Supplement to "What Model for Entry in First-Price Auctions? A Nonparametric Approach"

Vadim Marmer
University of British Columbia

Artyom Shneyerov
CIRANO, CIREQ, and Concordia University

Pai Xu
University of Hong Kong

March 18, 2012

S.1 Outline

We use MSX to abbreviate the main paper.

This supplement contains the following materials: smoothness results for the distributions of values and bids; proofs of consistency and asymptotic normality of the nonparametric estimators used in MSX; and a proof of the validity of the bootstrap critical values used in MSX. We also describe here how single-index models can be incorporated in MSX's approach to circumvent the curse of dimensionality when there are many covariates.

S.2 Smoothness Results for the distributions of values and bids

Lemma S.1 *The following results hold for the SEM and the S model.*

(a) *Let $f^*(v|N,x)$ denote the conditional PDF of valuations given $N_l = N, x_l = x$ and conditional on bidding. Then $f^*(\cdot|N,x)$ is strictly positive and bounded away from zero on its support $[\underline{v}(x), \overline{v}(x)]$, admits up to $R-1$ continuous* *derivatives on $[\underline{v}(x), \overline{v}(x)]$ for all $x \in \mathcal{X}$, $N \in \mathcal{N}$, and $f^*(v|N, \cdot)$ admits up to R continuous partial derivatives on the interior of $\mathcal{X}$ for all $v \in [\underline{v}(N,x), \overline{v}(N,x)]$, $N \in \mathcal{N}$.*

(b) *The conditional probability of entry $p(N,x)$ admits up to R continuous partial derivatives with respect to x on the interior of $\mathcal{X}$ for all $N \in \mathcal{N}$.*

Proof of Lemma S.1. Consider the SEM first. Note that $\bar{s}(N,x)$ is determined by

$$\int_{r(x)}^{\bar{v}(x)} (1 - F(v|s,x)) \lambda(v, \bar{s}, x)^{N-1} dv - k(x) = 0, \text{ where} \tag{S.1}$$

$$\lambda(v, \bar{s}, x) = F(\bar{s}|x) + \int_{\bar{s}}^{\bar{v}(x)} F(v|s,x) f(s|x) ds.$$

By Lemma A1(i) in GPV, $\bar{v}(x)$ admits up to R continuous partial derivatives on the interior of $\mathcal{X}$. Together with our Assumptions 8(f)-(h), this implies that the left-hand side in (S.1) is smooth up to order R in x. Moreover, its partial derivative with respect to $\bar{s}$ is

$$\int_{r(x)}^{\bar{v}(x)} (1 - F(v|x)) (f(\bar{s}|x) - F(v|\bar{s},x) f(\bar{s}|x))^{N-1} dv > 0.$$

The Implicit Function Theorem then implies that $\bar{s}(N,x)$ admits up to R continuous partial derivatives with respect to x on the interior of $\mathcal{X}$ for all $N \in \mathcal{N}$. The result follows then from the conditions of the lemma, and the definitions of $f^*(v|N,x)$ and $p(N,x)$:

$$\begin{aligned} f^*(v|N,x) &= \frac{\int_{s \geq \bar{s}(N,x)}^{\bar{v}(x)} f(v,s|x) ds}{\int_{v \geq r(x)}^{\bar{v}(x)} \int_{s \geq \bar{s}(N,x)}^{\bar{v}(x)} f(v,s|x) ds dv}, \\ p(N,x) &= \int_{r(x)}^{\bar{v}(x)} \int_{\bar{s}(N,x)}^{\bar{v}(x)} F(v,s|N,x) f(s|N,x) ds dv. \end{aligned}$$

In the S model, the cutoff $\bar{s}(N,x)$ is determined as an implicit function from the equation $(\bar{s} - r(x)) F(\bar{s}|x)^{N-1} - k(x) = 0$. The derivative with respect to $\bar{s}$ of the expression on the left-hand side is $F(\bar{s}|x)^{N-1} + (\bar{s} - r(x))(N-1) f(\bar{s}|x) F(\bar{s}|x)^{N-2} > 0$. Assumptions of the lemma also imply that the left-hand side of equation determining $\bar{s}(N,x)$ has continuous partial derivatives up to order R with respect to $\bar{s}$ and x.

The Implicit Function Theorem then implies that the solution $\bar{s}(N,x)$ has continuous derivatives up to order R in x on the interior of $\mathcal{X}$. The result follows since in this case for $v \geq \bar{s}(N,x)$, $f^*(v|N,x) = f(v|x)/p(N,x)$, and $p(N,x) = 1 - F(\bar{s}(N,x)|x)$. ■

We can now prove the following result about the order of smoothness of $g(b|N,x)$.[1]

Lemma S.2 *Suppose that Assumptions 8(f) and (g) hold. Then for all $N \in \mathcal{N}$, the conditional PDF of bids $g^*(b|N,x)$ is strictly positive and bounded away from zero on its support $[\underline{b}(N,x), \bar{b}(N,x)]$, and $g^*(\cdot|N,\cdot)$ admits up to R continuous partial derivatives on the interior of the set $\{(b,x) : x \in \mathcal{X}, b \in [\underline{b}(N,x), \bar{b}(N,x)]\}$.*

Proof of Lemma S.2. First, we establish the order of smoothness of inverse bidding strategy $\xi(v|N,x)$ is R in both v and x. It is straightforward to show that differential equation (3) can be re-written in terms of $\xi(b|N,x)$ for $b \in (\underline{b}(N,x), \bar{b}(N,x))$ as

$$\begin{aligned} \frac{\partial \xi(b|N,x)}{\partial b} &= \frac{1}{N-1} \frac{1}{\xi(b|N,x) - b} \frac{p(N,x) f^*(\xi(b|N,x)|N,x)}{1 - p(N,x) + p(N,x) F^*(\xi(b|N,x)|N,x)} \qquad \text{(S.2)} \\ &\equiv \Phi_N(\xi(b|N,x), x). \end{aligned}$$

By Assumption 8(f), $f^*(v|N,x)$ admits $R-1$ derivatives with respect to v and R derivatives with respect to x, while $p(N,x)$ admits R derivatives with respect to x. Therefore $\Phi_N(v,x)$ also admits $R-1$ derivatives with respect to its first argument and R derivatives with respect to second. A fundamental results in the theory of differential equations (see, for example, Theorem 2.6 in Anosov, Aranson, Arnold, Bronshtein, Grines, and Il'Yashenko (1997)) implies that $\xi(\cdot|N,x)$ admits R derivatives on $(\underline{b}(N,x), \bar{b}(N,x))$ as a solution of this differential equation. Also, $\xi(b|N,\cdot)$ admits R partial derivatives on the interior of $\mathcal{X}$. Next, since $G^*(b|N,x) = F^*(\xi(b|N,x)|N,x)$, we have $g(b|N,x) = f^*(\xi(b|N,x)|N,x)\,\partial\xi(b|N,x)/\partial b$. Substituting $\partial\xi(b|N,x)/\partial b$ from $(S.2)$ yields (note that $f^*(\xi(b)|N,x)$ cancels out):

$$g^*(b|N,x) = \frac{1}{N-1} \frac{1}{\xi(b|N,x) - b} \frac{p(N,x)}{1 - p(N,x) + p(N,x) F^*(\xi(b|N,x)|N,x)}.$$

The result follows from the just established order of smoothness of $\xi(\cdot|N,\cdot)$ and Assumptions 8(f) and (g). ■

[1] This result parallels Proposition 1(iii) in GPV.

S.3 Consistency and asymptotic normality of the estimators in MSX

For kernel estimation, we use kernel functions K satisfying the following standard assumption (see, for example, Newey (1994)).

The standard nonparametric regression arguments imply that the estimator of entry probabilities $\hat{p}(N,x)$ is asymptotically normal (see, for example, Pagan and Ullah (1999), Theorem 3.5, page 110):

Proposition S.1 *Assume that the bandwidth h satisfies $Lh^d \to \infty$ and $\sqrt{Lh^d}h^R \to 0$ as $L \to \infty$. Then, for x in the interior of $\mathcal{X}$ and under Assumptions 8 and 9, $\sqrt{Lh^d}(\hat{p}(N,x) - p(N,x))$ is asymptotically normal with mean zero and variance*

$$V_p(N,x) = \frac{p(N,x)(1-p(N,x))}{N\pi(N|x)\varphi(x)} \left(\int K(u)^2 du\right)^d .$$

Moreover, the estimators $\hat{p}(N,x)$ are asymptotically independent for any distinct $N, N' \in \{\underline{N}, ...\bar{N}\}$ and x, x' in the interior of $\mathcal{X}$.

Since the distribution of values and, consequently, the distribution bids have compact supports, the estimator of the PDF $g^*(b|N,x)$ in Section 4.3 is asymptotically biased near the boundaries of the bids' support. Our quantile approach allows one to avoid the problem by considering only inner intervals of the supports. Specifically given $N \in \mathcal{N}$ and x in the interior of $\mathcal{X}$, let $0 < \tau_1(N,x) < \tau_2(N,x) < 1$. In our approach, we use quantiles $Q^*(\tau|N,x)$ with $\tau \in [\tau_1(N,x), \tau_2(N,x)]$. Due to the assumptions on $f^*(v|N,x)$ and since the bidding function is monotone, there are $b_1(N,x)$ and $b_2(N,x)$ such that

$$[b_1(N,x), b_2(N,x)] \subset \left(\underline{b}(N,x), \bar{b}(N,x)\right), \text{ and} \tag{S.3}$$

$$[q^*(\tau_1(N,x)|N,x), q^*(\tau_2(N,x)|N,x)] \subset (b_1(N,x), b_2(N,x)). \tag{S.4}$$

By (S.3), the estimator $\hat{g}^*(b|N,x)$ in Section 4.3 consistently estimates $g^*(b|N,x)$ on the interval $[b_1(N,x), b_2(N,x)]$ as we show in the lemma below. Condition (S.4) is used for establishing consistency of $\hat{Q}^*(\hat{q}^*(\tau|N,x)|N,x)$.

