SUPPLEMENT TO "BOOTSTRAP-ASSISTED INFERENCE FOR GENERALIZED GRENANDER-TYPE ESTIMATORS"

BY MATIAS D. CATTANEO[1,a], MICHAEL JANSSON[2,b] AND KENICHI NAGASAWA[3,c]

[1]*Department of Operations Research and Financial Engineering, Princeton University,* [a]*cattaneo@princeton.edu*

[2]*Department of Economics, University of California at Berkeley,* [b]*mjansson@econ.berkeley.edu*

[3]*Department of Economics, University of Warwick,* [c]*kenichi.nagasawa@warwick.ac.uk*

Contents

MSC2020 subject classifications: Primary 62G09, 62G20; secondary 62G07, 62G08.
Keywords and phrases: Monotone estimation, bootstrapping, robust inference.

SA.1. Proofs.

SA.1.1. *Proof of Lemma A.1.* Lemma 4.2 of van der Vaart and van der Laan (2006) implies that, on (l,u),

$$\mathsf{GCM}_{[l,u]}(\Gamma \circ \Phi^-) = \mathsf{GCM}_{[l,u]}(\mathsf{LSC}(\Gamma \circ \Phi^-)).$$

Since $\Phi(\mathsf{x}) \in (l,u)$, we therefore have

$$\theta(\mathsf{x}) = \partial_- \mathsf{GCM}_{[l,u]}(\Gamma \circ \Phi^-) \circ \Phi(\mathsf{x}) = \partial_- \mathsf{GCM}_{[l,u]}(\mathsf{LSC}(\Gamma \circ \Phi^-)) \circ \Phi(\mathsf{x}),$$

so Lemma 4.1 of van der Vaart and van der Laan (2006) implies that

$$\theta(\mathsf{x}) > t \qquad \Longleftrightarrow \qquad y^\star < \Phi(\mathsf{x}),$$

where

$$y^\star = \max \operatorname*{argmax}_{y \in [l,u]} \left\{ ty - \mathsf{LSC}(\Gamma \circ \Phi^-)(y) \right\}.$$

Suppose $y^\star \in \Phi(I)$. Then $y^\star = \Phi(x^\star)$, where

$$x^\star = \Phi^-(y^\star) \in \operatorname*{argmax}_{x \in \Phi^-([l,u])} \left\{ t\Phi(x) - \mathsf{LSC}_\Phi(\Gamma)(x) \right\}.$$

In particular,

$$\theta(\mathsf{x}) > t \qquad \Longleftrightarrow \qquad y^\star = \Phi(x^\star) < \Phi(\mathsf{x}) \qquad \Longleftrightarrow \qquad x^\star < \Phi^-(\Phi(\mathsf{x})),$$

where, in fact,

$$x^\star = \max \operatorname*{argmax}_{x \in \Phi^-([l,u])} \left\{ t\Phi(x) - \mathsf{LSC}_\Phi(\Gamma)(x) \right\},$$

because if

$$x^\star < x' \in \operatorname*{argmax}_{x \in \Phi^-([l,u])} \left\{ t\Phi(x) - \mathsf{LSC}_\Phi(\Gamma)(x) \right\},$$

then, contradicting the definition of $y^\star$, we have

$$y' = \Phi(x') \in \operatorname*{argmax}_{y \in [l,u]} \left\{ ty - \mathsf{LSC}(\Gamma \circ \Phi^-)(y) \right\} \qquad \text{and} \qquad y' > y^\star.$$

The proof can therefore be completed by showing that $y^\star \in \Phi(I)$.

If $\Phi(I) \supseteq [l,u]$, then there is nothing to show, so suppose $\Phi(I) \not\supseteq [l,u]$. If $y \in \Phi(I)^c \cap [l,u]$, then, since $\Phi(I) \cap [l,u]$ is closed and $l,u \in \Phi(I)$, we have $[y-\eta, y+\eta] \cap \Phi(I) = \emptyset$ for some $\eta > 0$ with $[y-\eta, y+\eta] \subset [l,u]$. Therefore, the function $\mathsf{LSC}(\Gamma \circ \Phi^-)$ in constant on the interval $[y-\eta/2, y+\eta/2]$, implying in particular that

$$y^\star = \max \operatorname*{argmax}_{y' \in [l,u]} \left\{ ty' - \mathsf{LSC}(\Gamma \circ \Phi^-)(y') \right\} \neq y. \quad \square$$

