cdot)$ satisfies Assumption S2. Then the following results hold uniformly in $x \in \mathcal{X}$.*

(a) The bias of $\widehat{f}$ admits the expansion

$$\overline{f}(x) = f(x) + \left(\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}} \boldsymbol{K}_p(u)\, u^{p+1} du \right) \frac{f^{(p+1)}(x)}{(p+1)!} h^{p+1} + o\left(h^{p+1}\right).$$

(b) $\overline{f}'(x) = f'(x) + O(h^p)$.

(c) $\overline{f}''(x) = f''(x) + O\left(h^{p-1}\right)$.

Proof of Proposition A.1. Denote $\boldsymbol{\phi}(x) := \left(f^{(0)}(x)/0!, f^{(1)}(x)/1!, ..., f^{(p)}(x)/p!\right)^\top$. Let $\mathbf{H}$ be the $(p+1)$-dimensional diagonal matrix with diagonal elements $(1, h, ..., h^p)$. By change of variables and Taylor expansion,

$$\begin{aligned}
\overline{f}(x) &= \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \int_{\underline{x}}^{\overline{x}} \frac{1}{h} \boldsymbol{K}_p\left(\frac{y-x}{h}\right) f(y)\, dy \\
&= \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}} \boldsymbol{K}_p(u) \left\{ \boldsymbol{r}_p(u)^\top \mathbf{H} \boldsymbol{\phi}(x) + \frac{f^{(p+1)}(\dot{x})(hu)^{p+1}}{(p+1)!} \right\} du,
\end{aligned}$$

where $\dot{x}$ denotes the mean value that lies between x and $x + hu$. The conclusion in Part (a) follows from this result.

Now by (S34), (S35), integration by parts and tedious algebra, we have

$$\begin{aligned}
\overline{f}'(x) &= \int_{\underline{x}}^{\overline{x}} \frac{1}{h} \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p\left(\frac{y-x}{h}\right) f'(y)\, dy \\
&\quad + \frac{1}{h} \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right) \underline{I}(x) \\
&\quad - \frac{1}{h} \mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right) \overline{I}(x),
\end{aligned} \tag{S36}$$

where

$$\underline{I}(x) := f(\underline{x}) - \boldsymbol{r}_p^\top \left(\frac{\underline{x}-x}{h}\right) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \int_{\underline{x}}^{\overline{x}} \frac{1}{h} \boldsymbol{K}_p\left(\frac{y-x}{h}\right) f(y)\, dy,$$

and $\overline{I}(x)$ is defined similarly. By Taylor expansion,

$$\begin{aligned}
\underline{I}(x) &= f(\underline{x}) - \boldsymbol{r}_p^\top \left(\frac{\underline{x}-x}{h}\right) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \int_{\underline{x}}^{\overline{x}} \frac{1}{h} \boldsymbol{K}_p\left(\frac{y-x}{h}\right) f(y)\, dy \\
&= f(\underline{x}) - \boldsymbol{r}_p^\top (\underline{x}-x)\, \boldsymbol{\phi}(x) \\
&\quad - \boldsymbol{r}_p^\top \left(\frac{\underline{x}-x}{h}\right) \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x}) \int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}} \boldsymbol{K}_p(u) \frac{f^{(p+1)}(\dot{x})(hu)^{p+1}}{(p+1)!} du,
\end{aligned} \tag{S37}$$

where $\dot{x}$ denotes the mean value that lies between x and $x + hu$. Since

$$\left| K\left(\frac{\underline{x}-x}{h}\right) \right| \lesssim \mathbb{1}\left(|\underline{x}-x| \le h\right), \tag{S38}$$

by this result, (S30) and Taylor expansion,

$$\frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}\left(x \mid h, \underline{x}, \overline{x}\right) \boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right)\left\{f\left(\underline{x}\right)-\boldsymbol{r}_p^\top\left(\underline{x}-x\right)\boldsymbol{\phi}\left(x\right)\right\} = O\left(h^p\right),$$

uniformly in $x \in \mathcal{X}$. By this result, (S30) and (S37),

$$\frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}\left(x \mid h, \underline{x}, \overline{x}\right) \boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right)\underline{I}\left(x\right) = O\left(h^p\right), \tag{S39}$$

uniformly in $x \in \mathcal{X}$. Similarly,

$$\frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}\left(x \mid h, \underline{x}, \overline{x}\right) \boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right)\overline{I}\left(x\right) = O\left(h^p\right), \tag{S40}$$

uniformly in $x \in \mathcal{X}$. By Taylor expansion and (S30),

$$\begin{aligned}
&\int_{\underline{x}}^{\overline{x}} \frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}\left(x \mid h, \underline{x}, \overline{x}\right) \boldsymbol{K}_p\left(\frac{y-x}{h}\right) f'\left(y\right) dy \\
= \;& \mathbf{e}_1^\top \mathbf{W}_p^{-1}\left(x \mid h, \underline{x}, \overline{x}\right) \int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}} \boldsymbol{K}_p\left(u\right)\left\{\boldsymbol{r}_p\left(u\right)^\top \mathbf{H}\boldsymbol{\phi}'\left(x\right) + \frac{f^{(p+1)}\left(\dot{x}\right)\left(hu\right)^p}{p!}\right\} du \\
= \;& f'\left(x\right) + O\left(h^p\right),
\end{aligned}$$

uniformly in $x \in \mathcal{X}$, where $\boldsymbol{\phi}'\left(x\right) := \left(f^{(1)}\left(x\right)/0!, f^{(2)}\left(x\right)/1!, ..., f^{(p)}\left(x\right)/\left(p-1\right)!, 0\right)^\top$ and $\dot{x}$ is the mean value that lies between x and $x+hu$. The conclusion in Part (b) follows from this result, (S36), (S39) and (S40).

For Part (c), first note that we can write

$$\begin{aligned}
\overline{f}''\left(x\right) \;=\;& \frac{d}{dx}\int_{\underline{x}}^{\overline{x}} \frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}\left(x \mid h, \underline{x}, \overline{x}\right) \boldsymbol{K}_p\left(\frac{y-x}{h}\right) f'\left(y\right) dy \\
&+ \frac{1}{h}\frac{d}{dx}\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right)\underline{I}(x) \\
&- \frac{1}{h}\frac{d}{dx}\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right)\overline{I}(x).
\end{aligned} \tag{S41}$$

By integration by parts and (S33),

$$\begin{aligned}
&\frac{d}{dx}\int_{\underline{x}}^{\overline{x}} \frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{y-x}{h}\right) f'(y)\, dy \\
=\;& \int_{\underline{x}}^{\overline{x}} \frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{y-x}{h}\right) f''(y)\, dy \\
&+ \frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right)\dot{\underline{I}}(x) \\
&- \frac{1}{h}\mathbf{e}_1^\top \mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right)\dot{\overline{I}}(x),
\end{aligned} \tag{S42}$$

where
$$\underline{\dot{I}}(x) := f'(\underline{x}) - \boldsymbol{r}_p^\top\left(\frac{\underline{x}-x}{h}\right)\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\int_{\underline{x}}^{\overline{x}}\frac{1}{h}\boldsymbol{K}_p\left(\frac{y-x}{h}\right)f'(y)\,dy,$$

and $\dot{\overline{I}}(x)$ is defined similarly. By Taylor expansion and (S30),

$$\begin{aligned}
&\int_{\underline{x}}^{\overline{x}}\frac{1}{h}\mathbf{e}_1^\top\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{y-x}{h}\right)f''(y)\,dy \\
= \;& \mathbf{e}_1^\top\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}}\boldsymbol{K}_p(u)\left\{\boldsymbol{r}_p(u)^\top\mathbf{H}\boldsymbol{\phi}''(x) + \frac{f^{(p+1)}(\dot{x})(hu)^{p-1}}{(p-1)!}\right\}du \\
= \;& f''(x) + O\left(h^{p-1}\right),
\end{aligned}$$

uniformly in $x \in \mathcal{X}$, where $\boldsymbol{\phi}''(x) := \left(f^{(2)}(x)/0!, f^{(3)}(x)/1!, ..., f^{(p)}(x)/(p-2)!, 0, 0\right)^\top$ and $\dot{x}$ is the mean value that lies between x and $x + hu$. By using arguments similar to those used in the proof of (S39) (Taylor expansion, (S30) and (S38)), we have

$$\frac{1}{h}\mathbf{e}_1^\top\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right)\underline{\dot{I}}(x) = O\left(h^{p-1}\right),$$

uniformly in $x \in \mathcal{X}$. Similarly,

$$\frac{1}{h}\mathbf{e}_1^\top\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right)\dot{\overline{I}}(x) = O\left(h^{p-1}\right),$$

uniformly in $x \in \mathcal{X}$. By these results and (S42),

$$\frac{d}{dx}\int_{\underline{x}}^{\overline{x}}\frac{1}{h}\mathbf{e}_1^\top\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\boldsymbol{K}_p\left(\frac{y-x}{h}\right)f'(y)\,dy = f''(x) + O\left(h^{p-1}\right). \tag{S43}$$

By Taylor expansion with remainder terms written in their integral forms,

$$f(\overline{x}) - \boldsymbol{r}_p^\top(\overline{x}-x)\boldsymbol{\phi}(x) = \int_x^{\overline{x}}\frac{(\overline{x}-t)^p}{p!}f^{p+1}(t)\,dt,$$

and

$$\begin{aligned}
\overline{I}(x) \;=\; & \int_x^{\overline{x}}\frac{(\overline{x}-t)^p}{p!}f^{p+1}(t)\,dt \\
& - \boldsymbol{r}_p^\top\left(\frac{\overline{x}-x}{h}\right)\mathbf{W}_p^{-1}(x \mid h, \underline{x}, \overline{x})\int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}}\boldsymbol{K}_p(u)\left\{\int_x^{x+hu}f^{(p+1)}(t)\frac{(x+hu-t)^p}{p!}dt\right\}du.
\end{aligned}$$

By calculations,

$$\frac{d}{dx}\int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}}\boldsymbol{K}_p(u)\left\{\int_x^{x+hu}f^{(p+1)}(t)\frac{(x+hu-t)^p}{p!}dt\right\}du$$

$$
\begin{aligned}
= & -\frac{1}{h}\boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right)\int_x^{\overline{x}} f^{(p+1)}(t)\,\frac{(\overline{x}-t)^p}{p!}dt+\frac{1}{h}\boldsymbol{K}_p\left(\frac{\underline{x}-x}{h}\right)\int_x^{\underline{x}} f^{(p+1)}(t)\,\frac{(\underline{x}-t)^p}{p!}dt \\
& -\frac{f^{(p+1)}(x)}{p!}\int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}}\boldsymbol{K}_p(u)\,(hu)^p\,du+\int_{\frac{\underline{x}-x}{h}}^{\frac{\overline{x}-x}{h}}\boldsymbol{K}_p(u)\left\{\int_x^{x+hu} f^{(p+1)}(t)\,\frac{(x+hu-t)^{p-1}}{(p-1)!}dt\right\}du.
\end{aligned}
$$

By these results, (S30), (S33), and

$$
\begin{aligned}
\left|K\left(\frac{\overline{x}-x}{h}\right)\right| &\lesssim \mathbb{1}\left(|\overline{x}-x|\leq h\right), \\
\left|K'\left(\frac{\overline{x}-x}{h}\right)\right| &\lesssim \mathbb{1}\left(|\overline{x}-x|\leq h\right),
\end{aligned}
$$

we have

$$
\frac{1}{h}\frac{d}{dx}\mathbf{e}_1^\top\mathbf{W}_p^{-1}\left(x\mid h,\underline{x},\overline{x}\right)\boldsymbol{K}_p\left(\frac{\overline{x}-x}{h}\right)\overline{I}(x)=O\left(h^{p-1}\right),
$$

uniformly in $x\in\mathcal{X}$. A similar result holds for the second term on the right-hand side of (S42). The conclusion in Part (c) follows from these results, (S41) and (S43). ■

Consider

$$
\begin{aligned}
\mathfrak{K} &:= \left\{\mathcal{K}_p\left(\cdot,x\mid h,\underline{x},\overline{x}\right):x\in\mathcal{X}\right\}, \\
\sigma_{\mathfrak{K}}^2 &:= \sup_{x\in\mathcal{X}}\mathrm{E}\left[\mathcal{K}_p^2\left(X_i,x\mid h,\underline{x},\overline{x}\right)\right].
\end{aligned}
$$

Let $\left(\dot{\mathfrak{K}},\sigma_{\dot{\mathfrak{K}}}^2\right)$ and $\left(\ddot{\mathfrak{K}},\sigma_{\ddot{\mathfrak{K}}}^2\right)$ be defined similarly. By change of variables and (S30), $\sigma_{\mathfrak{K}}^2=O(h)$. Similarly, $\sigma_{\dot{\mathfrak{K}}}^2\vee\sigma_{\ddot{\mathfrak{K}}}^2=O(h)$. Let $K_j(t):=t^jK(t)$ and $c_K:=\sup_{t\in\mathbb{R}}K(t)$. It follows from Giné and Nickl (2015, Proposition 3.6.12) that for all $j=0,1,...,p$,

$$
\left\{t\mapsto K_j\left(\frac{t-x}{b}\right):x\in\mathcal{X},b>0\right\}
$$

is VC-type with respect to the constant envelope c_K. By (S30), Giné and Guillou (1999, Lemma 3 (b,c)) and Chernozhukov, Chetverikov, and Kato (2014, Corollary A.1(i)), $\mathfrak{K}$ is also VC-type with respect to the constant envelope $F_{\mathfrak{K}}:=(p+1)\varpi_pc_K$, where ϖ_p is defined in (S30). By similar arguments, $\dot{\mathfrak{K}}$ and $\ddot{\mathfrak{K}}$ are also VC-type with respect to some constant envelopes $F_{\dot{\mathfrak{K}}}$ and $F_{\ddot{\mathfrak{K}}}$.