In practice, $\tau_1(N,x)$ and $\tau_2(N,x)$ can be selected as follows. Following the discussion on page 531 of GPV, in the case with no covariates one can choose $\tau_1(N)$ and $\tau_2(N)$ such that

$$[\hat{q}^*(\tau_1(N)|N), \hat{q}^*(\tau_2(N)|N)] \subset (b_{\min}(N) + h, b_{\max}(N) - h),$$

where $b_{\min}(N)$ and $b_{\max}(N)$ denote the minimum and maximum bids respectively in the auctions with $N_l = N$. When there are covariates available, one can replace $b_{\min}(N)$ and $b_{\max}(N)$ with the corresponding minimum and maximum bids in the neighborhood of x as defined on page 541 of GPV.

Lemma S.3 *Under Assumptions 8 and 9, for all x in the interior of $\mathcal{X}$ and $N \in \mathcal{N}$,*

(a) $\hat{\varphi}(x) - \varphi(x) = O_p((Lh^d/\log L)^{-1/2} + h^R)$.

(b) $\hat{\pi}(N|x) - \pi(N|x) = O_p((Lh^d/\log L)^{-1/2} + h^R)$.

(c) $\hat{p}(N,x) - p(N,x) = O_p((Lh^d/\log L)^{-1/2} + h^R)$.

(d) $\sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} |\hat{G}^*(b|N,x) - G^*(b|N,x)| = O_p((Lh^d/\log L)^{-1/2} + h^R)$.

(e) $\sup_{\tau\in[\varepsilon,1-\varepsilon]} |\hat{q}^*(\tau|N,x) - q^*(\tau|N,x)| = O_p((Lh^d/\log L)^{-1/2} + h^R)$, *for any* $0 < \varepsilon < 1/2$.

(f) $\sup_{b\in[b_1(N,x),b_2(N,x)]} |\hat{g}^*(b|N,x) - g^*(b|N,x)| = O_p((Lh^{d+1}/\log L)^{-1/2} + h^R)$, *where* $b_1(N,x)$ *and* $b_2(N,x)$ *are defined in (S.3) and (S.4).*

(g) $\sup_{\tau\in[\tau_1(N,x),\tau_2(N,x)]} |\hat{Q}^*(\tau|N,x) - Q^*(\tau|N,x)| = O_p((Lh^{d+1}/\log L)^{-1/2} + h^R)$.

(h) $\hat{Q}^*(\hat{\beta}(\tau,N,x)|N,x) = Q^*(\beta(\tau,N,x)|N,x) + O_p((Lh^{d+1}/\log L)^{-1/2} + h^R)$ *uniformly in* τ *such that* $\beta(\tau,N,x) \in [\tau_1(N,x)+\varepsilon, \tau_2(N,x)-\varepsilon]$, *for any* $0 < \varepsilon < (\tau_2(N,x) - \tau_1(N,x))/2$.

Proof of Lemma S.3. Parts (a)-(c) of the lemma follow from Lemma B.3 of Newey (1994).

For part (d), define

$$G_0^*(b,N,x) = Np(N,x)\,\pi(N|x)\,G^*(b|N,x)\,\varphi(x),$$

and its estimator

$$\hat{G}_0^*(b,N,x) = \frac{1}{L}\sum_{l=1}^{L}\sum_{i=1}^{N_l} y_{il} 1\{N_l = N\}\, 1(b_{il} \leq b)\, K_{*h}(x_l - x),$$

where

$$K_{*h}(x_l - x) \quad = \quad \frac{1}{h^d} K_d\left(\frac{x_l - x}{h}\right), \text{ and}$$

$$K_d\left(\frac{x_l - x}{h}\right) = \prod_{k=1}^{d} K\left(\frac{x_{kl} - x_k}{h}\right). \tag{S.5}$$

Similarly to Lemma B.2 of Newey (1994), by Lemma S.2 and Assumptions 8(b), (c), and (g),

$$\sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} \left|G_0^*(b,N,x) - E\hat{G}_0^*(b,N,x)\right| = O\left(h^R\right). \tag{S.6}$$

Next, we show that

$$\sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} |\hat{G}_0^*(b,N,x) - E\hat{G}_0^*(b,N,x)| = O_p\left(\left(\frac{Lh^d}{\log L}\right)^{-1/2}\right). \tag{S.7}$$

We follow the approach of Pollard (1984). Consider, for given $N \in \mathcal{N}$ and x in the interior of $\mathcal{X}$, a class of functions $\mathcal{Z}$ indexed by h and b, with a representative function

$$z_l(b,N,x) = \sum_{i=1}^{N_l} y_{il} 1\{N_l = N\} 1(b_{il} \leq b) h^d K_{*h}(x_l - x).$$

By the result in Pollard (1984) (Problem 28), the class $\mathcal{Z}$ has polynomial discrimination. Theorem 37 in Pollard (1984) (see also Example 38) implies that for any sequences δ_L, α_L such that $L\delta_L^2\alpha_L^2/\log L \to \infty$, $Ez_l^2(b) \leq \delta_L^2$,

$$\alpha_L^{-1}\delta_L^{-2} \sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} |\frac{1}{L}\sum_{l=1}^{L} z_l(b,N,x) - Ez_l(b,N,x)| \to 0 \tag{S.8}$$

almost surely. We claim that this implies that

$$\left(\frac{Lh^d}{\log L}\right)^{1/2} \sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} |\hat{G}_0^*(b,N,x) - E\hat{G}_0^*(b,N,x)|$$

is bounded as $L \to \infty$ almost surely, which in turn implies the result in (S.7). The proof is by contradiction. Suppose not. Then there exist a sequence $\gamma_L \to \infty$ and a subsequence of L such that along this subsequence

$$\sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} |\hat{G}_0^*(b,N,x) - E\hat{G}_0^*(b,N,x)| \geq \gamma_L\left(\frac{Lh^d}{\log L}\right)^{-1/2}. \tag{S.9}$$

on a set of events $\Omega' \subset \Omega$ with a positive probability measure. Now if we let $\delta_L^2 = h^d$ and $\alpha_L = (\frac{Lh^d}{\log L})^{-1/2}\gamma_L^{1/2}$, then the definition of z implies that, along the subsequence on a set of events Ω',

$$\begin{aligned}
& \alpha_L^{-1}\delta_L^{-2} \sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} \left|\frac{1}{L}\sum_{l=1}^{L} z_l\left(b,N,x\right) - Ez_l\left(b,N,x\right)\right| \\
= & \left(\frac{Lh^d}{\log L}\right)^{1/2} \gamma_L^{-1/2} h^{-d} \sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} \left|\frac{1}{L}\sum_{l=1}^{L} z_l\left(b,N,x\right) - Ez_l\left(b,N,x\right)\right| \\
= & \left(\frac{Lh^d}{\log L}\right)^{1/2} \gamma_L^{-1/2} \sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} \left|\hat{G}_0^*\left(b,N,x\right) - E\hat{G}_0^*\left(b,N,x\right)\right| \\
\geq & \left(\frac{Lh^d}{\log L}\right)^{1/2} \gamma_L^{-1/2}\gamma_L \left(\frac{Lh^d}{\log L}\right)^{-1/2} \\
= & \gamma_L^{1/2} \to \infty,
\end{aligned}$$

where the inequality follows by $(S.9)$, a contradiction to $(S.8)$. This establishes $(S.7)$, so that $(S.6)$, $(S.7)$ and the triangle inequality together imply that

$$\sup_{b\in\left[\underline{b}(N,x),\bar{b}(N,x)\right]} \left|\hat{G}_0^*\left(b,N,x\right) - G_0^*\left(b,N,x\right)\right| = O_p\left(\left(\frac{Lh^d}{\log L}\right)^{-1/2} + h^R\right). \qquad \text{(S.10)}$$

To complete the proof, recall that from the definitions of $G_0^*\left(b,N,x\right)$ and $\hat{G}_0^*\left(b,N,x\right)$,

$$G^*\left(b|N,x\right) = \frac{G_0^*\left(b,N,x\right)}{p\left(N,x\right)\pi\left(N|x\right)\varphi\left(x\right)} \text{ and } \hat{G}^*\left(b|n,x\right) = \frac{\hat{G}_0^*\left(b,N,x\right)}{\hat{p}\left(N,x\right)\hat{\pi}\left(N|x\right)\hat{\varphi}\left(x\right)},$$

so that by the mean-value theorem,

$$\left|\hat{G}^*\left(b|N,x\right) - G^*\left(b|N,x\right)\right| \leq \tilde{C}\left(b,N,x\right) \left\| \begin{pmatrix} \hat{G}_0^*\left(b,N,x\right) - G_0^*\left(b,N,x\right) \\ \hat{p}\left(N,x\right) - p\left(N,x\right) \\ \hat{\pi}\left(N|x\right) - \pi\left(N|x\right) \\ \hat{\varphi}\left(x\right) - \varphi\left(x\right) \end{pmatrix} \right\|, \qquad \text{(S.11)}$$

where $\|\cdot\|$ denotes the Euclidean norm, $\tilde{C}\left(b,N,x\right)$ is given by

$$\left|\frac{1}{\tilde{p}\left(N,x\right)\tilde{\pi}\left(N,x\right)\tilde{\varphi}\left(x\right)}\right| \left\| \left(1, \frac{\tilde{G}_0^*\left(b,N,x\right)}{\tilde{p}\left(N,x\right)}, \frac{\tilde{G}_0^*\left(b,N,x\right)}{\tilde{\pi}\left(N,x\right)}, \frac{\tilde{G}_0^*\left(b,N,x\right)}{\tilde{\varphi}\left(x\right)}\right) \right\|,$$

and $\|(\tilde{G}_0^* - G_0^*, \tilde{p} - p, \tilde{\pi} - \pi, \tilde{\varphi} - \varphi)\| \leq \|(\hat{G}_0^* - G_0^*, \hat{p} - p, \hat{\pi} - \pi, \hat{\varphi} - \varphi)\|$. Further by Assumption 8(b), (c), and (g), and the results in parts (a)-(c) of the lemma, with probability approaching one $\tilde{\varphi}$, $\tilde{\pi}$, and $\tilde{p}$ are bounded away from zero. The desired result follows from (S.10), (S.11) and parts (a)-(c) of the lemma.