SA.1.2. *Proof of Lemma A.2.* We begin by adapting the arguments of Kim and Pollard (1990, pp. 196-198) to show that a maximizer of $\mathbb{G}(v)$ over $v \in \mathbb{R}$ exists and is unique with probability one. Let $\widetilde{\mathbb{G}}(v) = \mathbb{G}(v) - \mu(v)$ be the centered process and suppose that, for the same $c > 1/2$ as in Assumption A.2,

$$\mathbb{P}\left[\limsup_{|v| \to \infty} \frac{\widetilde{\mathbb{G}}(v)}{|v|^c} > \eta \right] = 0 \qquad \text{for any } \eta > 0. \tag{SA.1}$$

Then, with probability one, $\mathbb{G}(v) \to -\infty$ as $|v| \to \infty$, implying in turn that a maximizer of $\mathbb{G}(v)$ exists (because sample paths are continuous). Also, since

$$\mathbb{V}[\mathbb{G}(v) - \mathbb{G}(v')] = \mathcal{K}(v,v) + \mathcal{K}(v',v') - 2\mathcal{K}(v, v-v') = \mathcal{K}(v-v', v-v') > 0$$

for $v \neq v'$, Lemma 2.6 of Kim and Pollard (1990) implies that this maximizer is unique with probability one. In turn, (SA.1) follows from the Borel-Cantelli lemma because

$$\begin{aligned}\sum_{k=2}^{\infty} \mathbb{P}\left[\sup_{k-1 \leq |v| \leq k} \frac{\widetilde{\mathbb{G}}(v)}{|v|^c} > \eta\right] &\leq \sum_{k=2}^{\infty} \mathbb{P}\left[\sup_{|v| \leq k} \widetilde{\mathbb{G}}(v) > (k/2)^c \eta\right] \\ &= \sum_{k=2}^{\infty} \mathbb{P}\left[\sup_{|v| \leq 1} \widetilde{\mathbb{G}}(v) > k^{c-1/2} 2^{-c} \eta\right] \\ &\leq \frac{2^{4c/(2c-1)} \mathbb{E}\left[\sup_{|v| \leq 1} |\widetilde{\mathbb{G}}(v)|^{4c/(2c-1)}\right]}{\eta^{4c/(2c-1)}} \sum_{k=2}^{\infty} k^{-2} < \infty,\end{aligned}$$

where the equality uses the rescaling property $\mathcal{K}(v\tau, v'\tau) = \tau\mathcal{K}(v, v')$ and where the last inequality uses Jain and Marcus (1978, Corollary 4.7).

To show continuity of the function $x \mapsto \mathbb{P}[\operatorname{argmax}_{v \in \mathbb{R}} \mathbb{G}(v) \leq x]$, it suffices to show that $\mathbb{P}[\operatorname{argmax}_{v \in \mathbb{R}} \mathbb{G}(v) = x] = 0$ for every $x \in \mathbb{R}$. Fix $x \in \mathbb{R}$ and define $\tilde{Z}(x) = 0$ and

$$\tilde{Z}(v) = \frac{\mathbb{G}(v) - \mathbb{G}(x)}{\sqrt{\mathcal{K}(v-x, v-x)}}, \qquad v \neq x.$$

Then $\max_{v \in \mathbb{R}} \tilde{Z}(v) \geq 0$ and, for any set $\mathcal{V} \subset \mathbb{R}$,

$$\mathbb{P}\left[\operatorname*{argmax}_{v \in \mathbb{R}} \mathbb{G}(v) = x\right] = \mathbb{P}\left[\max_{v \in \mathbb{R}} \tilde{Z}(v) \leq 0\right] \leq \mathbb{P}\left[\max_{v \in \mathcal{V}} \tilde{Z}(v) \leq 0\right]. \tag{SA.2}$$

In the sequel, we show that the majorant in (SA.2) can be made arbitrarily small by choice of $\mathcal{V}$. In particular, for $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$, we construct $v_{1,N}^{\varepsilon}, \ldots, v_{N,N}^{\varepsilon}$ such that

$$\mathbb{E}\left[\tilde{Z}(v_{i,N}^{\varepsilon})\right] \geq -\varepsilon \quad \text{for every } 1 \leq i \leq N, \tag{SA.3}$$

and

$$\left|\mathbb{C}\text{ov}\left(\tilde{Z}(v_{i,N}^{\varepsilon}), \tilde{Z}(v_{j,N}^{\varepsilon})\right)\right| \leq \varepsilon \quad \text{for every } 1 \leq i < j \leq N. \tag{SA.4}$$