Proposition A.2. *Suppose that the assumptions in the statement of Proposition A.1 are satisfied and that* $\sqrt{|\log(h)|/(nh)}\downarrow 0$ *as* $n\uparrow\infty$. *Then the following results hold.*

(a) There exist some positive constants c_1,c_2,c_3 *which depend only on* $K(\cdot)$, *such that, for all* n *sufficiently large,*

$$
\Pr\left[\left\|\widehat{f}-\overline{f}\right\|_{\mathcal{X}}>\epsilon\right]\leq c_1\cdot\exp\left(-c_2\cdot\frac{(nh)\,\epsilon^2}{\left(\sigma_{\mathfrak{K}}^2/h\right)\vee F_{\mathfrak{K}}^2}\right) \tag{S44}
$$

if

$$c_3\left(\frac{\sigma_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\le\epsilon\le\frac{c_3}{2}.$$

(b) There exist positive constants c_1', c_2', c_3' which depend only on $K(\cdot)$, such that, for all n sufficiently large,

$$\Pr\left[\left\|\widehat{f}'-\overline{f}'\right\|_{\mathcal{X}}>\epsilon\right]\le c_1'\cdot\exp\left(-c_2'\cdot\frac{\left(nh^3\right)\epsilon^2}{\left(\sigma_{\dot{\mathfrak{K}}}^2/h\right)\vee F_{\dot{\mathfrak{K}}}^2}\right),$$

if

$$c_3'\left(\frac{\sigma_{\dot{\mathfrak{K}}}}{\sqrt{h}}\vee F_{\dot{\mathfrak{K}}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh^3}}\le\epsilon\le\frac{c_3'}{2h}.$$

(c) There exist positive constants c_1'', c_2'', c_3'' which depend only on $K(\cdot)$, such that, for all n sufficiently large,

$$\Pr\left[\left\|\widehat{f}''-\overline{f}''\right\|_{\mathcal{X}}>\epsilon\right]\le c_1''\cdot\exp\left(-c_2''\cdot\frac{\left(nh^5\right)\epsilon^2}{\left(\sigma_{\ddot{\mathfrak{K}}}^2/h\right)\vee F_{\ddot{\mathfrak{K}}}^2}\right),$$

if

$$c_3''\left(\frac{\sigma_{\ddot{\mathfrak{K}}}}{\sqrt{h}}\vee F_{\ddot{\mathfrak{K}}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh^5}}\le\epsilon\le\frac{c_3''}{2h^2}.$$

Proof of Proposition A.2. We apply Giné and Guillou (2002, Corollary 2.2) with $\mathcal{F}=\mathfrak{K}$, $\sigma=\sigma_{\mathfrak{K}}\vee\left(F_{\mathfrak{K}}\sqrt{h}\right)$ and $U=2F_{\mathfrak{K}}$. Note that $\log(U/\sigma)\le\log\left(2h^{-1/2}\right)$ and $\sqrt{n}\sigma\ge\sqrt{nh}F_{\mathfrak{K}}$ under these definitions. Therefore, we have

$$U\sqrt{\log\left(\frac{U}{\sigma}\right)}\le 2F_{\mathfrak{K}}\sqrt{\log\left(2h^{-1/2}\right)}\le\sqrt{nh}F_{\mathfrak{K}}\le\sqrt{n}\sigma,$$

when n is sufficiently large so that $\sqrt{\log\left(2h^{-1/2}\right)/(nh)}\le 1/2$. Therefore, Condition (2.5) in the statement of Giné and Guillou (2002, Corollary 2.2) is satisfied. Note that $\sigma^2/(Uh)\ge 1/2$ and

$$\frac{\sigma\sqrt{\log(U/\sigma)}}{\sqrt{n}h}\le\frac{\left(\sigma_{\mathfrak{K}}\vee\left(F_{\mathfrak{K}}\sqrt{h}\right)\right)\sqrt{\log\left(2h^{-1/2}\right)}}{\sqrt{n}h}=O\left(\sqrt{\frac{|\log(h)|}{nh}}\right).$$

Therefore, when n is sufficiently large,

$$\sqrt{n}\sigma\sqrt{\log\left(\frac{U}{\sigma}\right)}<\frac{n\sigma^2}{U}.$$

The conclusion in Part (a) follows from applying Giné and Guillou (2002, Corollary 2.2). The conclusions in Part (b) and Part (c) follow from the same arguments. ■

Corollary A.1. *Suppose that the assumptions in the statement of Proposition A.2 are satisfied. Then we have the following results. (a) There exists some $M > 0$ such that*

$$\Pr\left[\left\|\widehat{f}-\overline{f}\right\|_{\mathcal{X}} > M\left(\frac{\sigma_{\mathfrak{K}}}{\sqrt{h}} \vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right] = O\left(n^{-1}\right).$$

(b) There exists some $M' > 0$ such that

$$\Pr\left[\left\|\widehat{f}'-\overline{f}'\right\|_{\mathcal{X}} > M'\left(\frac{\sigma_{\dot{\mathfrak{K}}}}{\sqrt{h}} \vee F_{\dot{\mathfrak{K}}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh^3}}\right] = O\left(n^{-1}\right).$$

(c) There exists some $M'' > 0$ such that

$$\Pr\left[\left\|\widehat{f}''-\overline{f}''\right\|_{\mathcal{X}} > M''\left(\frac{\sigma_{\ddot{\mathfrak{K}}}}{\sqrt{h}} \vee F_{\ddot{\mathfrak{K}}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh^5}}\right] = O\left(n^{-1}\right).$$

Proof of Corollary A.1. For all $C > 1$, $C\left(\left(\sigma_{\mathfrak{K}}/\sqrt{h}\right) \vee F_{\mathfrak{K}}\right)\sqrt{\log\left(2h^{-1/2}\right)/(nh)} < 1/2$ if n is sufficiently large. Therefore, for $\epsilon = Cc_3\left(\left(\sigma_{\mathfrak{K}}/\sqrt{h}\right) \vee F_{\mathfrak{K}}\right)\sqrt{\log\left(2h^{-1/2}\right)/(nh)}$, (S44) holds if n is sufficiently large. And it is easy to see that the right-hand side of (S44) is $O\left(n^{-1}\right)$ if C is taken to be sufficiently large. The conclusions in Part (b) and Part (c) follow from similar arguments. ■

Let $\left\{X_1^\dagger, ..., X_n^\dagger\right\}$ be a nonparametric bootstrap sample from $\{X_1, ..., X_n\}$. Let $\widehat{f}_\dagger(x)$ be defined by the right-hand side of (S31) with X_i replaced by $X_i^\dagger$. Let

$$\begin{aligned}\mathfrak{Q} &:= \left\{\mathcal{K}_p^2(\cdot, x \mid h, \underline{x}, \overline{x}) : x \in \mathcal{X}\right\},\\ \sigma_{\mathfrak{Q}}^2 &:= \sup_{x\in\mathcal{X}} \mathrm{E}\left[\mathcal{K}_p^4(X, x \mid h, \underline{x}, \overline{x})\right].\end{aligned}$$

The following result is a bootstrap analogue of Corollary A.1.

Corollary A.2. *Assume that the assumptions in the statement of Proposition A.2 are satisfied. Then we have the following results. (a) There exists some positive constants (M_1, M_2) such that*

$$\Pr_\dagger\left[\left\|\widehat{f}_\dagger-\overline{f}\right\|_{\mathcal{X}} > M_1\left(\sqrt{\frac{\sigma_{\mathfrak{K}}^2}{h} + M_2\left(\frac{\sigma_{\mathfrak{Q}}}{\sqrt{h}} \vee F_{\mathfrak{K}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}} \vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right] = O_p\left(n^{-1}\right).$$

(b) There exists some positive constants (M_1', M_2') such that

$$\Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}'-\overline{f}'\right\|_{\mathcal{X}}>M_1'\left(\sqrt{\frac{\sigma_{\dot{\mathfrak{K}}}^2}{h}+M_2'\left(\frac{\sigma_{\dot{\mathfrak{Q}}}}{\sqrt{h}}\vee F_{\dot{\mathfrak{K}}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}}\vee F_{\dot{\mathfrak{K}}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh^3}}\right]$$
$$=O_p\left(n^{-1}\right).$$

(c) There exists some positive constants (M_1'', M_2'') such that

$$\Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}''-\overline{f}''\right\|_{\mathcal{X}}>M_1''\left(\sqrt{\frac{\sigma_{\ddot{\mathfrak{K}}}^2}{h}+M_2''\left(\frac{\sigma_{\ddot{\mathfrak{Q}}}}{\sqrt{h}}\vee F_{\ddot{\mathfrak{K}}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}}\vee F_{\ddot{\mathfrak{K}}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh^5}}\right]$$
$$=O_p\left(n^{-1}\right).$$

Proof of Corollary A.2. Let

$$\widehat{q}\,(x)\coloneqq\frac{1}{nh}\sum_{i=1}^{n}\mathcal{K}_p^2\left(X_i,x\mid h,\underline{x},\overline{x}\right)$$

and $\overline{q}\,(x)\coloneqq \mathrm{E}\left[\widehat{q}\,(x)\right]$. By Chernozhukov, Chetverikov, and Kato (2014, Corollary A.1(ii)), $\mathfrak{Q}$ is VC-type with respect to the constant envelope $F_{\mathfrak{K}}^2$. By change of variables and (S30), $\sigma_{\mathfrak{Q}}^2=O\left(h\right)$. By similar arguments as those used in the proof of Part (a) of Corollary A.1, for some $M_2>0$,

$$\Pr\left[\left\|\widehat{q}-\overline{q}\right\|_{\mathcal{X}}>M_2\left(\frac{\sigma_{\mathfrak{Q}}}{\sqrt{h}}\vee F_{\mathfrak{K}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right]=O\left(n^{-1}\right).\tag{S45}$$

Let

$$\widehat{\sigma}_{\mathfrak{K}}^2\coloneqq\sup_{x\in\mathcal{X}}\mathrm{E}_{\dagger}\left[\mathcal{K}_p^2\left(X_i^{\dagger},x\mid h,\underline{x},\overline{x}\right)\right].$$

Since $\|\widehat{q}\|_{\mathcal{X}}=\widehat{\sigma}_{\mathfrak{K}}^2/h$ and $\|\overline{q}\|_{\mathcal{X}}=\sigma_{\mathfrak{K}}^2/h$, it follows from (S45) and the triangle inequality that

$$\Pr\left[\left|\frac{\widehat{\sigma}_{\mathfrak{K}}^2}{h}-\frac{\sigma_{\mathfrak{K}}^2}{h}\right|>M_2\left(\frac{\sigma_{\mathfrak{Q}}}{\sqrt{h}}\vee F_{\mathfrak{K}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right]=O\left(n^{-1}\right)\tag{S46}$$

and $\left|\widehat{\sigma}_{\mathfrak{K}}^2/h-\sigma_{\mathfrak{K}}^2/h\right|=O_p\left(\sqrt{\log\left(2h^{-1/2}\right)/\left(nh\right)}\right)$. By Proposition A.2(a),

$$\Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}-\widehat{f}\right\|_{\mathcal{X}}>\epsilon\right]\leq c_1\cdot\exp\left(-c_2\cdot\frac{\left(nh\right)\epsilon^2}{\left(\widehat{\sigma}_{\mathfrak{K}}^2/h\right)\vee F_{\mathfrak{K}}^2}\right)$$

if

$$c_3\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\leq\epsilon\leq\frac{c_3}{2}.$$

By this result, for all $C > 1$,

$$\Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}-\widehat{f}\right\|_{\mathcal{X}} > Cc_3\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right] \leq c_1\left(2h^{-1/2}\right)^{-c_2c_3^2C^2} + \mathbb{1}\left(\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}} > \frac{1}{2C}\right).$$

By $\left|\widehat{\sigma}_{\mathfrak{K}}^2/h-\sigma_{\mathfrak{K}}^2/h\right| = O_p\left(\sqrt{\log\left(2h^{-1/2}\right)/(nh)}\right)$ and the fact that $\sigma_{\mathfrak{K}}^2/h = O(1)$, the second term on the right-hand side of the above inequality is zero wpa1. Therefore,

$$\Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}-\widehat{f}\right\|_{\mathcal{X}} > Cc_3\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right] = O_p\left(n^{-1}\right), \tag{S47}$$

if C is taken to be sufficiently large. By the triangle inequality,

$$\begin{aligned}\Pr_{\dagger}&\left[\left\|\widehat{f}_{\dagger}-\overline{f}\right\|_{\mathcal{X}} > Cc_3\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee\frac{\sigma_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right]\\ &\leq \Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}-\widehat{f}\right\|_{\mathcal{X}} > \frac{Cc_3}{2}\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right]\\ &\quad + \mathbb{1}\left(\left\|\widehat{f}-\overline{f}\right\|_{\mathcal{X}} > \frac{Cc_3}{2}\left(\frac{\sigma_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right).\end{aligned}$$

By Corollary A.1(a), the second term on the right-hand side of the above inequality is zero wpa1, if C is taken to be sufficiently large. By this result and (S47),

$$\Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}-\overline{f}\right\|_{\mathcal{X}} > Cc_3\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee\frac{\sigma_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right] = O_p\left(n^{-1}\right), \tag{S48}$$

if C is taken to be sufficiently large. Fix any $C > 0$ and take M_2 to be such that (S45) holds, we have

$$\begin{aligned}\Pr_{\dagger}&\left[\left\|\widehat{f}_{\dagger}-\overline{f}\right\|_{\mathcal{X}} > Cc_3\left(\sqrt{\frac{\sigma_{\mathfrak{K}}^2}{h}+M_2\left(\frac{\sigma_{\mathfrak{Q}}}{\sqrt{h}}\vee F_{\mathfrak{K}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right]\\ &\leq \Pr_{\dagger}\left[\left\|\widehat{f}_{\dagger}-\overline{f}\right\|_{\mathcal{X}} > Cc_3\left(\frac{\widehat{\sigma}_{\mathfrak{K}}}{\sqrt{h}}\vee\frac{\sigma_{\mathfrak{K}}}{\sqrt{h}}\vee F_{\mathfrak{K}}\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right]\end{aligned}$$

$$+\mathbb{1}\left(\left|\frac{\widehat{\sigma}_{\mathfrak{K}}^2}{h}-\frac{\sigma_{\mathfrak{K}}^2}{h}\right|>M_2\left(\frac{\sigma_{\mathfrak{Q}}}{\sqrt{h}}\vee F_{\mathfrak{K}}^2\right)\sqrt{\frac{\log\left(2h^{-1/2}\right)}{nh}}\right).$$

It follows from (S46) that the second term on the right-hand side of the above inequality is 0 wpa1. It follows from (S48) that the first term on the right-hand side of the above inequality is $O_p\left(n^{-1}\right)$, if C is taken to be sufficiently large. The conclusion in Part (a) can be deduced from the above inequality and these results.

The conclusions in Parts (b) and (c) follow from similar arguments. ∎

Appendix B Proofs of the results in Section S3.3

Denote $\mathcal{B}_n \coloneqq \left[\underline{b}_n, \overline{b}_n\right]$. Smoothness results similar to those in Guerre, Perrigne, and Vuong (2000, Proposition 1 and Lemmas A1 and A2) are summarized in the following lemma. Let

$$R(v\mid p_n,n)\coloneqq\frac{\int_v^{\overline{v}}H(t\mid p_n,n)\,dt}{H^2(v\mid p_n,n)}.$$

Let $\xi'(\cdot\mid p_n,n)$ denote the derivative of $\xi(\cdot\mid p_n,n)$. $R'(\cdot\mid p_n,n)$, $R''(\cdot\mid p_n,n)$, $H''(\cdot\mid p_n,n)$, $g'(\cdot\mid n)$, $g''(\cdot\mid n)$, $g'(\cdot,n)$ and $g''(\cdot,n)$ are defined similarly.