For part (e) of the lemma, since $\hat{G}^*(\cdot|N,x)$ is monotone by construction,

$$\begin{aligned} P\left(\hat{q}^*(\varepsilon|N,x) \leq \underline{b}(N,x)\right) &= P\left(\inf_b \left\{b : \hat{G}^*(b|N,x) \geq \varepsilon\right\} \leq \underline{b}(N,x)\right) \\ &= P\left(\hat{G}^*\left(\underline{b}(N,x)|N,x\right) \geq \varepsilon\right) \\ &= o(1), \end{aligned}$$

where the last equality is by the result in part (d). Similarly,

$$\begin{aligned} P\left(\hat{q}^*(1-\varepsilon|N,x) \geq \overline{b}(N,x)\right) &= P\left(\hat{G}^*\left(\overline{b}(N,x)|N,x\right) \leq 1-\varepsilon\right) \\ &= o(1). \end{aligned}$$

Hence, for all x in the interior of $\mathcal{X}$ and $N \in \mathcal{N}$, $\underline{b}(N,x) < \hat{q}^*(\varepsilon|N,x) < \hat{q}^*(1-\varepsilon|N,x) < \overline{b}(N,x)$ with probability approaching one. Since the distribution $G^*(b|N,x)$ is continuous in b, $G^*(q^*(\tau|N,x)|N,x) = \tau$, and for $\tau \in [\varepsilon, 1-\varepsilon]$, we can write the identity

$$G^*(\hat{q}^*(\tau|N,x)|N,x) - G^*(q^*(\tau|N,x)|N,x) = G^*(\hat{q}^*(\tau|N,x)|N,x) - \tau. \quad \text{(S.12)}$$

Next,

$$0 \leq \hat{G}^*(\hat{q}^*(\tau|N,x)|N,x) - \tau \leq \frac{(\sup_u K(u))^d}{\hat{p}(N,x)\,\hat{\pi}(N|x)\,\hat{\varphi}(x)\,NLh^d}, \quad \text{(S.13)}$$

where the first inequality is by Lemma 21.1(ii) in van der Vaart (1998), and the second inequality holds with probability one since $\hat{G}^*(\cdot|N,x)$ is an empirical CDF and the distribution of bids is continuous so that ties occur with probability zero. By (S.13) and the results in (a)-(c),

$$\hat{G}^*(\hat{q}^*(\tau|N,x)|N,x) = \tau + O_p\left(\left(Lh^d\right)^{-1}\right) \quad \text{(S.14)}$$

uniformly over $\tau \in [\varepsilon, 1-\varepsilon]$. Combining ($S$.12) and ($S$.14), and applying the mean-value theorem to the left-hand side of (S.12), we obtain

$$\hat{q}^*(\tau|N,x) - q^*(\tau|N,x)$$

$$= \frac{G^*\left(\hat{q}^*\left(\tau|N,x\right)|N,x\right) - \hat{G}^*\left(\hat{q}^*\left(\tau|N,x\right)|N,x\right)}{g^*\left(\widetilde{q}^*\left(\tau|N,x\right)|N,x\right)} + O_p\left(\left(Lh^d\right)^{-1}\right), \quad \text{(S.15)}$$

where $\widetilde{q}^*$ lies between $\hat{q}^*$ and q^* for all (τ, N, x). Now, by Lemma S.2, $g^*(b|N,x)$ is bounded away from zero, and the result in part (e) follows from (S.15) and part (d) of the lemma.

To prove part (f), by Lemma S.2, $g^*(\cdot|N,\cdot)$ admits up to R continuous bounded partial derivatives. Let

$$\begin{aligned} g_0^*(b,N,x) &= p(N,x)\,\pi(N|x)\,\varphi(x)\,g^*(b|N,x)\,, \text{ and} \qquad \text{(S.16)}\\ \hat{g}_0^*(b,N,x) &= \hat{p}(N,x)\,\hat{\pi}(N|x)\,\hat{\varphi}(x)\,\hat{g}^*(b|N,x)\,. \qquad \text{(S.17)} \end{aligned}$$

By Lemma B.3 of Newey (1994), $\hat{g}_0^*(b,N,x)$ is uniformly consistent in b over the interval $[b_1(N,x), b_2(N,x)]$. By the results in parts (a)-(c), the estimators $\hat{p}(N,x)$, $\hat{\pi}(N|x)$, and $\hat{\varphi}(x)$ converge at a faster rate than that of $\hat{g}_0^*(b,N,x)$. The desired result follows by the same argument as in the proof of part (d), equation $(S.11)$.

Next, we prove part (g). By Lemma S.2, $g^*(b|N,x) > c_g > 0$ for some constant c_g. Then,

$$\begin{aligned} &\left|\hat{Q}^*(\tau|N,x) - Q^*(\tau|N,x)\right| \\ \leq\ & \left|\hat{q}^*(\tau|N,x) - q^*(\tau|N,x)\right| + 2\frac{\left|\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right) - g^*\left(q^*(\tau|N,x)|N,x\right)\right|}{p(N,x)\,\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right)c_g} \\ &+\frac{\left|\hat{p}(N,x) - p(N,x)\right|}{\hat{p}(N,x)\,p(N,x)\,\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right)} \\ \leq\ & \left(1 + \frac{2\sup_{b\in[b_1(N,x),b_2(N,x)]}\left|\partial g^*(b|N,x)/\partial b\right|}{p(N,x)\,\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right)c_g}\right)\left|\hat{q}^*(\tau|n,x) - q^*(\tau|n,x)\right| \\ &+2\frac{\left|\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right) - g^*\left(\hat{q}^*(\tau|N,x)|N,x\right)\right|}{p(N,x)\,\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right)c_g} \\ &+\frac{\left|\hat{p}(N,x) - p(N,x)\right|}{\hat{p}(N,x)\,p(N,x)\,\hat{g}^*\left(\hat{q}^*(\tau|N,x)|N,x\right)}. \qquad \text{(S.18)} \end{aligned}$$

Define the event

$$E_L(N,x) = \left\{\hat{q}^*\left(\tau_1(N,x)|N,x\right) \geq b_1(N,x)\,, \hat{q}^*\left(\tau_2(N,x)|N,x\right) \leq b_2(N,x)\right\},$$

and let $\beta_L = \left(Lh^{d+1}/\log L\right)^{-1/2} + h^R$. By the result in part (e), $P\left(E_L^c(N,x)\right) = o(1)$. Hence, it follows from part (e) of the lemma that the estimator $\hat{g}^*(\hat{q}^*(\tau|N,x)|N,x)$

is bounded away from zero with probability approaching one, and the first term on the right-hand side of $(S.18)$ is $O_p(\beta_L)$ uniformly over $[\tau_1(N,x), \tau_2(N,x)]$. Next,

$$\begin{aligned}
& P\left(\sup_{\tau\in[\tau_1(N,x),\tau_2(N,x)]} \beta_L^{-1}\left|\hat{g}^*\left(\hat{q}^*\left(\tau|N,x\right)|N,x\right) - g^*\left(\hat{q}^*\left(\tau|N,x\right)|N,x\right)\right| > M\right) \\
\leq\ & P\left(\sup_{\tau\in[\tau_1(N,x),\tau_2(N,x)]} \beta_L^{-1}\left|\hat{g}^*\left(\hat{q}^*\left(\tau|N,x\right)|N,x\right) - g^*\left(\hat{q}^*\left(\tau|N,x\right)|N,x\right)\right| > M, E_L(x)\right) \\
& +P\left(E_L^c(x)\right) \\
\leq\ & P\left(\sup_{b\in[b_1(N,x),b_2(N,x)]} \beta_L^{-1}\left|\hat{g}^*\left(b|N,x\right) - g^*\left(b|N,x\right)\right| > M\right) + o(1). \qquad \text{(S.19)}
\end{aligned}$$

Part (g) follows from (S.18) and (S.19) by the results in parts (c) and (f) of the lemma.

For part (h), first for all $\tau \in [0,1]$,

$$\left|\hat{\beta}(\tau,N,x) - \beta(\tau,N,x)\right| \leq \left|\frac{\hat{p}\left(\bar{N},x\right)}{\hat{p}(N,x)} - \frac{p\left(\bar{N},x\right)}{p(N,x)}\right|,$$

and therefore by the result in part (c) of the lemma, $\sup_{\tau\in[0,1]}|\hat{\beta}(\tau,N,x) - \beta(\tau,N,x)| = O_p((Lh^d/\log L)^{-1/2} + h^R)$ for all $N \in \mathcal{N}$ and x in the interior of $\mathcal{X}$. The desired result then follows by the triangular inequality, uniform consistency of $\hat{\beta}(\tau,N,x)$, the result in part (g) of the lemma, differentiability of $Q^*(\cdot|N,x)$, the mean-value theorem, and since $f^*(\cdot|N,x)$ is bounded away from zero by Assumption 8(f). ■

Lemma S.4 *Suppose that Assumptions 8 and 9 hold, and that the bandwidth h is such that $Lh^{d+1} \to \infty$, $\sqrt{Lh^{d+1}}h^R \to 0$. Then,*

$$\sqrt{Lh^{d+1}}\left(\hat{g}^*(b|N,x) - g^*(b|N,x)\right) \to_d N\left(0, V_g(b,N,x)\right)$$

for $b \in [b_1(N,x), b_2(N,x)]$, x in the interior of $\mathcal{X}$, $N \in \mathcal{N}$, where $b_1(N,x)$ and $b_2(N,x)$ are defined in (S.3) and (S.4), and $V_g(b,N,x)$ is given by

$$V_g(N,b,x) = \frac{g^*(b|N,x)}{Np(N,x)\pi(N|x)\varphi(x)}\left(\int K(u)^2\,du\right)^{d+1}.$$

Furthermore, $\hat{g}^*(b|N_1,x)$ and $\hat{g}^*(b|N_2,x)$ are asymptotically independent for all $N_1 \neq N_2$, $N_1, N_2 \in \mathcal{N}$.

Proof of Lemma S.4. Consider $g_0^*(b,n,x)$ and $\hat{g}_0^*(b,n,x)$ defined in $(S.16)$ and $(S.17)$ respectively. It follows from parts (a)-(c) of Lemma S.3,

$$\begin{aligned}&\sqrt{Lh^{d+1}}\left(\hat{g}^*\left(b|N,x\right)-g^*\left(b|N,x\right)\right)\\&\qquad=\frac{1}{p\left(N,x\right)\pi\left(N|x\right)\varphi\left(x\right)}\sqrt{Lh^{d+1}}\left(\hat{g}_0^*\left(b,N,x\right)-g_0^*\left(b,N,x\right)\right)+o_p(1).\end{aligned}\tag{S.20}$$

Furthermore, as in Lemma B2 of Newey (1994), $E\hat{g}_0^*\left(b,N,x\right)-g_0^*\left(b,N,x\right)=O\left(h^R\right)$ uniformly in $b\in\left[b_1\left(N,x\right),b_2\left(N,x\right)\right]$ for all x in the interior of $\mathcal{X}$ and $N\in\mathcal{N}$. Thus, it remains to establish asymptotic normality of $\sqrt{Lh^{d+1}}\left(\hat{g}_0^*\left(b,N,x\right)-E\hat{g}_0^*\left(b,N,x\right)\right)$.