Defining $\mathcal{V}_N^{\varepsilon} = \{v_{1,N}^{\varepsilon}, \ldots, v_{N,N}^{\varepsilon}\}$, we therefore have

$$\mathbb{P}\left[\max_{v \in \mathbb{R}} \tilde{Z}(v) \leq 0\right] \leq \liminf_{\varepsilon \downarrow 0} \mathbb{P}\left[\max_{v \in \mathcal{V}_N^{\varepsilon}} \tilde{Z}(v) \leq 0\right] \leq \left|\int_{-\infty}^{0} \frac{\exp(-x^2/2)}{\sqrt{2\pi}} dx\right|^N = 2^{-N},$$

where the second inequality uses the fact that convergence of means and covariances of normal random vectors implies convergence in distribution. Since N is arbitrary, the left-hand side in the preceding display is zero. Letting $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$ be given, the proof can therefore be completed by exhibiting $\{v_{i,N}^{\varepsilon}\}_{i=1}^{N}$ satisfying (SA.3) and (SA.4).

Because $\mathcal{K}(\tau, \tau) = \tau\mathcal{K}(1,1)$ and $\lim_{\tau \downarrow 0}[\mu(x+\tau) - \mu(x)]/\sqrt{\tau} = 0$, there exists $\bar{\tau}_{\varepsilon}' > 0$ such that

$$\mathbb{E}\left[\tilde{Z}(x+\tau)\right] = \frac{\mu(x+\tau) - \mu(x)}{\sqrt{\mathcal{K}(\tau,\tau)}} = \frac{[\mu(x+\tau) - \mu(x)]/\sqrt{\tau}}{\sqrt{\mathcal{K}(1,1)}} \geq -\varepsilon$$

for every $\tau \in (0, \bar{\tau}_\varepsilon^{'})$. Also, because $\mathcal{K}(\tau_i, \tau_j) = \tau_i \mathcal{K}(1, \tau_j/\tau_i)$ and $\lim_{\tau \downarrow 0} \mathcal{K}(1,\tau)/\sqrt{\tau} = 0$, there exists $\bar{\tau}_\varepsilon^{''} > 0$ such that

$$\mathbb{C}\mathrm{ov}\left(\tilde{Z}(x+\tau_i), \tilde{Z}(x+\tau_j)\right) = \frac{\mathcal{K}(\tau_i,\tau_j)}{\sqrt{\mathcal{K}(\tau_i,\tau_i)\mathcal{K}(\tau_j,\tau_j)}} = \frac{\mathcal{K}(1,\tau_j/\tau_i)/\sqrt{\tau_j/\tau_i}}{\mathcal{K}(1,1)} \in [-\varepsilon, \varepsilon]$$

for all $\tau_i, \tau_j > 0$ with $\tau_j/\tau_i < \bar{\tau}_\varepsilon^{''}$. Now, if $v_{i,N}^\varepsilon = x + \bar{\tau}_\varepsilon^i/2$ for some $\bar{\tau}_\varepsilon < \min\{\bar{\tau}_\varepsilon^{'}, \bar{\tau}_\varepsilon^{''}, 1\}$, then $\{v_{i,N}^\varepsilon\}_{i=1}^N$ satisfies (SA.3) and (SA.4). □

SA.1.3. *Technical Lemmas.* In preparation for the proof of Theorem 1, this section presents six technical lemmas. The first lemma is a switching lemma, which will be used when characterizing the limiting distributions obtained in Theorem 1.

LEMMA SA-1. *Let $\Gamma : \mathbb{R} \to \mathbb{R}$ be a lower semi-continuous function that is bounded from below and satisfies $\lim_{|v|\to\infty} \Gamma(v)/|v| = \infty$. Then, for any $\mathsf{x}, t \in \mathbb{R}$,*

$$\partial_{-}\mathsf{GCM}_{\mathbb{R}}(\Gamma)(\mathsf{x}) > t \qquad \Longleftrightarrow \qquad \max \operatorname*{argmax}_{v\in\mathbb{R}} \{vt - \Gamma(v)\} < \mathsf{x}.$$

PROOF. Because $\lim_{|v|\to\infty} \Gamma(v)/|v| = \infty$, there exists a $K > |\mathsf{x}|$ such that if $|v| \geq K$, then $vt - \Gamma(v) < -\Gamma(0)$, implying in particular that

$$\operatorname*{argmax}_{v\in\mathbb{R}} \{vt - \Gamma(v)\} = \operatorname*{argmax}_{v\in[-c,c]} \{vt - \Gamma(v)\}$$

for every $c \geq K$. Also, by Lemma A.1. of Sen, Banerjee and Woodroofe (2010) there exists a $c > K$ such that $\mathsf{GCM}_{\mathbb{R}}(\Gamma) = \mathsf{GCM}_{[-c,c]}(\Gamma)$ on $[-K, K]$, implying in particular that

$$\partial_{-}\mathsf{GCM}_{\mathbb{R}}(\Gamma)(\mathsf{x}) = \partial_{-}\mathsf{GCM}_{[-c,c]}(\Gamma)(\mathsf{x}).$$