Lemma S4. *Suppose that the assumptions in the statement of Proposition S3 are satisfied. Then we have the following results. (a) $\xi(\cdot\mid p_n,n)$ is thrice continuously differentiable on $\mathcal{B}_n$ and $\xi'(\cdot\mid p_n,n)$ is bounded away from zero on $\mathcal{B}_n$. (b) $g(\cdot\mid n)$ is twice continuously differentiable on $\mathcal{B}_n$ and $g(\cdot\mid n)$ is bounded away from zero on $\mathcal{B}_n$.*

Proof of Lemma S4. For any $v\in(\underline{v},\overline{v})$, we have

$$\begin{aligned}\beta'(v\mid p_n,n) &= -\frac{\beta'(v\mid p_n,n)-v}{H(v\mid p_n,n)}\cdot H'(v\mid p_n,n)\\ &= -H'(v\mid p_n,n)\,R(v\mid p_n,n)\\ &= (n-1)\,p_nf^*(v\mid p_n)\,(1-p_nF^*(v\mid p_n))^{n-2}\,R(v\mid p_n,n)\,,\end{aligned}\tag{S49}$$

where

$$H'(v\mid p_n,n)=-(n-1)\,p_nf^*(v\mid p_n)\,(1-p_nF^*(v\mid p_n))^{n-2}\,.$$

Moreover, $\beta'(\cdot\mid p_n,n)$ is continuous on $(\underline{v},\overline{v})$ and $\beta'(v\mid p_n,n)>0$ for all $v\in(\underline{v},\overline{v})$. By L'Hopital rule, $\lim_{v\uparrow\overline{v}}R(v\mid p_n,n)=-\left(2H'(\overline{v}\mid p_n,n)\right)^{-1}>0$, where

$$H'(\overline{v}\mid p_n,n)=-(n-1)\,p_n\,(1-p_n)^{n-2}\,f^*(\overline{v}\mid p_n)<0,$$

and hence, $\lim_{v\uparrow\overline{v}}\beta'(v\mid p_n,n)=1/2$. A similar argument shows that $0<\lim_{v\downarrow\underline{v}}\beta'(v\mid p_n,n)<\infty$. Therefore, $\beta(\cdot\mid p_n,n)$ is continuously differentiable on $[\underline{v},\overline{v}]$ and $\beta'(\cdot\mid p_n,n)$ is bounded away from zero on $[\underline{v}, \overline{v}]$. By the inverse function theorem,

$$\xi'(b \mid p_n, n) = \frac{1}{\beta'(\xi(b \mid p_n, n) \mid p_n, n)}, \tag{S50}$$

for $b \in (\underline{b}_n, \overline{b}_n)$. It follows that $\xi(\cdot \mid p_n, n)$ is continuously differentiable on $\mathcal{B}_n$ and $\xi'(\cdot \mid p_n, n)$ is bounded away from zero on $\mathcal{B}_n$. By the quotient rule, for $v \in (\underline{v}, \overline{v})$,

$$R'(v \mid p_n, n) = -\frac{H^2(v \mid p_n, n) + 2H'(v \mid p_n, n)\left(\int_v^{\overline{v}} H(t \mid p_n, n)\, dt\right)}{H^3(v \mid p_n, n)}$$

and

$$\begin{aligned} R''(v \mid p_n, n) = -H^{-4}(v \mid p_n, n) \Bigg\{ & 2H''(v \mid p_n, n) H(v \mid p_n, n) \left(\int_v^{\overline{v}} H(t \mid p_n, n)\, dt\right) \\ & -3H^2(v \mid p_n, n) H'(v \mid p_n, n) - 6\left(\int_v^{\overline{v}} H(t \mid p_n, n)\, dt\right) \left(H'(v \mid p_n, n)\right)^2 \Bigg\}. \end{aligned}$$

Then, it is easy to check that the limits of $R'(v \mid p_n, n)$ and $R''(v \mid p_n, n)$ (as $v \downarrow \underline{v}$ or $v \uparrow \overline{v}$) all exist and are finite. It follows from this fact and (S49) that $\beta'(\cdot \mid p_n, n)$ is twice continuously differentiable on $[\underline{v}, \overline{v}]$. It follows from this fact and (S50) that $\xi'(\cdot \mid p_n, n)$ is twice continuously differentiable on $\mathcal{B}_n$.

The conclusion in Part (b) follows from results in Part (a) and

$$g(b \mid n) = f^*(\xi(b \mid p_n, n) \mid p_n)\, \xi'(b \mid p_n, n).$$

■

Denote

$$\widehat{G}(b, n) := \frac{1}{L} \sum_{l: N_l = n} \sum_{i=1}^{N_l^*} \mathbb{1}(B_{il} \le b),$$

and $G(b, n) := \mathrm{E}\left[\widehat{G}(b, n)\right] = r_n G(b \mid n)$, where

$$\begin{aligned} \pi_n &:= \Pr[N_l = n], \\ r_n &:= \mathrm{E}\left[\mathbb{1}(N_l = n) N_l^*\right] \\ &= \pi_n q_n n, \end{aligned}$$

and the second equality follows from LIE. Let

$$\widehat{g}(b, n) := \frac{1}{L} \sum_{l: N_l = n} \sum_{i=1}^{N_l^*} \frac{1}{h} \mathcal{K}_1\left(B_{il}, b \mid h, \widehat{\underline{b}}_n, \widehat{\overline{b}}_n\right) \tag{S51}$$

and $g(b,n) \coloneqq r_n g(b \mid n)$. We can now write $\widehat{G}(b \mid n) \coloneqq \widehat{G}(b,n)/\widehat{r}_n$ and $\widehat{g}(b \mid n) = \widehat{g}(b,n)/\widehat{r}_n$. Denote $\mathbb{G}(b \mid n) \coloneqq \widehat{G}(b \mid n) - G(b \mid n)$ and $\mathbb{H}(b \mid n) \coloneqq \widehat{g}(b \mid n) - g(b \mid n)$. Let $\mathbb{G}(b,n) \coloneqq \widehat{G}(b,n) - G(b,n)$ and $\mathbb{H}(b,n) \coloneqq \widehat{g}(b,n) - g(b,n)$. Let $\widehat{g}'(\cdot \mid n)$, $\widehat{g}''(\cdot \mid n)$, $\widehat{g}'(\cdot, n)$ and $\widehat{g}''(\cdot, n)$ be the first and second derivatives of $\widehat{g}(\cdot \mid n)$ and $\widehat{g}(\cdot, n)$. Let $\mathbb{H}'(\cdot \mid n)$ and $\mathbb{H}''(\cdot \mid n)$ be the first and second derivatives of $\mathbb{H}(\cdot \mid n)$. Let

$$\widehat{\pi}_n \coloneqq \frac{1}{L} \sum_{l=1}^{L} \mathbb{1}(N_l = n).$$

Then we can write $\widehat{q}_n = \widehat{r}_n / (n \widehat{\pi}_n)$. The following lemma collects results on the rates of convergence of $\widehat{q}_n$, $\widehat{G}(\cdot \mid n)$ and $\widehat{g}(\cdot \mid n)$ (and its derivatives).

Lemma S5. *Suppose that the assumptions in the statement of Proposition S3 are satisfied. Then we have the following results. (a)*

$$\Pr\left[|\widehat{p}_n - p_n| \geq \alpha_L^p\right] = O\left(L^{-1}\right),$$

for some $\alpha_L^p = O\left(\sqrt{\log(L)/L}\right)$. *(b)*

$$\Pr\left[\|\mathbb{G}(\cdot \mid n)\|_{\mathcal{B}_n} \geq \bar{\alpha}_L\right] = O\left(L^{-1}\right),$$

for some $\bar{\alpha}_L = O\left(\sqrt{\log(L)/L}\right)$. *(c)*

$$\Pr\left[\|\mathbb{H}(\cdot \mid n)\|_{\mathcal{B}_n} \geq \alpha_L\right] = O\left(L^{-1}\right),$$

for some $\alpha_L = O\left(\sqrt{\log(L)/(Lh)}\right)$ *and similar results with* $\alpha_L' = O\left(\sqrt{\log(L)/(Lh^3)}\right)$ *and* $\alpha_L'' = O\left(\sqrt{\log(L)/(Lh^5)}\right)$ *hold for* $\|\mathbb{H}'(\cdot \mid n)\|_{\mathcal{B}_n}$ *and* $\|\mathbb{H}''(\cdot \mid n)\|_{\mathcal{B}_n}$.

Proof of Lemma S5. By Bernstein's inequality (Giné and Nickl, 2015, Theorem 3.1.7),

$$\Pr\left[|\widehat{\pi}_n - \pi_n| \geq \sigma_{\pi,n} \sqrt{2 \cdot \frac{\log(L)}{L}}\right] = O\left(L^{-1}\right), \tag{S52}$$

where $\sigma_{\pi,n}^2 \coloneqq \pi_n - \pi_n^2$. By this result and simple calculation,

$$\Pr\left[\left|\frac{\pi_n}{\widehat{\pi}_n} - 1\right| \geq \frac{\sigma_{\pi,n}\sqrt{2 \cdot \log(L)/L}}{\pi_n - \sigma_{\pi,n}\sqrt{2 \cdot \log(L)/L}}\right] = O\left(L^{-1}\right). \tag{S53}$$

Similarly, by Bernstein's inequality,

$$\Pr\left[|\widehat{r}_n - r_n| \geq \sigma_{r,n} \sqrt{2 \cdot \frac{\log(L)}{L}}\right] = O\left(L^{-1}\right), \tag{S54}$$

where $\sigma_{r,n} \coloneqq \pi_n \mathrm{E}\left[(N_l^*)^2 \mid N_l = n\right] - r_n^2$. By this result, (S53) and

$$\widehat{q}_n - q_n = \frac{\widehat{r}_n - r_n}{n\pi_n} + \frac{\widehat{r}_n}{n\pi_n}\left(\frac{\pi_n}{\widehat{\pi}_n} - 1\right),$$

we have $\Pr\left[|\widehat{q}_n - q_n| \geq \alpha_L^q\right] = O\left(L^{-1}\right)$ for some $\alpha_L^q = O\left(\sqrt{\log(L)/L}\right)$. Then, by the mean value theorem and taking $\alpha_L^p \coloneqq \left\|\left(\varphi_n^{-1}\right)'\right\|_{\left[q_n \pm \alpha_L^q\right] \cap [0,1]} \alpha_L^q$, we have

$$\Pr\left[|\widehat{p}_n - p_n| \geq \alpha_L^p\right] \leq \Pr\left[|\widehat{q}_n - q_n| \geq \alpha_L^q\right] = O\left(L^{-1}\right).$$

By (S54) and simple calculation,

$$\Pr\left[\left|\frac{r_n}{\widehat{r}_n} - 1\right| \geq \frac{\sigma_{r,n}\sqrt{2 \cdot \log(L)/L}}{r_n - \sigma_{r,n}\sqrt{2 \cdot \log(L)/L}}\right] = O\left(L^{-1}\right). \tag{S55}$$

Let $(B_{1l}, B_{2l}, ..., B_{nl})$ be i.i.d., $\boldsymbol{B}_l \coloneqq (B_{1l}, ..., B_{nl}, N_l^*, N_l)^\top$ and

$$\mathcal{G}(\boldsymbol{B}_l; b) \coloneqq \mathbb{1}(N_l = n) \sum_{i=1}^{N_l^*} \mathbb{1}(B_{il} \leq b).$$

Then we can write $\widehat{G}(b, n) \coloneqq L^{-1} \sum_{l=1}^L \mathcal{G}(\boldsymbol{B}_l; b)$. By Kosorok (2008, Lemma 9.8), Giné and Nickl (2015, Theorem 3.6.9) and Nolan and Pollard (1987, Corollary 17), $\{\mathcal{G}(\cdot; b) : b \in \mathcal{B}_n\}$ is VC-type with respect to the constant envelope n. Then we apply Giné and Guillou (2002, Corollary 2.2) with $U = 2n$, $\sigma^2 = n^2\pi_n$ and t taken to be $C\sqrt{\log(L)\, L}$ for some positive constant C. Note that Equations (2.5) and (2.6) of Giné and Guillou (2002) hold for all L large enough. By Giné and Guillou (2002, Corollary 2.2), taking C to be sufficiently large, we have

$$\Pr\left[\|\mathbb{G}(\cdot, n)\|_{\mathcal{B}_n} \geq \breve{\alpha}_L\right] = O\left(L^{-1}\right),$$

for some $\breve{\alpha}_L = O\left(\sqrt{\log(L)/L}\right)$. The conclusion in Part (b) follows from this result, (S55) and

$$\mathbb{G}(b \mid n) = \frac{\mathbb{G}(b, n)}{r_n} + \frac{\widehat{G}(b, n)}{r_n}\left(\frac{r_n}{\widehat{r}_n} - 1\right).$$

By (straightforward adaptations of) Proposition A.1(a) and Corollary A.1(a),

$$\Pr\left[\|\mathbb{H}(\cdot, n)\|_{\mathcal{B}_n} \geq \alpha_L^g\right] = O\left(L^{-1}\right),$$

for some $\alpha_L^g = O\left(\sqrt{\log(L)/(Lh)}\right)$. The first conclusion in Part (c) follows from this result, (S55) and

$$\mathbb{H}(b \mid n) = \frac{\mathbb{H}(b, n)}{r_n} + \frac{\widehat{g}(b, n)}{r_n}\left(\frac{r_n}{\widehat{r}_n} - 1\right). \tag{S56}$$

The other results follow from similar arguments, Proposition A.1(b,c) and Corollary A.1(b,c). ■

Proof of Proposition S3. Let

$$\widehat{F}^*(v,n) \quad \coloneqq \quad \frac{1}{L}\sum_{l:N_l=n}\sum_{i=1}^{N_l^*}\mathbb{1}\left(\widehat{V}_{il}\leq v\right).$$

Now we can write $\widehat{F}^*(v\mid p_n) \coloneqq \widehat{F}^*(v,n)/\widehat{r}_n$. Denote

$$\widetilde{\xi}(b\mid n) \coloneqq b - \frac{\eta_n(p_n, G(b\mid n))}{(n-1)\,\widehat{g}(b\mid n)}.$$

Let $\widetilde{V}_{il} \coloneqq \widetilde{\xi}(B_{il}\mid N_l)$ and $\widetilde{F}^*(v\mid p_n) \coloneqq \widetilde{F}^*(v,n)/r_n$, where

$$\widetilde{F}^*(v,n) \coloneqq \frac{1}{L}\sum_{l:N_l=n}\sum_{i=1}^{N_l^*}\mathbb{1}\left(\widetilde{V}_{il}\leq v\right).$$

Then, we decompose

$$\widehat{F}^*(v\mid p_n) - F^*(v\mid p_n) = \left\{\widetilde{F}^*(v\mid p_n) - F^*(v\mid p_n)\right\} + \left\{\widehat{F}^*(v\mid p_n) - \widetilde{F}^*(v\mid p_n)\right\}, \tag{S57}$$

and

$$\widehat{F}^*(v\mid p_n) - \widetilde{F}^*(v\mid p_n) = \frac{\widehat{F}^*(v,n) - \widetilde{F}^*(v,n)}{r_n} + \frac{\widehat{F}^*(v,n)}{r_n}\left(\frac{r_n}{\widehat{r}_n} - 1\right). \tag{S58}$$