Define

$$\begin{aligned}w_{il,N}&=\sqrt{\frac{1}{h^{d+1}}}y_{il}1\left\{N_l=N\right\}K\left(\frac{b_{il}-b}{h}\right)K_d\left(\frac{x_l-x}{h}\right),\\\overline{w}_{L,N}&=\frac{1}{NL}\sum_{l=1}^{L}\sum_{i=l}^{N_l}w_{il,N},\end{aligned}$$

where K_d is defined in (S.5). With the above definitions we have that

$$\sqrt{NLh^{d+1}}\left(\hat{g}_0^*\left(b,N,x\right)-E\hat{g}_0^*\left(b,N,x\right)\right)=\sqrt{NL}\left(\overline{w}_{L,N}-E\overline{w}_{L,N}\right).\tag{S.21}$$

Then, by the Liapunov CLT (see, for example, Corollary 11.2.1 on page 427 of Lehman and Romano (2005)),

$$\sqrt{NL}\left(\overline{w}_{L,N}-E\overline{w}_{L,N}\right)/\sqrt{NLVar\left(\overline{w}_{L,N}\right)}\to_d N\left(0,1\right),\tag{S.22}$$

provided that $Ew_{il,N}^2<\infty$, and for some $\delta>0$,

$$\lim_{L\to\infty}\frac{1}{L^{\delta/2}}E\left|w_{il,N}-Ew_{il,N}\right|^{2+\delta}=0.$$

The last condition follows from the Liapunov's condition (equation (11.12) on page 427 of Lehman and Romano (2005)) and because $w_{il,N}$ are i.i.d. Next, $Ew_{il,N}$ is given by

$$\begin{aligned}&\sqrt{\frac{1}{h^{d+1}}}E\left(p\left(N,x_l\right)\pi\left(N|x_l\right)\int K\left(\frac{u-b}{h}\right)g^*\left(u|N,x_l\right)duK_d\left(\frac{x_l-x}{h}\right)\right)\\=&\sqrt{\frac{1}{h^{d+1}}}\int\int p\left(N,y\right)\pi\left(N|y\right)K\left(\frac{u-b}{h}\right)g^*\left(u|N,y\right)K_d\left(\frac{y-x}{h}\right)\varphi\left(y\right)dudy\end{aligned}$$

$$
\begin{aligned}
&= \sqrt{h^{d+1}} \\
&\quad \times \int\int p\left(N,x+hy\right)\pi\left(N|x+hy\right)K\left(u\right)g^{*}\left(b+hu|N,x+hy\right)K_{d}\left(y\right)\varphi\left(x+hy\right)dudy \\
&\to 0.
\end{aligned}
$$

Further, $Ew_{il,N}^{2}$ is given by

$$
\begin{aligned}
&\frac{1}{h^{d+1}}\int\int p\left(N,y\right)\pi\left(N|y\right)K^{2}\left(\frac{u-b}{h}\right)g^{*}\left(u|N,y\right)K_{d}^{2}\left(\frac{y-x}{h}\right)\varphi\left(y\right)dudy \\
&= \int\int p\left(N,x+hy\right)\pi\left(N|x+hy\right)K^{2}\left(u\right)g^{*}\left(b+hu|N,x+hy\right)K_{d}^{2}\left(y\right)\varphi\left(x+hy\right)dudy \\
&< \infty.
\end{aligned}
$$

Hence,

$$
NLVar\left(\overline{w}_{L,N}\right)\to p\left(N,x\right)\pi\left(N|x\right)g^{*}\left(b|N,x\right)\varphi\left(x\right)\left(\int K^{2}\left(u\right)du\right)^{d+1}du. \quad \text{(S.23)}
$$

Next, $E\left|w_{il,N}\right|^{2+\delta}$ is bounded by

$$
\begin{aligned}
&\frac{1}{h^{(d+1)(1+\delta/2)}}\int\int\left|K\left(\frac{u-b}{h}\right)\right|^{2+\delta}g^{*}\left(u|N,y\right)\left|K_{d}\left(\frac{y-x}{h}\right)\right|^{2+\delta}\varphi\left(y\right)dudy \\
&= \frac{1}{h^{(d+1)\delta/2}}\int\int\left|K\left(u\right)\right|^{2+\delta}g^{*}\left(b+hu|N,x+hy\right)\left|K_{d}\left(y\right)\right|^{2+\delta}\varphi\left(x+hy\right)dudy \\
&\leq \frac{1}{h^{(d+1)\delta/2}}\sup_{u\in[-1,1]}\left|K\left(u\right)\right|^{(d+1)(2+\delta)}\sup_{x\in\mathcal{X}}\varphi\left(x\right)\sup_{b\in[b_{1}(N,x),b_{2}(N,x)]}g^{*}\left(b|N,x\right) \\
&= \frac{C}{h^{(d+1)\delta/2}}.
\end{aligned}
$$

Lastly,

$$
\begin{aligned}
\frac{1}{L^{\delta/2}}E\left|w_{il,N}-Ew_{il,N}\right|^{2+\delta} &\leq \frac{2^{1+\delta}}{L^{\delta/2}}E\left|w_{il,N}\right|^{2+\delta} \\
&\leq \frac{2^{1+\delta}C}{\left(Lh^{d+1}\right)^{\delta/2}} \\
&\to 0, \qquad \text{(S.24)}
\end{aligned}
$$

since $Lh^{d+1}\to\infty$ by the assumption. The first result of the lemma follows now from $(S.20)$-$(S.24)$.

Next, note that the asymptotic covariance of $\overline{w}_{L,N_1}$ and $\overline{w}_{L,N_2}$ involves a product of the two indicator functions, $1\{N_l = N_1\} 1\{N_l = N_2\}$, which is zero for all $N_1 \neq N_2$. The joint asymptotic normality and asymptotic independence of $\hat{g}^* (b|N_1, x)$ and $\hat{g}^* (b|N_2, x)$ follows then by the Cramér-Wold device. ■

Proposition S.2 *Assume that the bandwidth h satisfies $Lh^{d+1} \to \infty$ and $\sqrt{Lh^{d+1}}h^R \to 0$. Then, for $\tau \in (0,1)$, x in the interior of $\mathcal{X}$, and under Assumptions 8 and 9,*

$$\sqrt{Lh^{d+1}} \left(\hat{Q}^* (\tau|N, x) - Q^* (\tau|N, x)\right) \to_d N (0, V_Q (N, \tau, x)),$$

$$\sqrt{Lh^{d+1}} \left(\hat{Q}^* \left(\hat{\beta} (\tau, N, x) |N, x\right) - Q^* (\beta (\tau, N, x) |N, x)\right) \to_d N (0, V_Q (N, \beta (\tau, N, x), x))),$$

where

$$V_Q (N, \tau, x) = \left(\frac{1 - p (N, x) (1 - \tau)}{(N - 1)p (N, x) g^{*2}(q^* (\tau|N, x) |N, x)}\right)^2 V_g (N, q^* (\tau|N, x), x),$$

and $V_g (N, \tau, x)$ is defined in Lemma S.4. Moreover, for any distinct $N, N' \in \{\underline{N}, ...\bar{N}\}$, $\tau, \tau' \in \Upsilon$, and x, x' in the interior of $\mathcal{X}$, the estimators $\hat{Q}^* (\tau|N, x)$ are asymptotically independent, as well as the estimators $\hat{Q}^* \left(\hat{\beta} (\tau, N, x) |N, x\right)$.

Proof of Proposition S.2. First, by Lemma S.3 (c), (e) and (f), and the mean-value theorem,

$$\hat{Q}^* (\tau|N, x) = Q^* (\tau|N, x) - \frac{1 - p (N, x) (1 - \tau)}{(N - 1)p (N, x) \widetilde{g}^{*2}(q^* (\tau|N, x) |N, x)} \times (\hat{g}^* (q^* (\tau|N, x)) - g^* (q^* (\tau|N, x))) + o_p \left(\frac{1}{\sqrt{Lh^{d+1}}}\right), \quad \text{(S.25)}$$

where $\widetilde{g}^*$ is a mean-value between g^* and $\hat{g}^*$ for $b = q^* (\tau|N, x)$. The result follows then by Lemma S.4. ■

S.4 Validity of the bootstrap

In this section, we establish the bootstrap validity for the SEM test in (29) in MSX. For the other tests, the proof is analogous and therefore omitted.

S.4.1 Notation and auxiliary results

In what follows, the statistics with superscript † denote the bootstrap analogues of the statistics computed using the original data. To simplify the notion, we will suppress the subscript indicating the bootstrap sample number for bootstrap objects (m). Let $P^\dagger$ denote probability conditional on the original sample. We use $E^\dagger$ and $Var^\dagger$ to denote expectation and variance under $P^\dagger$ respectively.

Let $\pi^\dagger$ denote the distribution of $N_l^\dagger$ implied by $P^\dagger$, i.e.

$$\begin{aligned}\pi^\dagger(N) &= P^\dagger(N_l^\dagger = N)\\ &= L^{-1}\sum_{l=1}^{L} 1(N_l = N)\\ &= \hat{\pi}\left(N\right),\end{aligned}$$

where $\pi(N) = P(N_l = N)$. Also, define

$$\begin{aligned}p^\dagger(N) &= P^\dagger(y_{il}^\dagger = 1|N)\\ &= \sum_{l=1}^{L}\left(n_l/N\right)P^\dagger\left(n_l^\dagger = n_l|N\right)\\ &= \frac{\sum_{l=1}^{L}\left(n_l/N\right)1\left\{N_l = N\right\}}{\sum_{l=1}^{L}1\left\{N_l = N\right\}}\\ &= \hat{p}\left(N\right),\end{aligned}$$

where $p\left(N\right) = P\left(y_{il} = 1|N_l = N\right)$.