Let $\mathbb{X}(b\mid n) \coloneqq \widetilde{\xi}(b\mid n) - \xi(b\mid p_n, n)$. Let $\mathbb{X}'(\cdot\mid n)$ and $\mathbb{X}''(\cdot\mid n)$ denote the first and second derivatives of $\mathbb{X}(\cdot\mid n)$. Then, by straightforward calculation, we have

$$\mathbb{X}(b\mid n) = -\frac{\eta_n(p_n, G(b\mid n))}{n-1}\left\{\frac{1}{\widehat{g}(b\mid n)} - \frac{1}{g(b\mid n)}\right\}, \tag{S59}$$

$$\begin{aligned}\mathbb{X}'(b\mid n) \quad = \quad & -\frac{D_2\eta_n(p_n, G(b\mid n))\,g(b\mid n)}{n-1}\left\{\frac{1}{\widehat{g}(b\mid n)} - \frac{1}{g(b\mid n)}\right\} \\ & + \frac{\eta_n(p_n, G(b\mid n))}{n-1}\left\{\frac{\widehat{g}'(b\mid n)}{\widehat{g}^2(b\mid n)} - \frac{g'(b\mid n)}{g^2(b\mid n)}\right\},\end{aligned} \tag{S60}$$

and

$$\begin{aligned}\mathbb{X}''(b\mid n) \quad = \quad & -\frac{D_2^2\eta_n(p_n,G(b\mid n))g^2(b\mid n) + D_2\eta_n(p_n,G(b\mid n))g'(b\mid n)}{n-1}\left\{\frac{1}{\widehat{g}(b\mid n)} - \frac{1}{g(b\mid n)}\right\} \\ & + \frac{2D_2\eta_n(p_n,G(b\mid n))g(b\mid n)}{n-1}\left\{\frac{\widehat{g}'(b\mid n)}{\widehat{g}^2(b\mid n)} - \frac{g'(b\mid n)}{g^2(b\mid n)}\right\} \\ & + \frac{\eta_n(p_n,G(b\mid n))}{n-1}\left\{\frac{\widehat{g}''(b\mid n)\widehat{g}^2(b\mid n) - 2(\widehat{g}'(b\mid n))^2\widehat{g}(b\mid n)}{\widehat{g}^4(b\mid n)}\right.\end{aligned}$$

$$
-\frac{g''(b \mid n) g^2(b \mid n) - 2(g'(b \mid n))^2 g(b \mid n)}{g^4(b \mid n)}\Bigg\}. \tag{S61}
$$

Denote

$$
\mathbb{K}\left(b \mid n\right) \coloneqq \eta_n\left(\widehat{p}_n, \widehat{G}\left(b \mid n\right)\right) - \eta_n\left(p_n, G\left(b \mid n\right)\right).
$$

By straightforward calculation, we have

$$
\left\|\mathbb{K}\left(\cdot \mid n\right)\right\|_{\mathcal{B}_n} \leq \left(\left\|D_1\eta_n\right\|_{\left[p_n \pm \alpha_L^p\right]\times[0,1]} \vee \left\|D_2\eta_n\right\|_{\left[p_n \pm \alpha_L^p\right]\times[0,1]}\right)\left(\left|\widehat{p}_n - p_n\right| + \left\|\mathbb{G}\left(\cdot \mid n\right)\right\|_{\mathcal{B}_n}\right),
$$

if $|\widehat{p}_n - p_n| < \alpha_L^p$. We have $\|D_1\eta_n\|_{[p_n\pm\alpha_L^p]\times[0,1]} = O(1)$ and $\|D_2\eta_n\|_{[p_n\pm\alpha_L^p]\times[0,1]} = O(1)$ by straightforward calculation. Therefore, by these results and Lemma S5(a,b),

$$
\Pr\left[\left\|\mathbb{K}\left(\cdot \mid n\right)\right\|_{\mathcal{B}_n} \geq \tilde{\alpha}_L\right] = O\left(L^{-1}\right), \tag{S62}
$$

for some $\tilde{\alpha}_L = O\left(\sqrt{\log(L)/L}\right)$. Let $\bar{\mathbb{T}} \coloneqq \mathbb{1}\left(\|\mathbb{K}(\cdot \mid n)\|_{\mathcal{B}_n} < \tilde{\alpha}_L\right)$. Let $(\mathbb{T}, \mathbb{T}', \mathbb{T}'')$ be defined by the same formula with $(\mathbb{K}(\cdot \mid n), \tilde{\alpha}_L)$ replaced by $(\mathbb{H}, \alpha_L)$, $(\mathbb{H}', \alpha_L')$ and $(\mathbb{H}'', \alpha_L'')$, where $(\alpha_L, \alpha_L', \alpha_L'')$ are defined in the statement of Lemma S5. Let $\mathbb{I} \coloneqq \bar{\mathbb{T}}\mathbb{T}\mathbb{T}'\mathbb{T}''$. It follows from (S62) and Lemma S5 that $\Pr[\mathbb{I} = 0] = O\left(L^{-1}\right)$.

Note that

$$
\widehat{\xi}\left(b \mid n\right) - \widetilde{\xi}\left(b \mid n\right) = -\frac{\mathbb{K}\left(b \mid n\right)}{(n-1)\,\widehat{g}\left(b \mid n\right)}.
$$

Denote $\delta_L \coloneqq \left\|\widehat{\xi}(\cdot \mid n) - \widetilde{\xi}(\cdot \mid n)\right\|_{\mathcal{B}_n}$. Then, by the triangle inequality,

$$
\left|\widehat{F}^*\left(v, n\right) - \widetilde{F}^*\left(v, n\right)\right| \leq \frac{1}{L}\sum_{l:N_l=n}\sum_{i=1}^{N_l^*}\left\{\mathbb{1}\left(\widetilde{V}_{il} \leq v + \delta_L\right) - \mathbb{1}\left(\widetilde{V}_{il} \leq v - \delta_L\right)\right\}.
$$

Let $\overline{v}_n \coloneqq \widetilde{\xi}\left(\overline{b}_n \mid n\right)$ and $\underline{v}_n \coloneqq \widetilde{\xi}\left(\underline{b}_n \mid n\right)$. By direct calculation, we have $\|\eta_n(p_n, \cdot)\|_{[0,1]} < \infty$ and $\|D_2\eta_n(p_n, \cdot)\|_{[0,1]} < \infty$. By these results, (S60) and Lemma S4(b), $\|\mathbb{X}'(\cdot \mid n)\|_{\mathcal{B}_n}$ is sufficiently small, when $\mathbb{I} = 1$ and L is large enough. By Lemma S4(a), $\widetilde{\xi}(\cdot \mid n)$ is strictly increasing on $\mathcal{B}_n$ and its inverse function $\widetilde{\beta}(\cdot \mid n) \coloneqq \widetilde{\xi}^{-1}(\cdot \mid n)$ exists, if $\mathbb{I} = 1$ and L is sufficiently large. $\widetilde{\beta}(\cdot \mid n)$ is a strictly increasing function on $[\underline{v}_n, \overline{v}_n]$. We can write

$$
\mathbb{1}\left(\widetilde{\xi}\left(B_{il} \mid n\right) \leq y\right) = \mathbb{1}\left(y \geq \overline{v}_n\right) + \mathbb{1}\left(y \in \left(\underline{v}_n, \overline{v}_n\right)\right)\mathbb{1}\left(B_{il} \leq \widetilde{\beta}\left(y \mid n\right)\right) \tag{S63}
$$

when $\widetilde{\xi}(\cdot \mid n)$ is strictly increasing. If $\mathbb{I} = 1$ and L is large enough, $\delta_L \lesssim \sqrt{\log(L)/L}$, $|\overline{v}_n - \overline{v}| \lesssim \sqrt{\log(L)/(Lh)}$ and $|\underline{v}_n - \underline{v}| \lesssim \sqrt{\log(L)/(Lh)}$. We have

$$
\mathbb{I} \cdot \frac{1}{L}\sum_{l:N_l=n}\sum_{i=1}^{N_l^*}\left\{\mathbb{1}\left(\widetilde{V}_{il} \leq v + \delta_L\right) - \mathbb{1}\left(\widetilde{V}_{il} \leq v - \delta_L\right)\right\}
$$

$$= \mathbb{I}\left\{\widehat{G}\left(\widetilde{\beta}\left(v+\delta_L \mid n\right), n\right) - \widehat{G}\left(\widetilde{\beta}\left(v-\delta_L \mid n\right), n\right)\right\},$$

when L is sufficiently large. Then,

$$\begin{aligned} \mathbb{I}\left|\widehat{F}^*\left(v,n\right) - \widetilde{F}^*\left(v,n\right)\right| &\leq \mathbb{I}\left\{\mathbb{G}\left(\widetilde{\beta}\left(v+\delta_L \mid n\right), n\right) - \mathbb{G}\left(\widetilde{\beta}\left(v-\delta_L \mid n\right), n\right)\right\} \\ &\quad + \mathbb{I}\left\{G\left(\widetilde{\beta}\left(v+\delta_L \mid n\right), n\right) - G\left(\widetilde{\beta}\left(v-\delta_L \mid n\right), n\right)\right\}, \end{aligned} \tag{S64}$$

when L is sufficiently large. By Dette, Neumeyer, and Pilz (2006, Lemma A.1), if $\mathbb{I} = 1$ and L is large enough, there exists $\lambda \in (0,1)$ such that

$$\widetilde{\beta}\left(v+\delta_L \mid n\right) - \beta\left(v+\delta_L \mid p_n, n\right) = -\frac{\mathbb{X}\left(b_\lambda \mid n\right)}{\xi'\left(b_\lambda \mid p_n, n\right) + \lambda \mathbb{X}'\left(b_\lambda \mid n\right)},$$

where $b_\lambda \coloneqq \left(\xi\left(\cdot \mid p_n, n\right) + \lambda\mathbb{X}\left(\cdot \mid n\right)\right)^{-1}\left(v+\delta_L\right)$ and by (S59), (S60), Lemmas S4 and S5,

$$\begin{aligned} \left|\widetilde{\beta}\left(v+\delta_L \mid n\right) - \beta\left(v+\delta_L \mid p_n, n\right)\right| &\leq \frac{\left\|\mathbb{X}\left(\cdot \mid n\right)\right\|_{\mathcal{B}_n}}{\inf_{b\in\mathcal{B}_n}\xi'\left(b \mid p_n, n\right) - \left\|\mathbb{X}'\left(\cdot \mid n\right)\right\|_{\mathcal{B}_n}} \\ &= O\left(\sqrt{\frac{\log\left(L\right)}{Lh}}\right). \end{aligned} \tag{S65}$$

By mean value expansion, if $\mathbb{I} = 1$ and L is large enough, $\left|\beta\left(v-\delta_L \mid p_n, n\right) - \beta\left(v \mid p_n, n\right)\right| \lesssim \sqrt{\log\left(L\right)/L}$. Therefore, if $\mathbb{I} = 1$ and L is large enough, we have $\left|\widetilde{\beta}\left(v+\delta_L \mid n\right) - \beta\left(v \mid p_n, n\right)\right| \lesssim \sqrt{\log\left(L\right)/\left(Lh\right)}$ and similarly, $\left|\widetilde{\beta}\left(v-\delta_L \mid n\right) - \beta\left(v \mid p_n, n\right)\right| \lesssim \sqrt{\log\left(L\right)/\left(Lh\right)}$. Therefore, when L is sufficiently large, the first term on the right-hand side of (S64) can be bounded by

$$\begin{aligned} &\mathbb{I}\left\{\mathbb{G}\left(\widetilde{\beta}\left(v+\delta_L \mid n\right), n\right) - \mathbb{G}\left(\widetilde{\beta}\left(v-\delta_L \mid n\right), n\right)\right\} \\ &\qquad \leq \sup_{(t_1,t_2)\in[-\epsilon_L,\epsilon_L]^2} \left|\mathbb{G}\left(\beta\left(v \mid p_n, n\right)+t_1, n\right) - \mathbb{G}\left(\beta\left(v \mid p_n, n\right)+t_2, n\right)\right|, \end{aligned} \tag{S66}$$

for some $\epsilon_L = O\left(\sqrt{\log\left(L\right)/\left(Lh\right)}\right)$. By calculation and the mean value theorem,

$$\begin{aligned} &\sup_{(t_1,t_2)\in[-\epsilon_L,\epsilon_L]^2} \mathrm{E}\left[\left(\mathcal{G}\left(\boldsymbol{B}_l;\beta\left(v \mid p_n, n\right)+t_1\right) - \mathcal{G}\left(\boldsymbol{B}_l;\beta\left(v \mid p_n, n\right)+t_2\right)\right)^2\right] \\ \leq &\sup_{(t_1,t_2)\in[0,\epsilon_L]^2} \mathrm{E}\left[\left(\mathcal{G}\left(\boldsymbol{B}_l;\beta\left(v \mid p_n, n\right)+t_1\right) - \mathcal{G}\left(\boldsymbol{B}_l;\beta\left(v \mid p_n, n\right)-t_2\right)\right)^2\right] \\ \leq &\sup_{(t_1,t_2)\in[0,\epsilon_L]^2} G\left(\beta\left(v \mid p_n, n\right)+t_1, n\right) - G\left(\beta\left(v \mid p_n, n\right)-t_2, n\right) \qquad \text{(S67)} \\ = &\; O\left(\sqrt{\frac{\log\left(L\right)}{Lh}}\right). \qquad \text{(S68)} \end{aligned}$$

Let $\mathfrak{G} \coloneqq \left\{\mathcal{G}\left(\cdot;\beta\left(v \mid p_n, n\right)+t\right) : t \in \left[-\epsilon_L, \epsilon_L\right]\right\}$. By Nolan and Pollard (1987, Corollary 17), the function class $\left\{f-g : f, g \in \mathfrak{G}\right\}$ is VC-type with respect to a constant envelope. By Chernozhukov, Chetverikov, and Kato (2014, Corollary 5.1) with $\mathcal{F}$ taken to be $\{f-g : f, g \in \mathfrak{G}\}$, F taken to be a constant envelope, and σ^2 taken to be the term on the left hand side of the first inequality in (S68), we have

$$\begin{aligned} &\mathrm{E}\left[\sup_{(t_1,t_2)\in[-\epsilon_L,\epsilon_L]^2} \left|\mathbb{G}\left(\beta\left(v \mid p_n, n\right)+t_1, n\right)-\mathbb{G}\left(\beta\left(v \mid p_n, n\right)+t_2, n\right)\right|\right] \\ &= O\left(L^{-1/2}\left(\frac{\log(L)}{Lh}\right)^{1/4}\right). \end{aligned} \tag{S69}$$

By the mean value and inverse function theorems, if $\mathbb{I}=1$ and L is sufficiently large,

$$\begin{aligned} \left|\widetilde{\beta}\left(v+\delta_L \mid n\right)-\widetilde{\beta}\left(v-\delta_L \mid n\right)\right| &\leq \frac{2\delta_L}{\inf_{b\in\mathcal{B}_n}\xi'\left(b \mid p_n, n\right)-\left\|\mathbb{X}'\left(\cdot \mid n\right)\right\|_{\mathcal{B}_n}} \\ &= O\left(\sqrt{\frac{\log(L)}{L}}\right). \end{aligned}$$