We say $\zeta_L = O_p^\dagger(\lambda_L)$ if for all $\varepsilon > 0$ there is $\Delta_\varepsilon > 0$ such that for all $L \geq L_\varepsilon$, $P(P^\dagger(|\zeta_L/\lambda_L| > \Delta_\varepsilon) > \varepsilon) < \varepsilon$. We say $\zeta_L = o_p^\dagger\left(\lambda_L\right)$ if $P^\dagger\left(|\zeta_L/\lambda_L| > \varepsilon\right) \to_p 0$ for all $\varepsilon > 0$ as $L \to \infty$.

Next, we present some simple results concerning the stochastic order (with respect to $P^\dagger$) of the bootstrap statistics. Let $\hat{\theta}_L$ be a statistic computed using the data in the original sample, and let $\hat{\theta}_L^\dagger$ be the bootstrap analogue of $\hat{\theta}_L$.

Lemma S.5 (a) *Suppose that* $\hat{\theta}_L = \theta + o_p\left(\delta_L\right)$ *and* $\hat{\theta}_L^\dagger = \hat{\theta}_L + o_p^\dagger\left(\delta_L\right)$. *Then,* $\hat{\theta}_L^\dagger = \theta + o_p^\dagger\left(\delta_L\right)$.

(b) *Suppose that* $\hat{\theta}_L = \theta + O_p\left(\delta_L\right)$ *and* $\hat{\theta}_L^\dagger = \hat{\theta}_L + O_p^\dagger\left(\delta_L\right)$. *Then,* $\hat{\theta}_L^\dagger = \theta + O_p^\dagger\left(\delta_L\right)$.

Proof. For part (a), since $\hat{\theta}_L$ is not random under $P^\dagger$,

$$P^\dagger\left(\delta_L^{-1}\left|\hat{\theta}_L^\dagger - \theta\right| > \varepsilon\right) \leq P^\dagger\left(\delta_L^{-1}\left|\hat{\theta}_L - \theta\right| > \frac{\varepsilon}{2}\right) + P^\dagger\left(\delta_L^{-1}\left|\hat{\theta}_L^\dagger - \hat{\theta}_L\right| > \frac{\varepsilon}{2}\right)$$

$$= 1\left(\delta_L^{-1}\left|\hat{\theta}_L - \theta\right| > \frac{\varepsilon}{2}\right) + o_p(1).$$

For the first summand, we have that for all $\varepsilon, \eta > 0$,

$$P\left(1\left(\delta_L^{-1}\left|\hat{\theta}_L - \theta\right| > \frac{\varepsilon}{2}\right) > \eta\right) = P\left(\delta_L^{-1}\left|\hat{\theta}_L - \theta\right| > \frac{\varepsilon}{2}\right) \to 0.$$

The proof of part (b) is similar. ■

Lemma S.6 *Suppose that* $E^\dagger(\hat{\theta}_L^\dagger)^2 = O_p(\lambda_L^2)$. *Then* $\hat{\theta}_L^\dagger = O_p^\dagger(\lambda_L)$.

Proof. Since $E^\dagger(\hat{\theta}_L^\dagger)^2 = O_p(\lambda_L^2)$, for all $\varepsilon > 0$ there is $\Delta_\varepsilon > 0$ such that $P(E^\dagger(\hat{\theta}_L^\dagger)^2 > \Delta_\varepsilon^2\lambda_L^2) < \varepsilon$. Let $\tilde{\Delta}_\varepsilon^2 = \Delta_\varepsilon^2/\varepsilon$. Then, we can write

$$P(E^\dagger(\hat{\theta}_L^\dagger)^2 > \tilde{\Delta}_\varepsilon^2\varepsilon\lambda_L^2) < \varepsilon \tag{S.26}$$

for all L large enough. By Markov's inequality,

$$P^\dagger\left(\left|\frac{\hat{\theta}_L^\dagger}{\lambda_L}\right| \geq \tilde{\Delta}_\varepsilon\right) \leq \frac{E^\dagger\left(\hat{\theta}_L^\dagger\right)^2}{\lambda_L^2\tilde{\Delta}_\varepsilon^2}.$$

Thus, for all $\varepsilon > 0$ there is $\tilde{\Delta}_\varepsilon$, such that for all L large enough,

$$P\left(P^\dagger\left(\left|\frac{\hat{\theta}_L^\dagger}{\lambda_L}\right| \geq \tilde{\Delta}_\varepsilon\right) > \varepsilon\right) \leq P\left(\frac{E^\dagger\left(\hat{\theta}_L^\dagger\right)^2}{\tilde{\Delta}_\varepsilon^2\lambda_L^2} > \varepsilon\right) < \varepsilon,$$

where the last inequality is by (S.26). ■

The following lemma is the bootstrap counterpart of Lemma S.3 in Section S.3. It will be used for justifying the asymptotic linearity of the bootstrap statistic in (S.45) below.

Lemma S.7 *Suppose that the assumptions of Lemma S.3 hold. Then, for all* x *in the interior of* $\mathcal{X}$ *and* $N \in \mathcal{N}$,

(a) $\hat{\varphi}^\dagger(x) = \hat{\varphi}(x) + O_p^\dagger\left(Lh^d\right)^{-1/2}$.

(b) $\hat{\pi}^\dagger(N|x) = \hat{\pi}(N|x) + O_p^\dagger\left(Lh^d\right)^{-1/2}$.

(c) $\hat{p}^{\dagger}(N,x) = \hat{p}(N,x) + O_p^{\dagger}((Lh^d/\log L)^{-1/2} + h^R)$.

(d) $\sup_{b \in [\underline{b}(N,x), \bar{b}(N,x)]} |\hat{G}^{*,\dagger}(b|N,x) - \hat{G}^{*}(b|N,x)| = O_p^{\dagger}((Lh^d/\log L)^{-1/2} + h^R)$.

(e) $\sup_{\tau \in [\varepsilon, 1-\varepsilon]} |\hat{q}^{*,\dagger}(\tau|N,x) - \hat{q}^{*}(\tau|N,x)| = O_p^{\dagger}((Lh^d/\log L)^{-1/2} + h^R)$, *for any* $0 < \varepsilon < 1/2$.

(f) $\sup_{b \in [b_1(N,x), b_2(N,x)]} |\hat{g}^{*,\dagger}(b|N,x) - \hat{g}^{*}(b|N,x)| = O_p^{\dagger}((Lh^{d+1}/\log L)^{-1/2} + h^R)$, *where* $b_1(N,x)$ *and* $b_2(N,x)$ *are defined in (S.3) and (S.4) in MSX.*

(g) $\sup_{\tau \in [\tau_1(N,x), \tau_2(N,x)]} |\hat{Q}^{*,\dagger}(\tau|N,x) - \hat{Q}^{*}(\tau|N,x)| = O_p^{\dagger}((Lh^{d+1}/\log L)^{-1/2} + h^R)$.

(h) $\hat{Q}^{*,\dagger}(\hat{\beta}^{\dagger}(\tau,N,x)|N,x) = \hat{Q}^{*}(\beta(\tau,N,x)|N,x) + O_p^{\dagger}((Lh^{d+1}/\log L)^{-1/2} + h^R)$ *uniformly in* τ *such that* $\beta(\tau,N,x) \in [\tau_1(N,x)+\varepsilon, \tau_2(N,x)-\varepsilon]$, *for any* $0 < \varepsilon < (\tau_2(N,x) - \tau_1(N,x))/2$.

Proof. Part (a) follows from the uniform strong approximation in Chen and Lo (1997), Proposition 3.2.

For part (b), write

$$\hat{\pi}(N|x) = \hat{\pi}(N,x)\,\hat{\varphi}(x), \text{ where}$$
$$\hat{\pi}(N,x) = \frac{1}{Lh^d}\sum_{l=1}^{L} 1(N_l = N)\prod_{k=1}^{d} K\left(\frac{x_{kl} - x_k}{h}\right).$$

By Proposition 3.2 in Chen and Lo (1997), $(Lh^d)^{1/2}(\hat{\pi}(N,x) - E\hat{\pi}(N,x)) = O_p^{\dagger}(1)$. By the Taylor expansion of $\hat{\pi}^{\dagger}(N|x)$, the result in part (a), and since $\hat{\varphi}(x)$ is bounded away from zero with probability approaching one by Assumption 8(b) and Lemma S.3(a),

$$\begin{aligned}
\left(Lh^d\right)^{1/2}\left(\hat{\pi}^{\dagger}(N|x) - \hat{\pi}(N|x)\right) =& \frac{1}{\hat{\varphi}(x)}\left(Lh^d\right)^{1/2}\left(\hat{\pi}^{\dagger}(N,x) - \hat{\pi}(N,x)\right) \\
&- \frac{\hat{\pi}^{\dagger}(N,x)}{\left(\hat{\varphi}(x)\right)^2}\left(Lh^d\right)^{1/2}\left(\hat{\varphi}^{\dagger}(x) - \hat{\varphi}(x)\right) \\
&+ o\left(\left(Lh^d\right)^{1/2}\left(\hat{\varphi}^{\dagger}(x) - \hat{\varphi}(x)\right)\right) \\
=& O_p^{\dagger}(1).
\end{aligned}$$

The proof of part (c) is similar to that of part (b) and therefore omitted.