By the above result and Lemma S4(b),

$$\mathbb{I}\left|G\left(\widetilde{\beta}\left(v+\delta_L \mid n\right), n\right)-G\left(\widetilde{\beta}\left(v-\delta_L \mid n\right), n\right)\right| \lesssim \sqrt{\frac{\log(L)}{L}}, \tag{S70}$$

if L sufficiently large. Now by this result, (S64), (S66) and (S69),

$$\mathbb{I}\left|\widehat{F}^*\left(v, n\right)-\widetilde{F}^*\left(v, n\right)\right| = O_p\left(\sqrt{\frac{\log(L)}{L}}\right).$$

It follows from this result, (S55), (S58) and $\Pr[\mathbb{I}=0]=O\left(L^{-1}\right)$ that

$$\widehat{F}^*\left(v \mid p_n\right)-\widetilde{F}^*\left(v \mid p_n\right) = O_p\left(\sqrt{\frac{\log(L)}{L}}\right).$$

By this result and (S57), we have

$$\widehat{F}^*\left(v \mid p_n\right)-F^*\left(v \mid p_n\right) = \left\{\widetilde{F}^*\left(v \mid p_n\right)-F^*\left(v \mid p_n\right)\right\}+O_p\left(\sqrt{\frac{\log(L)}{L}}\right).$$

By using (S63), write

$$\mathbb{I}\left(\widetilde{F}^*\left(v \mid p_n\right)-F^*\left(v \mid p_n\right)\right) = \mathbb{I}\cdot\frac{\mathbb{G}\left(\widetilde{\beta}\left(v \mid n\right), n\right)}{r_n}+\mathbb{I}\left\{G\left(\widetilde{\beta}\left(v \mid n\right) \mid n\right)-G\left(\beta\left(v \mid p_n, n\right) \mid n\right)\right\},$$

where the last equality holds when L is sufficiently large. The first term on the right-hand side can be bounded by $\left\|\mathbb{G}\left(\cdot, n\right)\right\|_{\mathcal{B}_n} = O_p\left(\sqrt{\log(L)/L}\right)$. By similar arguments in the proof of (S65),

$\left|\widetilde{\beta}(v \mid n)-\beta(v \mid p_n, n)\right| \lesssim \sqrt{\log(L)/(Lh)}$ if $\mathbb{I}=1$ and L is sufficiently large. By using this result and the mean value theorem, we have

$$\begin{aligned}\mathbb{I}&\left\{G\left(\widetilde{\beta}(v \mid n) \mid n\right)-G(\beta(v \mid p_n, n) \mid n)\right\} \\ &= \mathbb{I}\cdot g(\beta(v \mid p_n, n) \mid n)\left(\widetilde{\beta}(v \mid n)-\beta(v \mid p_n, n)\right)+O\left(\frac{\log(L)}{Lh}\right).\end{aligned}$$

By Dette, Neumeyer, and Pilz (2006, Lemma A.1), if $\mathbb{I}=1$ and L is sufficiently large, there exists $\lambda\in(0,1)$ such that

$$\begin{aligned}\widetilde{\beta}(v \mid n)-\beta(v \mid p_n, n) &= -\frac{\mathbb{X}(\beta(v \mid p_n, n) \mid n)}{\xi'(\beta(v \mid p_n, n) \mid p_n, n)}-\frac{2\mathbb{X}\left(\tilde{b}_\lambda \mid n\right)\mathbb{X}'\left(\tilde{b}_\lambda \mid n\right)}{\left(\xi'\left(\tilde{b}_\lambda \mid p_n, n\right)+\lambda\mathbb{X}'\left(\tilde{b}_\lambda \mid n\right)\right)^2} \\ &\quad+\frac{\mathbb{X}^2\left(\tilde{b}_\lambda \mid n\right)\left(\xi''\left(\tilde{b}_\lambda \mid p_n, n\right)+\lambda\mathbb{X}''\left(\tilde{b}_\lambda \mid n\right)\right)}{\left(\xi'\left(\tilde{b}_\lambda \mid p_n, n\right)+\lambda\mathbb{X}'\left(\tilde{b}_\lambda \mid n\right)\right)^3},\end{aligned}$$

where $\tilde{b}_\lambda := (\xi(\cdot \mid p_n, n)+\lambda\mathbb{X}(\cdot \mid n))^{-1}(v)$. It now follows from the above result, (S59), (S60), (S61) and Lemmas S4 and S5 that

$$\widehat{F}^*(v \mid p_n)-F^*(v \mid p_n)=-\frac{g(\beta(v \mid p_n, n) \mid n)}{\xi'(\beta(v \mid p_n, n) \mid p_n, n)}\mathbb{X}(\beta(v \mid p_n, n) \mid n)+O_p\left(\frac{\log(L)}{Lh}\right). \tag{S71}$$

By (S59),

$$\widehat{g}^{-1}(b \mid n)-g^{-1}(b \mid n)=-\frac{\mathbb{H}(b \mid n)}{g^2(b \mid n)}-\frac{\mathbb{H}(b \mid n)}{g^2(b \mid n)}\left(\frac{g(b \mid n)}{\widehat{g}(b \mid n)}-1\right),$$

Lemma S4(b) and Lemma S5(c), we have

$$\mathbb{X}(\beta(v \mid p_n, n) \mid n)=\frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))}{(n-1)g^2(\beta(v \mid p_n, n) \mid n)}\mathbb{H}(\beta(v \mid p_n, n) \mid n)+O_p\left(\frac{\log(L)}{Lh}\right).$$

Then, by this result and (S71),

$$\begin{aligned}\widehat{F}^*(v \mid p_n)-F^*(v \mid p_n) &= -\frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))}{(n-1)\xi'(\beta(v \mid p_n, n) \mid p_n, n)g(\beta(v \mid p_n, n) \mid n)}\mathbb{H}(\beta(v \mid p_n, n) \mid n) \\ &\quad+O_p\left(\frac{\log(L)}{Lh}\right).\end{aligned} \tag{S72}$$

For $b\in\left[\underline{\widehat{b}}_n+h, \widehat{\overline{b}}_n-h\right]$, we have

$$\widehat{g}(b, n)=\frac{1}{L}\sum_{l:N_l=n}\sum_{i=1}^{N_l^*}\frac{1}{h}K\left(\frac{B_{il}-b}{h}\right).$$

By Taylor expansion,

$$\begin{aligned}\mathrm{E}\left[\widehat{g}\left(\beta\left(v \mid p_n, n\right), n\right)\right] =& g\left(\beta\left(v \mid p_n, n\right), n\right) \\ &+ \frac{1}{2} g''\left(\beta\left(v \mid p_n, n\right) \mid n\right) r_n \left(\int u^2 K\left(u\right) du\right) h^2 + o\left(h^2\right). \end{aligned} \tag{S73}$$

Let

$$\begin{aligned} J_l\left(v, n\right) \coloneqq& \mathbb{1}\left(N_l = n\right) \sum_{i=1}^{N_l^*} \frac{1}{\sqrt{h}} K\left(\frac{B_{il} - \beta\left(v \mid p_n, n\right)}{h}\right) \\ &- \mathrm{E}\left[\mathbb{1}\left(N_l = n\right) \sum_{i=1}^{N_l^*} \frac{1}{\sqrt{h}} K\left(\frac{B_{il} - \beta\left(v \mid p_n, n\right)}{h}\right)\right], \end{aligned}$$

and $\sigma_g^2\left(v, n\right) \coloneqq \mathrm{E}\left[\left(\sum_{l=1}^{L}\left(J_l\left(v, n\right)/\sqrt{L}\right)\right)^2\right] = \mathrm{E}\left[J_l^2\left(v, n\right)\right]$. Then we can write

$$\sqrt{Lh}\left(\widehat{g}\left(\beta\left(v \mid p_n, n\right), n\right) - \mathrm{E}\left[\widehat{g}\left(\beta\left(v \mid p_n, n\right), n\right)\right]\right) = \sum_{l=1}^{L} \frac{J_l\left(v, n\right)}{\sqrt{L}}.$$

By LIE and Taylor expansion, we have

$$\sigma_g^2\left(v, n\right) = g\left(\beta\left(v \mid p_n, n\right), n\right) \int K^2\left(u\right) du + o\left(1\right) \tag{S74}$$

and

$$\sum_{l=1}^{L} \mathrm{E}\left[\left|\frac{J_l\left(v, n\right)}{\sigma_g\left(v, n\right)}\right|^3\right] = O\left(\left(Lh\right)^{-1/2}\right),$$

which shows that Lyapunov's condition holds. By Lyapunov's central limit theorem (Severini, 2005, Theorem 12.2), $\sum_{l=1}^{L} J_l\left(v, n\right)/\sigma_g\left(v, n\right) \to_d \mathrm{N}\left(0, 1\right)$. By this result, (S74) and Slutsky's theorem,

$$\begin{aligned} &\sqrt{Lh}\Big(\widehat{g}\big(\beta(v \mid p_n, n), n\big) - \mathrm{E}\big[\widehat{g}\big(\beta(v \mid p_n, n), n\big)\big]\Big) \\ &\qquad \to_d \mathrm{N}\Big(0,\, g\big(\beta(v \mid p_n, n), n\big) \int K^2(u)\, du\Big). \end{aligned} \tag{S75}$$

By (S55), (S56) and Lemma S5(c),

$$\mathbb{H}\left(\beta\left(v \mid p_n, n\right) \mid n\right) = \frac{\widehat{g}\left(\beta\left(v \mid p_n, n\right), n\right) - g\left(\beta\left(v \mid p_n, n\right), n\right)}{r_n} + O_p\left(\frac{\log\left(L\right)}{L}\right). \tag{S76}$$

By this result, (S73) and (S75),

$$\sqrt{Lh}\left(\mathbb{H}\left(\beta\left(v \mid p_n, n\right) \mid n\right) - \frac{1}{2} g''\left(\beta\left(v \mid p_n, n\right) \mid n\right)\left(\int u^2 K\left(u\right) du\right) h^2\right)$$

$$\to_d \mathrm{N}\left(0, \frac{g\left(\beta\left(v \mid p_n, n\right) \mid n\right)}{r_n} \int K^2(u)\, du\right).$$

The conclusion in Part (a) follows from this result and (S72).

It follows from straightforward calculations that $\mathrm{E}\left[J_l(v, n) J_l(v', n')\right] = o(1)$ for all $(v, n) \neq (v', n')$. The Lyapunov condition for the multi-dimensional Lyapunov central limit theorem can also be easily verified. Then by these results,

$$\begin{pmatrix} \sqrt{Lh}\left(\widehat{g}\left(\beta\left(v_1 \mid p_{n_1}, n_1\right), n_1\right) - \mathrm{E}\left[\widehat{g}\left(\beta\left(v_1 \mid p_{n_1}, n_1\right), n_1\right)\right]\right) \\ \vdots \\ \sqrt{Lh}\left(\widehat{g}\left(\beta\left(v_J \mid p_{n_M}, n_M\right), n_M\right) - \mathrm{E}\left[\widehat{g}\left(\beta\left(v_J \mid p_{n_M}, n_M\right), n_M\right)\right]\right) \end{pmatrix}$$
$$\to_d \mathrm{N}\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} g\left(\beta\left(v_1 \mid p_{n_1}, n_1\right), n_1\right) & & \\ & \ddots & \\ & & g\left(\beta\left(v_J \mid p_{n_M}, n_M\right), n_M\right) \end{pmatrix} \int K^2(u)\, du\right).$$

The conclusion in Part (b) follows from this result, (S72), (S73) and (S76). ■

Proof of Proposition S4. For notational simplicity, write $\widehat{\boldsymbol{\vartheta}} := \widehat{\boldsymbol{\vartheta}}\left(\widehat{\mathbf{W}}\right)$. Since $Q(\cdot, \cdot; \cdot)$ is continuously differentiable on $[\epsilon, 1-\epsilon]^2 \times \Theta$, under Assumption S3, the uniform convergence

$$\sup_{\boldsymbol{\vartheta} \in \Theta \times \mathbb{F}} \left\| \boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \boldsymbol{\vartheta}\right) - \boldsymbol{\Upsilon}\left(\boldsymbol{F}^*, \boldsymbol{p}_0, \boldsymbol{\vartheta}\right) \right\| \to_p 0 \tag{S77}$$

follows from the consistency of $\widehat{\boldsymbol{F}}^*$ and $\widehat{\boldsymbol{p}}$. Let

$$D_0(\boldsymbol{\vartheta}; \mathbf{W}) := \boldsymbol{\Upsilon}^\top\left(\boldsymbol{F}^*, \boldsymbol{p}_0, \boldsymbol{\vartheta}\right) \mathbf{W} \boldsymbol{\Upsilon}\left(\boldsymbol{F}^*, \boldsymbol{p}_0, \boldsymbol{\vartheta}\right),$$

and it follows that

$$\boldsymbol{\vartheta}_0 = \operatorname*{arg\,min}_{\boldsymbol{\vartheta} \in \Theta \times \mathbb{F}} D_0(\boldsymbol{\vartheta}; \mathbf{W}_0).$$

By the reverse triangle inequality,

$$\left| \sqrt{\widehat{D}(\boldsymbol{\vartheta}; \mathbf{W}_0)} - \sqrt{D_0(\boldsymbol{\vartheta}; \mathbf{W}_0)} \right|$$
$$\leq \left(\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \boldsymbol{\vartheta}\right) - \boldsymbol{\Upsilon}\left(\boldsymbol{F}^*, \boldsymbol{p}_0, \boldsymbol{\vartheta}\right)\right)^\top \mathbf{W}_0 \left(\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*, \widehat{\boldsymbol{p}}, \boldsymbol{\vartheta}\right) - \boldsymbol{\Upsilon}\left(\boldsymbol{F}^*, \boldsymbol{p}_0, \boldsymbol{\vartheta}\right)\right).$$

It follows from this result and (S77) that

$$\sup_{\boldsymbol{\vartheta} \in \Theta \times \mathbb{F}} \left| \sqrt{\widehat{D}(\boldsymbol{\vartheta}; \mathbf{W}_0)} - \sqrt{D_0(\boldsymbol{\vartheta}; \mathbf{W}_0)} \right| \to_p 0.$$

It follows from the reverse triangle inequality and (S77) that

$$\sup_{\boldsymbol{\vartheta}\in\Theta\times\mathbb{F}}\left|\widehat{D}\left(\boldsymbol{\vartheta};\widehat{\mathbf{W}}\right)-\widehat{D}\left(\boldsymbol{\vartheta};\mathbf{W}_0\right)\right|\to_p 0.$$

It follows from these results and the triangle inequality that

$$\sup_{\boldsymbol{\vartheta}\in\Theta\times\mathbb{F}}\left|\widehat{D}\left(\boldsymbol{\vartheta};\widehat{\mathbf{W}}\right)-D_0\left(\boldsymbol{\vartheta};\mathbf{W}_0\right)\right|\to_p 0.$$

Consistency of $\widehat{\boldsymbol{\vartheta}}$ follows from this result and the standard arguments used in the proof of the consistency of M-estimators (see, e.g., Hansen, 2022, Theorem 22.1). Compactness of $\Theta\times\mathbb{F}$ and continuity of $D_0\left(\cdot;\mathbf{W}_0\right)$ ensure that the second requirement in the statement of Hansen (2022, Theorem 22.1) is satisfied.

It follows from consistency of $\widehat{\boldsymbol{\vartheta}}$ and the assumption that $\boldsymbol{\vartheta}_0$ is an interior point that $\widehat{\boldsymbol{\vartheta}}$ satisfies the FOCs wpa1. Then, we have

$$\begin{aligned}
o_p\left((Lh)^{-1/2}\right) &= \frac{\partial}{\partial\boldsymbol{\vartheta}}\boldsymbol{\Upsilon}^\top\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)\widehat{\mathbf{W}}\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)\\
&= \frac{\partial}{\partial\boldsymbol{\vartheta}}\boldsymbol{\Upsilon}^\top\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)\widehat{\mathbf{W}}\left\{\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)-\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)+\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)\right\},
\end{aligned}$$

and therefore,

$$\begin{aligned}
&\frac{\partial}{\partial\boldsymbol{\vartheta}}\boldsymbol{\Upsilon}^\top\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)\widehat{\mathbf{W}}\left(\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)-\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)\right)\\
&\qquad= -\frac{\partial}{\partial\boldsymbol{\vartheta}}\boldsymbol{\Upsilon}^\top\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\widehat{\boldsymbol{\vartheta}}\right)\widehat{\mathbf{W}}\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)+o_p\left((Lh)^{-1/2}\right).
\end{aligned}$$

Therefore, by the mean value theorem,

$$\begin{aligned}
&\left(\frac{\partial}{\partial\boldsymbol{\vartheta}}\boldsymbol{\Upsilon}^\top\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)\widehat{\mathbf{W}}\frac{\partial}{\partial\boldsymbol{\vartheta}^\top}\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\dot{\boldsymbol{\vartheta}}\right)\right)\left(\widehat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta}_0\right)\\
&\qquad= -\frac{\partial}{\partial\boldsymbol{\vartheta}}\boldsymbol{\Upsilon}^\top\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)\widehat{\mathbf{W}}\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)+o_p\left((Lh)^{-1/2}\right), \quad \text{(S78)}
\end{aligned}$$

where $\dot{\boldsymbol{\vartheta}}$ denotes the mean value. By Proposition S3, the fact that $\widehat{\boldsymbol{p}}-\boldsymbol{p}_0=O_p\left(L^{-1/2}\right)$ and the mean value theorem,

$$\begin{aligned}
\boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right) &= \boldsymbol{\Upsilon}\left(\widehat{\boldsymbol{F}}^*,\widehat{\boldsymbol{p}},\boldsymbol{\vartheta}_0\right)-\boldsymbol{\Upsilon}\left(\boldsymbol{F}^*,\boldsymbol{p}_0,\boldsymbol{\vartheta}_0\right)\\
&= \boldsymbol{\Psi}_1\left(\widehat{\boldsymbol{F}}^*-\boldsymbol{F}^*\right)+o_p\left((Lh)^{-1/2}\right).
\end{aligned}$$

Then by this result, (S78), $\widehat{\mathbf{W}}\to_p\mathbf{W}_0$, $\widehat{\boldsymbol{F}}^*\to_p\boldsymbol{F}^*$ and $\widehat{\boldsymbol{p}}\to_p\boldsymbol{p}_0$, we have

$$\widehat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta}_0=-\left(\boldsymbol{\Pi}_0^\top\mathbf{W}_0\boldsymbol{\Pi}_0\right)^{-1}\boldsymbol{\Pi}_0^\top\mathbf{W}_0\boldsymbol{\Psi}_1\left(\widehat{\boldsymbol{F}}^*-\boldsymbol{F}^*\right)+o_p\left((Lh)^{-1/2}\right). \quad \text{(S79)}$$

The second conclusion follows from this result and Proposition S3. ■

Denote $\mathbb{G}_{\dagger}(\cdot \mid n) \coloneqq \widehat{G}_{\dagger}(\cdot \mid n) - G(\cdot \mid n)$, $\mathbb{G}_{\dagger}(\cdot, n) \coloneqq \widehat{G}_{\dagger}(\cdot, n) - G(\cdot, n)$, $\mathbb{H}_{\dagger}(\cdot \mid n) \coloneqq \widehat{g}_{\dagger}(\cdot \mid n) - g(\cdot \mid n)$ and $\mathbb{H}_{\dagger}(\cdot, n) \coloneqq \widehat{g}_{\dagger}(\cdot, n) - g(\cdot, n)$. Let $\widehat{g}'_{\dagger}(\cdot \mid n)$, $\widehat{g}''_{\dagger}(\cdot \mid n)$, $\widehat{g}'_{\dagger}(\cdot, n)$ and $\widehat{g}''_{\dagger}(\cdot, n)$ be the first and second derivatives of $\widehat{g}_{\dagger}(\cdot \mid n)$ and $\widehat{g}_{\dagger}(\cdot, n)$. Let $\mathbb{H}'_{\dagger}(\cdot \mid n)$ and $\mathbb{H}''_{\dagger}(\cdot \mid n)$ denote the first and second derivatives of $\mathbb{H}_{\dagger}(\cdot \mid n)$. The following lemma is a bootstrap analogue of Lemma S5.

Lemma S6. *Suppose that the assumptions in the statement of Proposition S3 are satisfied. Then we have the following results. (a)*

$$\Pr_{\dagger}\left[\left|\widehat{p}_n^{\dagger} - p_n\right| \geq \alpha_{\dagger,L}^{p}\right] = O_p\left(L^{-1}\right),$$

for some deterministic sequence $\alpha_{\dagger,L}^{p} = O\left(\sqrt{\log(L)/L}\right)$. *(b)*

$$\Pr_{\dagger}\left[\left\|\mathbb{G}_{\dagger}(\cdot \mid n)\right\|_{\mathcal{B}_n} \geq \bar{\alpha}_{\dagger,L}\right] = O_p\left(L^{-1}\right),$$

for some deterministic sequence $\bar{\alpha}_{\dagger,L} = O\left(\sqrt{\log(L)/L}\right)$. *(c)*

$$\Pr_{\dagger}\left[\left\|\mathbb{H}_{\dagger}(\cdot \mid n)\right\|_{\mathcal{B}_n} \geq \alpha_{\dagger,L}\right] = O_p\left(L^{-1}\right),$$

for some deterministic sequence $\alpha_{\dagger,L} = O\left(\sqrt{\log(L)/(Lh)}\right)$ *and similar results with deterministic sequences* $\alpha'_{\dagger,L} = O\left(\sqrt{\log(L)/(Lh^3)}\right)$ *and* $\alpha''_{\dagger,L} = O\left(\sqrt{\log(L)/(Lh^5)}\right)$ *hold for* $\left\|\mathbb{H}'_{\dagger}(\cdot \mid n)\right\|_{\mathcal{B}_n}$ *and* $\left\|\mathbb{H}''_{\dagger}(\cdot \mid n)\right\|_{\mathcal{B}_n}$.

Proof of Lemma S6. Let

$$\widehat{\pi}_n^{\dagger} \coloneqq \frac{1}{L}\sum_{l=1}^{L} \mathbb{1}\left(N_l^{\dagger} = n\right).$$

Then, we can write $\widehat{q}_n^{\dagger} = \widetilde{r}_n^{\dagger} / \left(n\widehat{\pi}_n^{\dagger}\right)$. By Bernstein's inequality, we have

$$\Pr_{\dagger}\left[\left|\widetilde{r}_n^{\dagger} - \widehat{r}_n\right| \geq \widehat{\sigma}_{r,n}\sqrt{\frac{2 \cdot \log(L)}{L}}\right] = O_p\left(L^{-1}\right). \tag{S80}$$

By Bernstein's inequality and (S54),

$$\Pr\left[\left|\widehat{\sigma}_{r,n}^2 - \sigma_{r,n}^2\right| \geq C\sqrt{\frac{\log(L)}{L}}\right] = O\left(L^{-1}\right)$$

for some $C>0$. Then by this result, (S54), (S80) and the triangle inequality, we have

$$\Pr_{\dagger}\left[\left|\widetilde{r}_n^{\dagger}-r_n\right|\geq 2\sqrt{\sigma_{r,n}^2+C\sqrt{\frac{\log\left(L\right)}{L}}}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right]=O_p\left(L^{-1}\right). \tag{S81}$$

It follows from similar arguments that

$$\Pr_{\dagger}\left[\left|\widehat{\pi}_n^{\dagger}-\pi_n\right|\geq\alpha_{\dagger,L}^{\pi}\right]=O_p\left(L^{-1}\right)$$

for some deterministic sequence $\alpha_{\dagger,L}^{\pi}=O\left(\sqrt{\log\left(L\right)/L}\right)$. By simple calculation,

$$\Pr_{\dagger}\left[\left|\frac{\pi_n}{\widehat{\pi}_n^{\dagger}}-1\right|\geq\frac{\alpha_{\dagger,L}^{\pi}}{\pi_n-\alpha_{\dagger,L}^{\pi}}\right]=O_p\left(L^{-1}\right).$$

It follows from this result, (S81) and

$$\widehat{q}_n^{\dagger}-q_n=\frac{\widetilde{r}_n^{\dagger}-r_n}{n\pi_n}+\frac{\widetilde{r}_n^{\dagger}}{n\pi_n}\left(\frac{\pi_n}{\widehat{\pi}_n^{\dagger}}-1\right)$$

that $\Pr_{\dagger}\left[\left|\widehat{q}_n^{\dagger}-q_n\right|\geq\alpha_{\dagger,L}^{q}\right]=O_p\left(L^{-1}\right)$ for some deterministic sequence $\alpha_{\dagger,L}^{q}=O\left(\sqrt{\log\left(L\right)/L}\right)$. The conclusion in Part (a) follows from this result.

By (S80) and the fact that for all $x\in\mathbb{R}$ and $c>0$, $|x\vee(-c)|\geq c$ if and only if $|x|\geq c$, we have

$$\begin{aligned}\Pr_{\dagger}\left[\left|\widehat{r}_n^{\dagger}-\widehat{r}_n\right|\geq\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right] &= \Pr_{\dagger}\left[\left|\widetilde{r}_n^{\dagger}-\widehat{r}_n\right|\geq\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right]\\ &= O_p\left(L^{-1}\right).\end{aligned} \tag{S82}$$

Then it follows from the same arguments as those used to prove (S81) that

$$\Pr_{\dagger}\left[\left|\widehat{r}_n^{\dagger}-r_n\right|\geq\alpha_{\dagger,L}^{r}\right]=O_p\left(L^{-1}\right),$$

for some deterministic sequence $\alpha_{\dagger,L}^{r}=O\left(\sqrt{\log\left(L\right)/L}\right)$, and by simple calculation,

$$\Pr_{\dagger}\left[\left|\frac{r_n}{\widehat{r}_n^{\dagger}}-1\right|\geq\frac{\alpha_{\dagger,L}^{r}}{r_n-\alpha_{\dagger,L}^{r}}\right]=O_p\left(L^{-1}\right). \tag{S83}$$

Then we apply Giné and Guillou (2002, Corollary 2.2) with $U=2n$, $\sigma^2=n^2\widehat{\pi}_n$ and t taken to be $C\sqrt{\log\left(L\right)L}$ for some positive constant C. Let $\alpha_L^{\pi}\coloneqq\sigma_{\pi,n}\sqrt{2\left(\log\left(L\right)/L\right)}$. Then, by Giné and

Guillou (2002, Corollary 2.2), taking C to be sufficiently large, we have

$$\Pr_{\dagger}\left[\left\|\widehat{G}_{\dagger}\left(\cdot,n\right)-\widehat{G}\left(\cdot,n\right)\right\|_{\mathcal{B}_n}>C\sqrt{\frac{\log\left(L\right)}{L}}\right]\mathbb{1}\left(\left|\widehat{\pi}_n-\pi_n\right|<\alpha_L^{\pi}\right)=O_p\left(L^{-1}\right).$$

Note that by (S52), we have $\mathbb{1}\left(\left|\widehat{\pi}_n-\pi_n\right|<\alpha_L^{\pi}\right)=1$ wpa1. Therefore,

$$\Pr_{\dagger}\left[\left\|\widehat{G}_{\dagger}\left(\cdot,n\right)-\widehat{G}\left(\cdot,n\right)\right\|_{\mathcal{B}_n}>C\sqrt{\frac{\log\left(L\right)}{L}}\right]=O_p\left(L^{-1}\right),$$

if C is sufficiently large. The conclusion in Part (a) follows from this result, (S83) and

$$\mathbb{G}_{\dagger}\left(b\mid n\right)=\frac{\mathbb{G}_{\dagger}\left(b,n\right)}{r_n}+\frac{\widehat{G}_{\dagger}\left(b,n\right)}{r_n}\left(\frac{r_n}{\widehat{r}_n^{\dagger}}-1\right).$$

The other conclusions follow from using (S83) and Corollary A.2. ■

Proof of Proposition S5. Denote

$$\widetilde{\xi}_{\dagger}\left(b\mid n\right):=b-\frac{\eta_n\left(p_n,G\left(b\mid n\right)\right)}{\left(n-1\right)\widehat{g}_{\dagger}\left(b\mid n\right)}.$$

Let $\widetilde{V}_{il}^{\dagger}:=\widetilde{\xi}_{\dagger}\left(B_{il}^{\dagger}\mid N_l^{\dagger}\right)$. Let $\widetilde{F}_{\dagger}^{*}\left(v,n\right)$ be defined by the right-hand side of (S28) with $\widehat{V}_{il}^{\dagger}$ replaced by $\widetilde{V}_{il}^{\dagger}$ and let $\widetilde{F}_{\dagger}^{*}\left(v\mid p_n\right):=\widetilde{F}_{\dagger}^{*}\left(v,n\right)/r_n$. Let $\mathbb{X}_{\dagger}\left(b\mid n\right):=\widetilde{\xi}_{\dagger}\left(b\mid n\right)-\xi\left(b\mid p_n,n\right)$. Let $\mathbb{X}_{\dagger}'\left(\cdot\mid n\right)$ and $\mathbb{X}_{\dagger}''\left(\cdot\mid n\right)$ denote the first and second derivatives of $\mathbb{X}_{\dagger}\left(\cdot\mid n\right)$. Then, we have

$$\mathbb{X}_{\dagger}\left(b\mid n\right)=-\frac{\eta_n\left(p_n,G\left(b\mid n\right)\right)}{n-1}\left\{\frac{1}{\widehat{g}_{\dagger}\left(b\mid n\right)}-\frac{1}{g\left(b\mid n\right)}\right\},\tag{S84}$$

$$\begin{aligned}\mathbb{X}_{\dagger}'\left(b\mid n\right) &= -\frac{D_2\eta_n\left(p_n,G\left(b\mid n\right)\right)g\left(b\mid n\right)}{n-1}\left\{\frac{1}{\widehat{g}_{\dagger}\left(b\mid n\right)}-\frac{1}{g\left(b\mid n\right)}\right\}\\ &\quad+\frac{\eta_n\left(p_n,G\left(b\mid n\right)\right)}{n-1}\left\{\frac{\widehat{g}_{\dagger}'\left(b\mid n\right)}{\widehat{g}_{\dagger}^2\left(b\mid n\right)}-\frac{g'\left(b\mid n\right)}{g^2\left(b\mid n\right)}\right\},\end{aligned}\tag{S85}$$

and

$$\begin{aligned}\mathbb{X}_{\dagger}''(b\mid n) &= -\frac{D_2^2\eta_n(p_n,G(b\mid n))g^2(b\mid n)+D_2\eta_n(p_n,G(b\mid n))g'(b\mid n)}{n-1}\left\{\frac{1}{\widehat{g}_{\dagger}(b\mid n)}-\frac{1}{g(b\mid n)}\right\}\\ &\quad+\frac{2D_2\eta_n(p_n,G(b\mid n))g(b\mid n)}{n-1}\left\{\frac{\widehat{g}_{\dagger}'(b\mid n)}{\widehat{g}_{\dagger}^2(b\mid n)}-\frac{g'(b\mid n)}{g^2(b\mid n)}\right\}\\ &\quad+\frac{\eta_n(p_n,G(b\mid n))}{n-1}\left\{\frac{\widehat{g}_{\dagger}''(b\mid n)\widehat{g}_{\dagger}^2(b\mid n)-2(\widehat{g}_{\dagger}'(b\mid n))^2\widehat{g}_{\dagger}(b\mid n)}{\widehat{g}_{\dagger}^4(b\mid n)}\right.\end{aligned}$$

$$-\frac{g''(b\mid n)g^2(b\mid n)-2(g'(b\mid n))^2g(b\mid n)}{g^4(b\mid n)}\bigg\}. \tag{S86}$$