We prove part (d) next. The proof is similar to the proof of Lemma B.1 in Newey (1994). For fixed x in the interior of $\mathcal{X}$ and $N \in \mathcal{N}$, write

$$\hat{G}^*(b^*, N, x) = \hat{G}(b^*|N, x)\,\hat{p}(N, x)\,\hat{\pi}(N|x)\,\hat{\varphi}(x),$$

so that

$$\begin{aligned}\hat{G}^*(b, N, x) &= \frac{1}{NL}\sum_{l=1}^{L}\sum_{i=1}^{n_l} T_{il},\\ T_{il} &= \frac{1}{h^d} y_{il} 1(b_{il} \leq b)\, 1\{N_l = N\} \textstyle\prod_{k=1}^{d} K\left(\frac{x_{kl} - x_k}{h}\right),\end{aligned} \tag{S.27}$$

and let

$$\begin{aligned}\hat{G}^{*,\dagger}(b, N, x) &= \frac{1}{nL}\sum_{l=1}^{L}\sum_{i=1}^{n_l} T_{il}^{\dagger}(b),\\ T_{il}^{\dagger}(b) &= \frac{1}{h^d} y_{il}^{\dagger} 1\left(b_{il}^{\dagger} \leq b\right) 1\left\{N_l^{\dagger} = N\right\} \textstyle\prod_{k=1}^{d} K\left(\frac{x_{kl}^{\dagger} - x_k}{h}\right).\end{aligned}$$

Next, for the chosen values N and x, let

$$\begin{aligned}I &= \left[\underline{b}(N, x), \bar{b}(N, x)\right],\\ I &= \cup_{k=1}^{J_L} I_k,\end{aligned}$$

where the sub-intervals I_k's are non-overlapping and of length

$$s_L = \frac{\log L}{L}. \tag{S.28}$$

Denote as c_k the center of I_k. Note that I, I_k, c_k depend on N and x. Denote as $\kappa(b)$ the interval containing b, i.e. $b \in I_{\kappa(b)}$. Since

$$\hat{G}^*(b, N, x) = E^{\dagger} T_{il}^{\dagger}(b),$$

we can write

$$\begin{aligned}\hat{G}^{*,\dagger}(b, n, x) - \hat{G}^*(b, n, x) &= A_L^{\dagger}(b) - B_L^{\dagger}(b) + C_L^{\dagger}(b), \text{ where}\\ A_L^{\dagger}(b) &= \frac{1}{NL}\sum_{l=1}^{L}\sum_{i=1}^{N_l}\left(T_{il}^{\dagger}(b) - T_{il}^{\dagger}\left(c_{\kappa(b)}\right)\right),\end{aligned}$$

$$B_L^\dagger(b) = \frac{1}{NL}\sum_{l=1}^{L}\sum_{i=1}^{n_l}\left(E^\dagger T_{il}^\dagger(b) - E^\dagger T_{il}^\dagger\left(c_{\kappa(b)}\right)\right),$$
$$C_L^\dagger(b) = \frac{1}{NL}\sum_{l=1}^{L}\sum_{i=1}^{n_l}\left(T_{il}^\dagger\left(c_{\kappa(b)}\right) - E^\dagger T_{il}^\dagger\left(c_{\kappa(b)}\right)\right).$$

In the above decomposition, $A_L^\dagger(b)$ is the average of the deviations of $T_{il}^\dagger(b)$ from its value computed using the center of the interval containing b, and $B_L^\dagger(b)$ is the expected value under $P^\dagger$ of $A_L^\dagger(b)$. The terms $\sup_{b\in I}|A_L^\dagger(b)|$ and $\sup_{b\in I}|B_L^\dagger(b)|$ are small when s_L is small.

For $A_L^\dagger$ we have

$$\begin{aligned}
&\left|T_{il}^\dagger(b) - T_{il}^\dagger\left(c_{\kappa(b)}\right)\right| \\
&\leq h^{-d}(\sup K)^d y_{il}^\dagger 1\left(N_l^\dagger = N\right)\left|1\left(b_{il}^\dagger \leq b\right) - 1\left(b_{il}^\dagger \leq c_{\kappa(b)}\right)\right| \\
&\leq h^{-d}(\sup K)^d y_{il}^\dagger 1\left(N_l^\dagger = N\right)1\left(b_{il}^\dagger \in I_{\kappa(b)}\right),
\end{aligned} \tag{S.29}$$

where the second inequality holds because $|1(b_{il}^\dagger \leq b) - 1(b_{il}^\dagger \leq c_{\kappa(b)})|$ is equal to zero if $b_{il}^\dagger \notin I_{\kappa(b)}$ and is at most 1 if $b_{il}^\dagger \in I_{\kappa(b)}$. Thus,

$$\left|A_L^\dagger(b)\right| \leq h^{-d}(\sup K)^d \frac{1}{NL}\sum_{l=1}^{L}\sum_{i=1}^{N} y_{il}^\dagger 1\left(N_l^\dagger = N\right)1\left(b_{il}^\dagger \in I_{\kappa(b)}\right). \tag{S.30}$$

Next,

$$\begin{aligned}
&E^\dagger\left(\frac{1}{NL}\sum_{l=1}^{L}\sum_{i=1}^{N} y_{il}^\dagger 1\left(N_l^\dagger = N\right)1\left(b_{il}^\dagger \in I_k\right)\right) \\
&= E^\dagger\left(y_{il}^\dagger 1\left(N_l^\dagger = N\right)1\left(b_{il}^\dagger \in I_k\right)\right) \\
&= E^\dagger\left(y_{il}^\dagger 1\left(N_l^\dagger = N\right)P^\dagger\left(b_{il}^\dagger \in I_k | y_{il}^\dagger = 1, N_l^\dagger = N\right)\right) \\
&= E^\dagger\left(1\left(N_l^\dagger = N\right)p^\dagger(N)P^\dagger\left(b_{il}^\dagger \in I_k | y_{il}^\dagger = 1, N_l^\dagger = N\right)\right) \\
&= \pi^\dagger(N)p^\dagger(N)P^\dagger\left(b_{il}^\dagger \in I_k | y_{il}^\dagger = 1, N_l^\dagger = N\right).
\end{aligned}$$

Further,

$$E^\dagger\Bigg[\frac{1}{nL}\sum_{l=1}^{L}\sum_{i=1}^{N} y_{il}^\dagger 1\left(N_l^\dagger = N\right)1\left(b_{il}^\dagger \in I_k\right)$$

$$- P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N)\Big]^2 \leq$$

$$\leq \frac{P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N)}{NL}, \tag{S.31}$$

and by Lemma S.6,

$$\begin{aligned}
&\frac{1}{nL} \sum_{l=1}^{L} \sum_{i=1}^{N} y_{il}^{\dagger} 1\left(N_l^{\dagger} = N\right) 1\left(b_{il}^{\dagger} \in I_k\right) \\
&= P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N) \\
&\quad + O_p^{\dagger}\left(\frac{P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N)}{NL}\right)^{1/2} \\
&= P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N) \\
&\quad \times \left(1 + O_p^{\dagger}\left(\frac{1}{P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N) NL}\right)^{1/2}\right).
\end{aligned} \tag{S.32}$$

Now, by a similar argument,

$$\begin{aligned}
&P^{\dagger}\left(b_{il}^{\dagger} \in I_k | y_{il}^{\dagger} = 1, N_l^{\dagger} = N\right) \pi^{\dagger}(N) p^{\dagger}(N) \\
&= \frac{1}{NL} \sum_{l=1}^{L} \sum_{i=1}^{N} y_{il} 1(N_l = N) 1(b_{il} \in I_k) \\
&= P(b_{il} \in I_k | y_{il} = 1, N_l = N) \pi(N) p(N) \\
&\quad \times \left(1 + O_p\left(\frac{1}{P(b_{il} \in I_k | y_{il} = 1, N_l = N) \pi(N) p(N) NL}\right)^{1/2}\right) \\
&\leq \sup_{k=1,\ldots,J_L} P(b_{il} \in I_k | y_{il} = 1, N_l = N) \pi(N) p(N) \\
&\quad \times \left(1 + O_p\left(\frac{1}{\inf_{k=1,\ldots,J_L} P(b_{il} \in I_k | y_{il} = 1, N_l = N) \pi(N) p(N) NL}\right)^{1/2}\right).
\end{aligned} \tag{S.33}$$

Furthermore, for all I_k's

$$\left(\inf_{b\in I, x\in\mathcal{X}} g^*\left(b|N,x\right)\right) s_L \le P\left(b_{il}\in I_k|y_{il}=1, N_l=N\right) \le \left(\sup_{b\in I,x\in\mathcal{X}} g^*\left(b|N,x\right)\right) s_L. \tag{S.34}$$

Equations (S.30)-(S.34) together imply that

$$\begin{aligned}\left|\sup_{b\in I} A_L^{\dagger}\left(b\right)\right| &= O_p^{\dagger}\left(h^{-d} s_L \left(1+O_p\left(\frac{1}{s_L L}\right)^{1/2}\right)\right)\\ &= O_p^{\dagger}\left(\frac{\log L}{Lh^d}\right),\end{aligned} \tag{S.35}$$

where the last equality is by (S.28).

By (S.29), (S.33), and (S.34), for $B_L^{\dagger}\left(b\right)$ we have

$$\begin{aligned}\left|\sup_{b\in I} B_L^{\dagger}\left(b\right)\right| &\le \sup_{b\in I} E^{\dagger}\left|T_{il}^{\dagger}\left(b\right)-T_{il}^{\dagger}\left(c_{\kappa(b)}\right)\right|\\ &\le h^{-d}\left(\sup K\right)^d \pi^{\dagger}\left(n\right) \sup_{k=1,\ldots,J_L} P^{\dagger}\left(b_{il}^{\dagger}\in I_k|y_{il}^{\dagger}=1, N_l^{\dagger}=N\right)\\ &= O_p^{\dagger}\left(\frac{\log L}{Lh^d}\right).\end{aligned} \tag{S.36}$$

Note that $C_L^{\dagger}(b)$ depends on b only through c_k's, and therefore

$$\sup_{b\in I}|C_L^{\dagger}(b)| \le \max_{k=1,\ldots,J_L}|C_L^{\dagger}(c_k)|. \tag{S.37}$$

A Bonferroni inequality implies that for any $\Delta>0$,

$$\begin{aligned}&P^{\dagger}\left(\left(\frac{Lh^d}{\log L}\right)^{1/2}\max_{k=1,\ldots,J_L}|C_L^{\dagger}(b)|>\Delta\right)\le\\ &\qquad\le\sum_{k=1}^{J_L}P^{\dagger}\left(\left|\sum_{l=1}^{L}\sum_{i=1}^{N}\left(T_{il}^{\dagger}\left(c_k\right)-E^{\dagger}T_{il}^{\dagger}\left(c_k\right)\right)\right|>\Delta NL\left(\frac{\log L}{Lh^d}\right)^{1/2}\right).\end{aligned} \tag{S.38}$$

By (S.27), $|T_{il}^{\dagger}(c_k)|\le h^{-d}(\sup K)^d$ and

$$\left|T_{il}^{\dagger}\left(c_k\right)-E^{\dagger}T_{il}^{\dagger}\left(c_k\right)\right|\le 2(\sup K)^d h^{-d}.$$

Further, by (S.31)-(S.34), there is a constant $0 < D_1 < \infty$ such that

$$\begin{aligned} Var^{\dagger}\left(T_{il}^{\dagger}\left(c_k\right)\right) &\leq D_1 h^{-2d} s_L \left(1 + o_p\left(1\right)\right) \\ &= D_1 h^{-d}\left(\log L / \left(Lh^d\right)\right)\left(1 + o_p\left(1\right)\right). \end{aligned}$$