Denote

$$\mathbb{K}_\dagger\left(b\mid n\right)\coloneqq\eta_n\left(\widehat{p}_n^\dagger,\widehat{G}_\dagger\left(b\mid n\right)\right)-\eta_n\left(p_n,G\left(b\mid n\right)\right).$$

By Lemma S6(a,b) and similar arguments as those used to prove (S62),

$$\Pr_\dagger\left[\left\|\mathbb{K}_\dagger\left(\cdot\mid n\right)\right\|_{\mathcal{B}_n}>\tilde{\alpha}_{\dagger,L}\right]=O_p\left(L^{-1}\right),$$

for some deterministic sequence $\tilde{\alpha}_{\dagger,L}=O\left(\sqrt{\log\left(L\right)/L}\right)$. Let $\bar{\mathbb{T}}_\dagger\coloneqq\mathbb{1}\left(\left\|\mathbb{K}_\dagger\left(\cdot\mid n\right)\right\|_{\mathcal{B}_n}<\tilde{\alpha}_{\dagger,L}\right)$. Let $\left(\mathbb{T}_\dagger,\mathbb{T}_\dagger',\mathbb{T}_\dagger''\right)$ be defined by the same formula with $(\mathbb{K}_\dagger,\tilde{\alpha}_{\dagger,L})$ replaced by $(\mathbb{H}_\dagger,\alpha_{\dagger,L})$, $\left(\mathbb{H}_\dagger',\alpha_{\dagger,L}'\right)$ and $\left(\mathbb{H}_\dagger'',\alpha_{\dagger,L}''\right)$. Let $\mathbb{I}_\dagger\coloneqq\bar{\mathbb{T}}_\dagger\mathbb{T}_\dagger\mathbb{T}_\dagger'\mathbb{T}_\dagger''$. Then, we have $\Pr_\dagger\left[\mathbb{I}_\dagger=0\right]=O_p\left(L^{-1}\right)$.

Decompose

$$\widehat{F}_\dagger^*\left(v\mid p_n\right)-\widetilde{F}_\dagger^*\left(v\mid p_n\right)=\frac{\widehat{F}_\dagger^*\left(v,n\right)-\widetilde{F}_\dagger^*\left(v,n\right)}{r_n}+\frac{\widehat{F}_\dagger^*\left(v,n\right)}{r_n}\left(\frac{r_n}{\widehat{r}_n^\dagger}-1\right). \tag{S87}$$

It now follows that

$$\begin{aligned}\mathrm{E}_\dagger\left[\left(\frac{r_n}{\widehat{r}_n^\dagger}-1\right)^2\right]\ \leq\ &\frac{\mathrm{E}_\dagger\left[\left(\left(\widetilde{r}_n^\dagger-r_n\right)\vee\left(\widehat{r}_n-r_n-\widehat{\sigma}_{r,n}\sqrt{2\cdot\log\left(L\right)/L}\right)\right)^2\right]}{\left(\widehat{r}_n-\widehat{\sigma}_{r,n}\sqrt{2\cdot\log\left(L\right)/L}\right)^2}\\&+\mathbb{1}\left(\widehat{r}_n\leq\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right)\mathrm{E}_\dagger\left[\left(\frac{r_n}{\widehat{r}_n^\dagger}-1\right)^2\right],\end{aligned}$$

where the second term on the right-hand side equals zero wpa1. For the first term, we have

$$\begin{aligned}\mathrm{E}_\dagger\left[\left(\left(\widetilde{r}_n^\dagger-r_n\right)\vee\left(\widehat{r}_n-r_n-\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right)\right)^2\right]&\leq\mathrm{E}_\dagger\left[\left(\widetilde{r}_n^\dagger-r_n\right)^2\right]\\&+\Pr_\dagger\left[\left|\widetilde{r}_n^\dagger-\widehat{r}_n\right|>\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right]\left(\widehat{r}_n-r_n-\widehat{\sigma}_{r,n}\sqrt{\frac{2\cdot\log\left(L\right)}{L}}\right)^2.\end{aligned}$$

Note that $\mathrm{E}_\dagger\left[\left(\widetilde{r}_n^\dagger-r_n\right)^2\right]=O_p\left(L^{-1}\right)$. It now follows from this result, (S82) and the above inequalities that $\mathrm{E}_\dagger\left[\left(r_n/\widehat{r}_n^\dagger-1\right)^2\right]=O_p\left(L^{-1}\right)$.

Denote $\delta_L^\dagger \coloneqq \left\|\widehat{\xi}_\dagger(\cdot \mid n) - \widetilde{\xi}_\dagger(\cdot \mid n)\right\|_{\mathcal{B}_n}$. Then,

$$\left|\widehat{F}_\dagger^*(v,n) - \widetilde{F}_\dagger^*(v,n)\right| \leq \frac{1}{L}\sum_{l:N_l^\dagger=n}\sum_{i=1}^{N_l^{*\dagger}}\left\{\mathbb{1}\left(\widetilde{V}_{il}^\dagger \leq v+\delta_L^\dagger\right) - \mathbb{1}\left(\widetilde{V}_{il}^\dagger \leq v-\delta_L^\dagger\right)\right\}.$$

Let $\overline{v}_n^\dagger \coloneqq \widetilde{\xi}_\dagger\left(\overline{b}_n \mid n\right)$ and $\underline{v}_n^\dagger \coloneqq \widetilde{\xi}_\dagger\left(\underline{b}_n \mid n\right)$. If $\mathbb{I}_\dagger = 1$ and L is large enough, $\delta_L^\dagger \lesssim \sqrt{\log(L)/L}$, $\left|\overline{v}_n^\dagger - \overline{v}\right| \lesssim \sqrt{\log(L)/(Lh)}$ and $\left|\underline{v}_n^\dagger - \underline{v}\right| \lesssim \sqrt{\log(L)/(Lh)}$. By (S85), if $\mathbb{I}_\dagger = 1$ and L is large enough, $\left\|\mathbb{X}_\dagger'(\cdot \mid n)\right\|_{\mathcal{B}_n}$ is sufficiently small, the inverse $\widetilde{\beta}_\dagger(\cdot \mid n) \coloneqq \widetilde{\xi}_\dagger^{-1}(\cdot \mid n)$ exists and $\widetilde{\beta}_\dagger(\cdot \mid n)$ is a strictly increasing function on $\left[\underline{v}_n^\dagger, \overline{v}_n^\dagger\right]$ with $\widetilde{\beta}_\dagger\left(\underline{v}_n^\dagger\right) = \underline{b}_n$ and $\widetilde{\beta}_\dagger\left(\overline{v}_n^\dagger \mid n\right) = \overline{b}_n$. Then, when L is sufficiently large,

$$\begin{aligned}\mathbb{I}_\dagger\left|\widehat{F}_\dagger^*(v,n) - \widetilde{F}_\dagger^*(v,n)\right| \quad \leq \quad & \mathbb{I}_\dagger\left\{\mathbb{G}_\dagger\left(\widetilde{\beta}_\dagger\left(v+\delta_L^\dagger \mid n\right),n\right) - \mathbb{G}_\dagger\left(\widetilde{\beta}_\dagger\left(v-\delta_L^\dagger \mid n\right),n\right)\right\} \\ & + \mathbb{I}_\dagger\left\{G\left(\widetilde{\beta}_\dagger\left(v+\delta_L^\dagger \mid n\right),n\right) - G\left(\widetilde{\beta}_\dagger\left(v-\delta_L^\dagger \mid n\right),n\right)\right\}. \quad \text{(S88)}\end{aligned}$$

By arguments similar to those in the proof of (S70),

$$\mathbb{I}_\dagger\left|G\left(\widetilde{\beta}_\dagger\left(v+\delta_L^\dagger \mid n\right),n\right) - G\left(\widetilde{\beta}_\dagger\left(v-\delta_L^\dagger \mid n\right),n\right)\right| \lesssim \sqrt{\frac{\log(L)}{L}},$$

if L sufficiently large. For the first term on the right-hand side of (S88),

$$\begin{aligned}&\mathbb{I}_\dagger\left\{\mathbb{G}_\dagger\left(\widetilde{\beta}_\dagger\left(v+\delta_L^\dagger \mid n\right),n\right) - \mathbb{G}_\dagger\left(\widetilde{\beta}_\dagger\left(v-\delta_L^\dagger \mid n\right),n\right)\right\}^2 \\ &\qquad \leq \left(\sup_{(t_1,t_2)\in[-\epsilon_L,\epsilon_L]^2}\left|\mathbb{G}_\dagger\left(\beta(v \mid p_n,n)+t_1,n\right) - \mathbb{G}_\dagger\left(\beta(v \mid p_n,n)+t_2,n\right)\right|\right)^2,\end{aligned}$$

for some deterministic sequence $\epsilon_L = O\left(\sqrt{\log(L)/(Lh)}\right)$. By arguments similar to those in the proof of (S69),

$$\begin{aligned}&\mathrm{E}_\dagger\left[\sup_{(t_1,t_2)\in[-\epsilon_L,\epsilon_L]^2}\left|\mathbb{G}_\dagger\left(\beta(v \mid p_n,n)+t_1,n\right) - \mathbb{G}_\dagger\left(\beta(v \mid p_n,n)+t_2,n\right)\right|\right] \\ &= O_p\left(L^{-1/2}\left(\frac{\log(L)}{Lh}\right)^{1/4}\right).\end{aligned}$$

By this result and Ledoux and Talagrand (1991, Theorem 6.20),

$$\begin{aligned}&\mathrm{E}_\dagger\left[\left(\sup_{(t_1,t_2)\in[-\epsilon_L,\epsilon_L]^2}\left|\mathbb{G}_\dagger\left(\beta(v \mid p_n,n)+t_1,n\right) - \mathbb{G}_\dagger\left(\beta(v \mid p_n,n)+t_2,n\right)\right|\right)^2\right] \\ &= O_p\left(L^{-1}\left(\frac{\log(L)}{Lh}\right)^{1/2}\right).\end{aligned}$$

It now follows that

$$\begin{aligned}
\mathrm{E}_{\dagger}\left[\left(\widehat{F}_{\dagger}^{*}(v,n)-\widetilde{F}_{\dagger}^{*}(v,n)\right)^{2}\right] &\lesssim \mathrm{E}_{\dagger}\left[\mathbb{I}_{\dagger}\left(\widehat{F}_{\dagger}^{*}(v,n)-\widetilde{F}_{\dagger}^{*}(v,n)\right)^{2}\right]+\mathrm{Pr}_{\dagger}\left[\mathbb{I}_{\dagger}=0\right] \\
&= O_{p}\left(\frac{\log(L)}{L}\right). \qquad \text{(S89)}
\end{aligned}$$

Then it follows from this result, (S87) and $\mathrm{E}_{\dagger}\left[\left(r_{n}/\widehat{r}_{n}^{\dagger}-1\right)^{2}\right]=O_{p}\left(L^{-1}\right)$ that

$$\mathrm{E}_{\dagger}\left[\left(\widehat{F}_{\dagger}^{*}(v\mid p_{n})-\widetilde{F}_{\dagger}^{*}(v\mid p_{n})\right)^{2}\right]=O_{p}\left(\frac{\log(L)}{L}\right). \qquad \text{(S90)}$$

Write

$$\begin{aligned}
\widetilde{F}_{\dagger}^{*}(v\mid p_{n}) &= \mathbb{I}_{\dagger}\cdot\widetilde{F}_{\dagger}^{*}(v\mid p_{n})+(1-\mathbb{I}_{\dagger})\,\widetilde{F}_{\dagger}^{*}(v\mid p_{n}) \\
&= \mathbb{I}_{\dagger}\left(G\left(\widetilde{\beta}_{\dagger}(v\mid n)\mid n\right)-G\left(\beta(v\mid p_{n},n)\mid n\right)\right) \\
&\quad +\mathbb{I}_{\dagger}\cdot G\left(\beta(v\mid p_{n},n)\mid n\right)+\mathbb{I}_{\dagger}\cdot\frac{\mathbb{G}_{\dagger}\left(\widetilde{\beta}_{\dagger}(v\mid n),n\right)}{r_{n}}+(1-\mathbb{I}_{\dagger})\,\widetilde{F}_{\dagger}^{*}(v\mid p_{n}). \qquad \text{(S91)}
\end{aligned}$$