We therefore can apply Bernstein's inequality (Pollard, 1984, page 193) to obtain

$$\begin{aligned} & P^{\dagger}\left(\left|\sum_{l=1}^{L}\sum_{i=1}^{N}\left(T_{il}^{\dagger}\left(c_k\right) - E^{\dagger}T_{il}^{\dagger}\left(c_k\right)\right)\right| > \Delta N L \left(\frac{\log L}{Lh^d}\right)^{1/2}\right) \\ & \leq 2\exp\left(-\frac{1}{2}\frac{\Delta^2 N^2 L^2 \frac{\log L}{Lh^d}}{NLD_1 h^{-d}\left(1 + o_p\left(1\right)\right)\frac{\log L}{Lh^d} + \left(2/3\right)\Delta N (\sup K)^d h^{-d} L \left(\frac{\log L}{Lh^d}\right)^{1/2}}\right) \\ & = 2\exp\left(-\frac{1}{2}\frac{\Delta^2 N \left(\log L\right)^{1/2}\left(Lh^d\right)^{1/2}}{D_1\left(\log L / \left(Lh^d\right)\right)^{1/2}\left(1 + o_p\left(1\right)\right) + \left(2/3\right)\Delta (\sup K)^d}\right) \\ & = 2\exp\left(-\frac{\Delta N}{\left(4/3\right)\left(\sup K\right)^d + o_p\left(1\right)}\left(\log L\right)^{1/2}\left(Lh^d\right)^{1/2}\right), \end{aligned} \tag{S.39}$$

where the equality in the last line is due to $Lh^d / \log L \to \infty$. The inequalities in (S.37)-(S.39) together with (S.28) imply that there is a constant $0 < D_2 < \infty$ such that

$$\begin{aligned} & P^{\dagger}\left(\left(\frac{Lh^d}{\log L}\right)^{1/2}\sup_{b\in I}\left|C_L^{\dagger}(b)\right| > \Delta\right) \\ & \leq 2J_L \exp\left(-\frac{\Delta N}{\left(4/3\right)\left(\sup K\right)^d + o_p\left(1\right)}\left(\log L\right)^{1/2}\left(Lh^d\right)^{1/2}\right) \\ & \leq D_2 s_L^{-1}\exp\left(-\frac{\Delta N}{\left(4/3\right)\left(\sup K\right)^d + o_p\left(1\right)}\left(\log L\right)^{1/2}\left(Lh^d\right)^{1/2}\right) \\ & \leq D_2 \exp\left(\log L\left(1 - \frac{\Delta N}{\left(4/3\right)\left(\sup K\right)^d + o_p\left(1\right)}\left(\frac{Lh^d}{\log L}\right)^{1/2}\right)\right) \\ & = o_p\left(1\right), \end{aligned}$$

where the equality in the last line is by $Lh^d / \log L \to \infty$. By a similar argument as in the proof of Lemma S.6,

$$\sup_{b\in I}\left|C_L^{\dagger}(b)\right| = o_p^{\dagger}\left(\frac{Lh^d}{\log L}\right)^{-1/2}. \tag{S.40}$$

The result of part (d) follows from (S.35), (S.36), and (S.40).

The proof of part (e) is similar to that of Lemma S.3(e). First, by similar arguments as in the proof of Lemma S.3(e), one can show that $\underline{b}(N,x) \leq \hat{q}^{*,\dagger}(\varepsilon|B,x) \leq \hat{q}^{\dagger}(1-\varepsilon|n,x) \leq \bar{b}(N,x)$ with probability $P^{\dagger}$ approaching one (in probability), and that uniformly over $\tau \in [\varepsilon, 1-\varepsilon]$,

$$\hat{G}^{*,\dagger}\left(\hat{q}^{\dagger}(\tau|N,x)|N,x\right) = \tau + O_p^{\dagger}\left(Lh^d\right)^{-1}$$

Next,

$$\begin{aligned}
&G^*\left(\hat{q}^{*,\dagger}(\tau|N,x)|N,x\right) - \hat{G}^{*,\dagger}\left(\hat{q}^{*,\dagger}(\tau|N,x)|N,x\right) \\
&= G^*\left(\hat{q}^{*,\dagger}(\tau|N,x)|N,x\right) - \tau + O_p^{\dagger}\left(Lh^d\right)^{-1} \\
&= G^*\left(\hat{q}^{*,\dagger}(\tau|N,x)|N,x\right) - G^*\left(q^*\left(\tau|N,x\right)|N,x\right) + O_p^{\dagger}\left(Lh^d\right)^{-1} \\
&= g^*\left(\tilde{q}^{*,\dagger}\left(\tau|N,x\right)|N,x\right)\left(\hat{q}^{*,\dagger}(\tau|N,x) - q^*\left(\tau|N,x\right)\right) + O_p^{\dagger}\left(Lh^d\right)^{-1},
\end{aligned}$$

where $\tilde{q}^{\dagger}$ denotes the mean value, or

$$\begin{aligned}
&\hat{q}^{*,\dagger}(\tau|N,x) - q^*\left(\tau|N,x\right) \\
&= \frac{G^*\left(\hat{q}^{*,\dagger}(\tau|N,x)|N,x\right) - \hat{G}^{*,\dagger}\left(\hat{q}^{*,\dagger}(\tau|N,x)|N,x\right)}{g^*\left(\tilde{q}^{*,\dagger}\left(\tau|N,x\right)|N,x\right)} + O_p^{\dagger}\left(Lh^d\right)^{-1}.
\end{aligned}$$

By part (d) of this lemma, Lemma S.3(d) in MSX, and Lemma S.5(b),

$$\sup_{\tau\in[\varepsilon,1-\varepsilon]}\left|\hat{q}^{*,\dagger}(\tau|N,x) - q^*\left(\tau|N,x\right)\right| = O_p^{\dagger}\left(Lh^d\right)^{-1}. \tag{S.41}$$

As in the proof of Lemma S.5 and since $\hat{q}^*\left(\tau|N,x\right)$ is non-random under $P^{\dagger}$, for all $\epsilon > 0$ there is $\Delta_\epsilon > 0$ such that

$$\begin{aligned}
&P\left(P^{\dagger}\left(Lh^d\left|\hat{q}^*\left(\tau|N,x\right) - q^*\left(\tau|N,x\right)\right| > \Delta_\epsilon\right) > \epsilon\right) \\
&= P\left(1\left(Lh^d\left|\hat{q}^*\left(\tau|N,x\right) - q^*\left(\tau|N,x\right)\right| > \Delta_\epsilon\right) > \epsilon\right) \\
&= P\left(Lh^d\left|\hat{q}^*\left(\tau|N,x\right) - q^*\left(\tau|N,x\right)\right| > \Delta_\epsilon\right) \\
&< \epsilon,
\end{aligned} \tag{S.42}$$

where the inequality in the last line is by S.3(e). Furthermore, the last result holds uniformly in $\tau \in [\varepsilon, 1-\varepsilon]$. The result in part (e) of the lemma then follows by (S.41) and (S.42).[2]

The result in part (f) is implied by Proposition 3.2 in Chen and Lo (1997). The proof of parts (g) and (h) is similar to that of Lemma S.3(g) and (h). ∎

[2]Note that (S.42) establishes a trivial result that, if $\hat{\theta}_L = \theta + O_p\left(\lambda_L\right)$, then $\hat{\theta}_L = \theta + O_p^{\dagger}\left(\lambda_L\right)$ (recall that $\hat{\theta}_L$ is computed using the data in the original sample).

S.4.2 Main result

As is standard in the literature on statistical testing, see for example Chapter 21.4 in Gourieroux and Monfort (1995), in the SEM test we control the maximum asymptotic rejection probability under our composite null hypothesis, and replace the statistic $T^{SEM}(x)$ with its infeasible re-centered version $\bar{T}^{SEM}(x)$,

$$\begin{aligned}\bar{T}^{SEM}(x) &= \sup_{\tau \in \Upsilon} \sqrt{Lh^{d+1}} \sum_{N=\underline{N}}^{\bar{N}-1} \sum_{N'=N+1}^{\bar{N}} \frac{\left[\hat{\Delta}(\tau, N, N', x) - \Delta(\tau, N, N', x)\right]_-}{\hat{\sigma}(\tau, N, N', x)}, \text{ where} \\ \Delta(\tau, N, N', x) &= Q^*(\tau|N', x) - Q^*(\tau|N, x).\end{aligned}$$

Since under the null hypothesis, $\Delta(\tau, N, N', x) \leq 0$, it follows that $P(T^{SEM}(x) > u) \leq P(\bar{T}^{SEM}(x) > u)$ for all $u \in \mathbb{R}$. Therefore, when the test based on $\bar{T}^{SEM}(x)$ has asymptotic size α, the asymptotic size of the test $T^{SEM}(x)$ is less or equal to α. Thus, it suffices to show that

$$\sup_u \left| P(\bar{T}^{SEM}(x) > u) - P^\dagger(T_m^{\dagger,SEM}(x) > u) \right| \to_p 0, \tag{S.43}$$

where $P^\dagger$ denotes the bootstrap distribution conditional on the original data.

To show (S.43), we proceed as follows. First, the results in Lemma S.3 imply the following delta-method expansion:

$$\begin{aligned}&\sqrt{Lh^{d+1}} \frac{\left[\hat{\Delta}(\tau, N, N', x) - \Delta(\tau, N, N', x)\right]_-}{\hat{\sigma}(\tau, N, N', x)} = \frac{\kappa(\tau, N, N', x)}{\sigma(\tau, N, N', x)} \\ &\times \sqrt{Lh^{d+1}} \left[\hat{g}^*(q^*(\tau|N', x)) - g^*(q^*(\tau|N', x)) - \hat{g}^*(q^*(\tau|N, x)) + g^*(q^*(\tau|N, x))\right]_- \\ &+ o_p(1),\end{aligned} \tag{S.44}$$

where $\kappa(\tau, N, N', x)$ is determined by the term in the front of $\hat{g}^*(q^*(\tau|N, x)) - g^*(q^*(\tau|N, x))$ in (S.25), and the $o_p(1)$ term is uniform in $\tau \in [\tau_1(N, x), \tau_2(N, x)]$.