By Dette, Neumeyer, and Pilz (2006, Lemma A.1), if $\mathbb{I}_{\dagger}=1$ and L is large enough, there exists $\lambda\in(0,1)$ such that

$$\begin{aligned}
\widetilde{\beta}_{\dagger}(v\mid n)-\beta(v\mid p_{n},n) &= -\frac{\mathbb{X}_{\dagger}\left(\beta(v\mid p_{n},n)\mid n\right)}{\xi'\left(\beta(v\mid p_{n},n)\mid p_{n},n\right)}-2\frac{\mathbb{X}_{\dagger}(b_{\lambda}\mid n)\,\mathbb{X}_{\dagger}'(b_{\lambda}\mid n)}{\left(\xi'(b_{\lambda}\mid p_{n},n)+\lambda\mathbb{X}_{\dagger}'(b_{\lambda}\mid n)\right)^{2}} \\
&\quad +\frac{\mathbb{X}_{\dagger}^{2}(b_{\lambda}\mid n)\left(\xi''(b_{\lambda}\mid p_{n},n)+\lambda\mathbb{X}_{\dagger}''(b_{\lambda}\mid n)\right)}{\left(\xi'(b_{\lambda}\mid p_{n},n)+\lambda\mathbb{X}_{\dagger}'(b_{\lambda}\mid n)\right)^{3}}, \qquad \text{(S92)}
\end{aligned}$$

where $b_{\lambda}:=\left(\xi(\cdot\mid p_{n},n)+\lambda\mathbb{X}_{\dagger}(\cdot\mid n)\right)^{-1}(v)$. By (S84), (S85), (S86), (S92) and the mean value theorem, we have

$$\begin{aligned}
\mathbb{I}_{\dagger}\left(G\left(\widetilde{\beta}_{\dagger}(v\mid n)\mid n\right)-G\left(\beta(v\mid p_{n},n)\mid n\right)\right) &= \mathbb{I}_{\dagger}\cdot g\left(\beta(v\mid p_{n},n)\mid n\right)\left(\widetilde{\beta}_{\dagger}(v\mid n)-\beta(v\mid p_{n},n)\right) \\
&\quad +O\left(\frac{\log(L)}{Lh}\right)
\end{aligned}$$

and

$$\mathbb{I}_{\dagger}\left(\widetilde{\beta}_{\dagger}(v\mid n)-\beta(v\mid p_{n},n)\right)=-\mathbb{I}_{\dagger}\cdot\frac{\mathbb{X}_{\dagger}\left(\beta(v\mid p_{n},n)\mid n\right)}{\xi'\left(\beta(v\mid p_{n},n)\mid p_{n},n\right)}+O\left(\frac{\log(L)}{Lh}\right).$$

It is easy to check that

$$\mathbb{I}_{\dagger}\cdot\mathbb{X}_{\dagger}\left(\beta(v\mid p_{n},n)\mid n\right)=\mathbb{I}_{\dagger}\cdot\frac{\eta_{n}\left(p_{n},G\left(\beta(v\mid p_{n},n)\mid n\right)\right)\mathbb{H}_{\dagger}\left(\beta(v\mid p_{n},n)\mid n\right)}{(n-1)\,g^{2}\left(\beta(v\mid p_{n},n)\mid n\right)}+O\left(\frac{\log(L)}{Lh}\right).$$

By these results and (S91), we have

$$\begin{aligned}\widetilde{F}_{\dagger}^{*}(v \mid p_n) &= -\mathbb{I}_{\dagger} \cdot \frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))\, \mathbb{H}_{\dagger}(\beta(v \mid p_n, n) \mid n)}{(n-1)\, g(\beta(v \mid p_n, n) \mid n)\, \xi'(\beta(v \mid p_n, n) \mid p_n, n)} \\ &\quad + \mathbb{I}_{\dagger} \cdot G(\beta(v \mid p_n, n) \mid n) + \mathbb{I}_{\dagger} \cdot \frac{\mathbb{G}_{\dagger}\left(\widetilde{\beta}_{\dagger}(v \mid n), n\right)}{r_n} + (1 - \mathbb{I}_{\dagger})\, \widetilde{F}_{\dagger}^{*}(v \mid p_n) \\ &\quad + O\left(\frac{\log(L)}{Lh}\right).\end{aligned}$$

Then, by this result and

$$\widehat{g}_{\dagger}(\beta(v \mid p_n, n) \mid n) = \frac{\widehat{g}_{\dagger}(\beta(v \mid p_n, n), n)}{r_n} + \frac{\widehat{g}_{\dagger}(\beta(v \mid p_n, n), n)}{r_n}\left(\frac{r_n}{\widehat{r}_n^{\dagger}} - 1\right),$$

we can write

$$\begin{aligned}\widetilde{F}_{\dagger}^{*}(v \mid p_n) &= -\frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))\, \widehat{g}_{\dagger}(\beta(v \mid p_n, n), n)}{(n-1)\, \xi'(\beta(v \mid p_n, n) \mid p_n, n)\, g(\beta(v \mid p_n, n), n)} \\ &\quad + V_1^{\dagger} + V_2^{\dagger} + V_3^{\dagger} + V_4^{\dagger} + V_5^{\dagger} + O\left(\frac{\log(L)}{Lh}\right),\end{aligned} \tag{S93}$$

where

$$\begin{aligned}V_1^{\dagger} &:= (1 - \mathbb{I}_{\dagger}) \frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))\, \widehat{g}_{\dagger}(\beta(v \mid p_n, n), n)}{(n-1)\, \xi'(\beta(v \mid p_n, n) \mid p_n, n)\, g(\beta(v \mid p_n, n), n)} \\ V_2^{\dagger} &:= -\mathbb{I}_{\dagger} \cdot \frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))\, \widehat{g}_{\dagger}(\beta(v \mid p_n, n), n)}{(n-1)\, \xi'(\beta(v \mid p_n, n) \mid p_n, n)\, g(\beta(v \mid p_n, n), n)}\left(\frac{r_n}{\widehat{r}_n^{\dagger}} - 1\right) \\ V_3^{\dagger} &:= \mathbb{I}_{\dagger}\left\{G(\beta(v \mid p_n, n) \mid n) + \frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))}{(n-1)\, \xi'(\beta(v \mid p_n, n) \mid p_n, n)}\right\} \\ V_4^{\dagger} &:= \mathbb{I}_{\dagger} \cdot \frac{\mathbb{G}_{\dagger}\left(\widetilde{\beta}_{\dagger}(v \mid n), n\right)}{r_n} \\ V_5^{\dagger} &:= (1 - \mathbb{I}_{\dagger})\, \widetilde{F}_{\dagger}^{*}(v \mid p_n).\end{aligned}$$

It follows from the Cauchy-Schwarz inequality that

$$\begin{aligned}&\mathrm{E}_{\dagger}\left[(1 - \mathbb{I}_{\dagger})^2 \widehat{g}_{\dagger}^2(\beta(v \mid p_n, n), n)\right] \\ &\quad \lesssim \sqrt{\Pr_{\dagger}[\mathbb{I}_{\dagger} = 0] \cdot \mathrm{E}_{\dagger}\left[(\widehat{g}_{\dagger}(\beta(v \mid p_n, n), n) - \widehat{g}(\beta(v \mid p_n, n), n))^4\right]} \\ &\qquad + \Pr_{\dagger}[\mathbb{I}_{\dagger} = 0]\, \widehat{g}^2(\beta(v \mid p_n, n), n).\end{aligned} \tag{S94}$$

By the Rosenthal inequality,

$$\mathrm{E}_{\dagger}\left[(\widehat{g}_{\dagger}(\beta(v \mid p_n, n), n) - \widehat{g}(\beta(v \mid p_n, n), n))^4\right] = O_p\left((Lh)^{-2}\right).$$

By this result, $\Pr_{\dagger}[\mathbb{I}_{\dagger} = 0] = O_p\left(L^{-1}\right)$ and (S94), we have $\mathrm{Var}_{\dagger}\left[V_1^{\dagger}\right] = O_p\left(L^{-1}\right)$. It follows from similar arguments and $\mathrm{E}_\dagger\left[\left(r_n/\widehat{r}_n^\dagger - 1\right)^2\right] = O_p\left(L^{-1}\right)$ that $\mathrm{Var}_\dagger\left[V_2^\dagger\right] = O_p\left(L^{-1}\right)$. It follows from

$$\mathrm{Var}_\dagger[\mathbb{I}_\dagger] = \Pr_\dagger[\mathbb{I}_\dagger = 0] \cdot \Pr_\dagger[\mathbb{I}_\dagger = 1] = O_p(L^{-1})$$

that $\mathrm{Var}_\dagger[V_3^\dagger] = O_p(L^{-1})$. By Chernozhukov, Chetverikov, and Kato (2014, Corollary 5.1), we have $\mathrm{E}[\|\mathbb{G}(\cdot, n)\|_{\mathcal{B}_n}] = O(L^{-1/2})$ and $\mathrm{E}_\dagger[\|\widehat{G}_\dagger(\cdot, n) - \widehat{G}(\cdot, n)\|_{\mathcal{B}_n}] = O_p(L^{-1/2})$. By these results and Ledoux and Talagrand (1991, Theorem 6.20), we have

$$\mathrm{Var}_\dagger\left[\mathrm{V}_4^\dagger\right] \lesssim \mathrm{E}_\dagger\left[\left(\|\mathbb{G}_\dagger(\cdot, n)\|_{\mathcal{B}_n}\right)^2\right] = O_p\left(L^{-1}\right).$$

Since $0 \leq \widetilde{F}_\dagger^*(v \mid p_n) \leq n/r_n$,

$$\mathrm{Var}_\dagger\left[V_5^\dagger\right] \leq \mathrm{E}\left[(1 - \mathbb{I}_\dagger)^2\left(\widetilde{F}_\dagger^*(v \mid p_n)\right)^2\right] \lesssim \Pr_\dagger[\mathbb{I}_\dagger = 0] = O_p\left(L^{-1}\right).$$

It follows from the above results, (S90), (S93) and the Cauchy-Schwarz inequality that

$$\mathrm{Var}_\dagger\left[\widehat{F}_\dagger^*(v \mid p_n)\right] = \mathrm{Var}_\dagger\left[\frac{\eta_n(p_n, G(\beta(v \mid p_n, n) \mid n))\,\widehat{g}_\dagger(\beta(v \mid p_n, n), n)}{(n-1)\,\xi'(\beta(v \mid p_n, n) \mid p_n, n)\,g(\beta(v \mid p_n, n), n)}\right] + o_p\left((Lh)^{-1}\right). \quad \text{(S95)}$$

By simple calculation, we have

$$\begin{aligned}
&\mathrm{Var}_\dagger\left[\widehat{g}_\dagger(\beta(v \mid p_n, n), n)\right] \\
&\quad = \frac{1}{L}\left\{\frac{1}{L}\sum_{l:N_l=n}\left(\sum_{i=1}^{N_l^*}\frac{1}{h}\mathcal{K}_1\left(B_{il}, \beta(v \mid p_n, n) \mid h, \widehat{\underline{b}}_n, \widehat{\overline{b}}_n\right)\right)^2 - \widehat{g}^2(\beta(v \mid p_n, n), n)\right\}.
\end{aligned}$$

The conclusion follows from this result, (S95), $\widehat{g}(\beta(v \mid p_n, n), n) \to_p g(\beta(v \mid p_n, n), n)$ and the fact that

$$\frac{1}{Lh}\sum_{l:N_l=n}\left(\sum_{i=1}^{N_l^*}\mathcal{K}_1\left(B_{il}, \beta(v \mid p_n, n) \mid h, \widehat{\underline{b}}_n, \widehat{\overline{b}}_n\right)\right)^2 \to_p g(\beta(v \mid p_n, n), n)\int K^2(u)\,du.$$

■

References

Bickel, P. J., and K. A. Doksum (2015): *Mathematical Statistics: Basic Ideas and Selected Topics, Volumes I-II Package*. Chapman and Hall/CRC.

Chen, S. X., and T. Huang (2007): "Nonparametric Estimation of Copula Functions for Dependence Modelling," *Canadian Journal of Statistics*, 35(2), 265–282.

Chen, X., M. L. Gentry, T. Li, and J. Lu (2026): "Identification and Inference in First-Price Auctions with Risk Averse Bidders and Selective Entry," *Review of Economic Studies*, 93(1), 366–403.

Cheng, M.-Y., J. Fan, and J. S. Marron (1997): "On Automatic Boundary Corrections," *Annals of Statistics*, 25(4), 1691–1708.

Chernozhukov, V., D. Chetverikov, and K. Kato (2014): "Gaussian approximation of suprema of empirical processes," *Annals of Statistics*, 42(4), 1564–1597.

Dette, H., N. Neumeyer, and K. F. Pilz (2006): "A Simple Nonparametric Estimator of a Strictly Monotone Regression Function," *Bernoulli*, 12(1), 1–23.

Gentry, M., and T. Li (2014): "Identification in Auctions with Selective Entry," *Econometrica*, 82(1), 315–344.

Giné, E., and A. Guillou (1999): "Laws of the Iterated Logarithm for Censored Data," *Annals of Probability*, 27(4), 2042–2067.

——— (2002): "Rates of Strong Uniform Consistency for Multivariate Kernel Density Estimators," *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 38(6), 907–921.

Giné, E., and R. Nickl (2015): *Mathematical Foundations of Infinite-Dimensional Statistical Models*, vol. 40. Cambridge University Press.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

Hahn, J., and Z. Liao (2021): "Bootstrap Standard Error Estimates and Inference," *Econometrica*, 89(4), 1963–1977.

Hansen, B. (2022): *Econometrics*. Princeton University Press.

Hickman, B. R., and T. P. Hubbard (2015): "Replacing Sample Trimming with Boundary Correction in Nonparametric Estimation of First-Price Auctions," *Journal of Applied Econometrics*, 30(5), 739–762.

Jones, M. C. (1993): "Simple Boundary Correction for Kernel Density Estimation," *Statistics and Computing*, 3(3), 135–146.

Kosorok, M. R. (2008): *Introduction to Empirical Processes and Semiparametric Inference*, vol. 61. Springer.

Ledoux, M., and M. Talagrand (1991): *Probability in Banach Spaces: Isoperimetry and Processes*. Springer Science & Business Media.

Lejeune, M., and P. Sarda (1992): "Smooth Estimators of Distribution and Density Functions," *Computational Statistics & Data Analysis*, 14(4), 457–471.

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

Lewbel, A. (2007): "A Local Generalized Method of Moments Estimator," *Economics Letters*, 94(1), 124–128.

——— (2019): "The Identification Zoo: Meanings of Identification in Econometrics," *Journal of Economic Literature*, 57(4), 835–903.

Li, T., and X. Zheng (2009): "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions," *Review of Economic Studies*, 76(4), 1397–1429.

Ma, J., V. Marmer, and A. Shneyerov (2019): "Inference for First-Price Auctions with Guerre, Perrigne, and Vuong's Estimator," *Journal of Econometrics*, 211(2), 507–538.

Ma, J., V. Marmer, A. Shneyerov, and P. Xu (2021): "Monotonicity-Constrained Nonparametric Estimation and Inference for First-Price Auctions," *Econometric Reviews*, 40(10), 944–982.

Marmer, V., and A. Shneyerov (2012): "Quantile-Based Nonparametric Inference for First-Price Auctions," *Journal of Econometrics*, 167(2), 345–357.

Marmer, V., A. Shneyerov, and P. Xu (2013): "What Model for Entry in First-Price Auctions? A Nonparametric Approach," *Journal of Econometrics*, 176(1), 46–58.

Nelsen, R. B. (2006): *An Introduction to Copulas*. Springer.

Nolan, D., and D. Pollard (1987): "U-Processes: Rates of Convergence," *Annals of Statistics*, 15(2), 780–799.

Samuelson, W. F. (1985): "Competitive Bidding with Entry Costs," *Economics Letters*, 17(1-2), 53–57.

Severini, T. A. (2005): *Elements of Distribution Theory*, no. 17. Cambridge University Press.

Zincenko, F. (2024): "Estimation and Inference of Seller's Expected Revenue in First-Price Auctions," *Journal of Econometrics*, 241(1), 105734.