Given the results in Lemma S.7 for the bootstrap analogues of the original sample statistics and using the same arguments as in the proof of Proposition S.2 in Section S.3, and applying Lemma S.5, one can show that the bootstrap version of (S.44) holds as well:

$$\begin{aligned}&\sqrt{Lh^{d+1}} \frac{\left[\hat{\Delta}_m^\dagger(\tau, N, N', x) - \hat{\Delta}(\tau, N, N', x)\right]_-}{\hat{\sigma}^\dagger(\tau, N, N', x)} = \frac{\kappa(\tau, N, N', x)}{\sigma(\tau, N, N', x)} \\ &\times \sqrt{Lh^{d+1}} \left[\hat{g}_m^{*,\dagger}(q^*(\tau|N', x)) - \hat{g}^*(q^*(\tau|N', x)) - \hat{g}_m^{*,\dagger}(q^*(\tau|N, x)) + \hat{g}^*(q^*(\tau|N, x))\right]_-\end{aligned}$$

$$+\, o_p^{\dagger}(1), \tag{S.45}$$

where the $o_p^{\dagger}(1)$ term is again uniform in $\tau \in [\tau_1(N,x), \tau_2(N,x)]$.

Next, we use the uniform strong approximation for the bootstrap of Chen and Lo (1997). Provided that $Lh^{d+1} \to \infty$ and by Proposition 3.2 in Chen and Lo (1997), one can construct $\tilde{g}^*(\cdot|N,x)$ independent of the original data, such that $\tilde{g}^*(\cdot|N,x) =^d \hat{g}^*(\cdot|N,x)$, and for almost all sample paths,

$$\sup_{x\in\mathcal{X}^0} \sup_{b\in[b_1(N,x),b_2(N,x)]} \sqrt{Lh^{d+1}} \left|\hat{g}_m^{*,\dagger}(b|N,x) - \hat{g}^*(b|N,x) - (\tilde{g}^*(b|N,x) - g^*(b|N,x))\right| = O(\delta_L), \tag{S.46}$$

where $\delta_L \to 0$ is a sequence of constants.[3] In the above result, $[b_1(N,x), b_2(N,x)]$ and $\mathcal{X}^0$ are the inner compact subsets of the support of bids and x respectively.

Define

$$\frac{\left[\tilde{\Delta}(\tau,N,N',x) - \Delta(\tau,N,N',x)\right]_-}{\sigma(\tau,N,N',x)} = \frac{\kappa(\tau,N,N',x)}{\sigma(\tau,N,N',x)} \times \left[\tilde{g}^*(q^*(\tau|N',x)) - g^*(q^*(\tau|N',x)) - \tilde{g}^*(q^*(\tau|N,x)) + g^*(q^*(\tau|N,x))\right]_-,$$

and

$$\tilde{T}^{SEM}(x) = \sup_{\tau\in\Upsilon} \sqrt{Lh^{d+1}} \sum_{N=\underline{N}}^{\bar{N}-1} \sum_{N'=N+1}^{\bar{N}} \frac{\left[\tilde{\Delta}(\tau,N,N',x) - \Delta(\tau,N,N',x)\right]_-}{\sigma(\tau,N,N',x)}.$$

By the results in Lemma S.4, (S.44), and the Continuous Mapping Theorem,

$$\bar{T}^{SEM}(x) \to_d \mathcal{T}, \tag{S.47}$$

where $\mathcal{T}$ is a random variable with a continuous CDF. Since $\tilde{g}^*(\cdot|N,x) =^d \hat{g}^*(\cdot|N,x)$ by construction, we have as well that

$$\tilde{T}^{SEM}(x) \to_d \mathcal{T}. \tag{S.48}$$

Next, by (S.45) and (S.46),

$$T_m^{\dagger,SEM}(x) - \tilde{T}^{SEM}(x) = o_p^{\dagger}(1). \tag{S.49}$$

[3] See (14) in Chen and Lo (1997) for the precise definition of δ_L.

Since $\tilde{g}^*(\cdot|N,x)$ is independent of the original data by construction, $P^\dagger(\tilde{T}^{SEM}(x) \leq u) = P(\tilde{T}^{SEM}(x) \leq u)$ for all $u \in \mathbb{R}$. This, together with (S.49) and the fact that the CDF of $\mathcal{T}$ is continuous, implies that

$$P(\tilde{T}^{SEM}(x) > u) - P^\dagger(T_m^{\dagger,SEM}(x) > u) \to_p 0 \tag{S.50}$$

for all $u \in \mathbb{R}$.[4] Lastly, by (S.47), (S.48), and (S.50), we have that for all $u \in \mathbb{R}$,

$$P(\bar{T}^{SEM}(x) > u) - P^\dagger(T_m^{\dagger,SEM}(x) > u) \to_p 0. \tag{S.51}$$

The result in (S.43) now follows by the pointwise convergence in (S.51) and Pólya's Theorem (Shao and Tu, 1995, page 447).

S.5 Circumventing the curse of dimensionality: A single index approach

Consider a single index model

$$\begin{aligned} F(v,s|x) &= F(v,s|x'\beta), \\ r &= r(x'\beta). \end{aligned}$$

(In this section, we abuse the notation slightly by often keeping it the same for the distribution conditional on the single index, as well as for other relevant objects, r etc.) Here $x'\beta$ is a *single index* that captures the dependence of signals and valuations on covariates, and $\beta \in \mathbb{R}^d$ is a vector of coefficients, identifiable up to a common scale normalization. For simplicity, assume that the entry cost k does not vary with x.[5] Equation (18) implies that the signal cutoffs $\bar{s}$ are also functions of the single index $x'\beta$, say $\bar{s}(N, x'\beta)$, and therefore the distribution of active bidders' valuations $F^*(v|x'\beta)$ also depends on x only through $x'\beta$. Equation (17) implies that the bidding strategy $B(\cdot|N,x'\beta)$ also depends on x only through the single index $x'\beta$. Moreover, since the bidding strategy is monotone increasing, the quantiles of bids are equal to

$$\begin{aligned} q^*(\tau|N,x) &= B(Q^*(\tau|N,x'\beta)\,|N,x'\beta) \\ &\equiv \tilde{q}^*(\tau|N,x'\beta), \end{aligned}$$

[4] This can be shown similarly to Theorem 22.4 on page 349 in Davidson (1994).

[5] This assumption may be plausible in some applications. For example, Bajari, Hong, and Ryan (2010) in their empirical study of highway procurement auctions assume that entry costs do not depend on auction characteristics.

i.e. also depend on x only through the single index $x'\beta$. Consequently, we can estimate β from the bids data by any of the methods proposed in the literature on single index quantile regression.

In particular, we can use an average derivative estimator. For $u = x'\beta$, we have

$$\frac{\partial q^{*}\left(\tau|N,x\right)}{\partial x} = \beta \frac{\partial \tilde{q}^{*}\left(\tau|N,u\right)}{\partial u}, \tag{S.52}$$

and β can be estimated as an average quantile derivative. (Recall that β is only identifiable up to a scale normalization). Equation $(S.52)$ implies that β is *proportional* to the average derivative

$$\int \frac{\partial q^{*}\left(\tau|N,x\right)}{\partial x} w\left(x\right)\varphi\left(x\right)dx, \tag{S.53}$$

where $w\left(\cdot\right)$ is a nonnegative, smooth weighting function with compact support within $\mathcal{X}$. Since β is only identifiable up to a scalar multiple, we can normalize β by setting it *equal* to $(S.53)$. Taking into account this compact support assumption for $w\left(\cdot\right)$ and using integration by parts in $(S.53)$,

$$\begin{aligned} \beta &= -\int q^{*}\left(\tau|N,x\right)\frac{\partial\left[w\left(x\right)\varphi\left(x\right)\right]}{\partial x}dx \\ &= -\int q^{*}\left(\tau|N,x\right)\left(\varphi\left(x\right)\frac{\partial w\left(x\right)}{\partial x} + w\left(x\right)\frac{\partial\varphi\left(x\right)}{\partial x}\right)dx. \end{aligned} \tag{S.54}$$

Chaudhuri, Doksum, and Samarov (1997) propose an estimator of β based on a finite sample analogue to the average derivative $(S.54)$:

$$\hat{\beta} = -\frac{1}{L}\sum_{l=1}^{L}\frac{\hat{q}^{*}\left(\tau|N,x\right)}{\hat{\varphi}\left(x_l\right)}\left(\hat{\varphi}\left(x_l\right)\frac{\partial w\left(x_l\right)}{\partial x} + w\left(x_l\right)\frac{\partial\hat{\varphi}\left(x_l\right)}{\partial x}\right), \tag{S.55}$$

and provide conditions under which this estimator is root-L consistent, $\hat{\beta} = \beta + O_p\left(L^{-1/2}\right)$. Since the convergence of the estimator $\hat{\beta}$ is root-L, which is faster than nonparametric, the asymptotics of our estimators $\hat{Q}^{*}\left(\tau|N,x\right)$ etc. will remain unaffected if we use the single index $x'\hat{\beta}$ in place of x in the implementation of the nonparametric estimators described in the main text.

References

Anosov, D. V., S. K. Aranson, V. I. Arnold, I. U. Bronshtein, V. Z. Grines, and Y. S. Il'Yashenko (1997): *Ordinary Differential Equations and Smooth Dynamical Systems.* Springer Verlag, Berlin.

Bajari, P., H. Hong, and S. Ryan (2010): "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 78(5), 1529–1568.

Chaudhuri, P., K. Doksum, and A. Samarov (1997): "On Average Derivative Quantile Regression," *Annals of Statistics*, 25(2), 715–744.

Chen, K., and S. H. Lo (1997): "On a Mapping Approach to Investigating the Bootstrap Accuracy," *Probability Theory and Related Fields*, 107(2), 197–217.

Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press, New York.

Gourieroux, C., and A. Monfort (1995): *Statistics and Econometric Models*. Cambridge University Press, Cambridge.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–74.

Lehman, E. L., and J. P. Romano (2005): *Testing Statistical Hypotheses*. Springer, New York.

Newey, W. K. (1994): "Kernel Estimation of Partial Means and a General Variance Estimator," *Econometric Theory*, 10(2), 233–253.

Pagan, A., and A. Ullah (1999): *Nonparametric Econometrics*. Cambridge University Press, New York.

Pollard, D. (1984): *Convergence of Stochastic Processes*. Springer-Verlag, New York.

Shao, J., and D. Tu (1995): *The Jackknife and Bootstrap*. Springer-Verlag, New York.

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