For any such c, the conclusion of the lemma is equivalent to the statement

$$\partial_{-}\mathsf{GCM}_{[-c,c]}(\Gamma)(\mathsf{x}) > t \qquad \Longleftrightarrow \qquad \max \operatorname*{argmax}_{v\in[-c,c]} \{vt - \Gamma(v)\} < \mathsf{x},$$

whose validity follows from Lemma 4.1 of van der Vaart and van der Laan (2006). □

The proof of Theorem 1 furthermore employs various approximations to functionals of the form $\mathsf{LSC}_\Phi(f)$. The approximations in question are obtained using Lemmas SA-2, SA-3, SA-5, and SA-6. In all cases, the approximations are based on the representation

$$\mathsf{LSC}_\Phi(f)(x) = \liminf_{\epsilon\downarrow 0} \inf_{x'\in\mathcal{X}_\Phi^\epsilon(x)} f(x'), \tag{SA.5}$$

where $\mathcal{X}_\Phi^\epsilon(x) = (\Phi^-(\Phi(x)-\epsilon), \Phi^-(\Phi(x))] \cup (\Phi^-(\Phi(x)+), \Phi^-(\Phi(x)+\epsilon))$.

The following lemma uses (SA.5) and the special structure of Γ_0 to obtain a simple "global" bound on the error of the approximation $\mathsf{LSC}_\Phi(\Gamma_0) \approx \Gamma_0$.

LEMMA SA-2. *Suppose Assumption A holds and suppose Φ is non-decreasing and right-continuous on I. Then, for every $x \in I$,*

$$\left|\mathsf{LSC}_\Phi(\Gamma_0)(x) - \Gamma_0(x)\right| \leq 2\left(\sup_{x'\in I} |\theta_0(x')|\right)\left(\sup_{x'\in I} |\Phi(x') - \Phi_0(x')|\right).$$

PROOF. By (SA.5) and continuity of Γ_0,

$$\mathsf{LSC}_\Phi(\Gamma_0)(x) = \min\left[\Gamma_0(\Phi^-(\Phi(x))), \Gamma_0(\Phi^-(\Phi(x)+))\right],$$

while, by Assumption A,

$$|\Gamma_0(x') - \Gamma_0(x)| \leq \left(\sup_{x'' \in I} |\theta_0(x'')|\right) |\Phi_0(x') - \Phi_0(x)|.$$

Now, using $\Phi \circ \Phi^- \circ \Phi = \Phi$,

$$\begin{aligned} |\Phi_0(\Phi^-(\Phi(x))) - \Phi_0(x)| &= |\Phi_0(\Phi^-(\Phi(x))) - \Phi(\Phi^-(\Phi(x))) + \Phi(x) - \Phi_0(x)| \\ &\leq 2\left(\sup_{x' \in I} |\Phi(x') - \Phi_0(x')|\right). \end{aligned}$$

Also,

$$|\Phi_0(\Phi^-(\Phi(x)+)) - \Phi_0(x)| \leq 2\left(\sup_{x' \in I} |\Phi(x') - \Phi_0(x')|\right)$$

because, for every $\eta > 0$,

$$\begin{aligned} 0 &\leq \Phi_0(\Phi^-(\Phi(x)+)) - \Phi_0(x) \leq \Phi_0(\Phi^-(\Phi(x)+\eta)) - \Phi_0(x) \\ &\leq \Phi_0\left(\Phi^-\left(\Phi_0(x) + \sup_{x' \in I} |\Phi(x') - \Phi_0(x')| + \eta\right)\right) - \Phi_0(x) \\ &\leq \Phi_0\left(\Phi_0^-\left(\Phi_0(x) + 2\sup_{x' \in I} |\Phi(x') - \Phi_0(x')| + \eta\right)\right) - \Phi_0(x) \\ &\leq 2\sup_{x' \in I} |\Phi(x') - \Phi_0(x')| + \eta, \end{aligned}$$

where the last inequality uses continuity of Φ_0. $\square$

A simple "global" bound on the error of the approximation $\mathsf{LSC}_\Phi(f) \approx f$ is available also in the important special case where f is proportional to Φ.

LEMMA SA-3. *Suppose Assumption A holds and suppose Φ is non-decreasing and right-continuous on I. Then, for every $x \in I$ and every $\theta \in \mathbb{R}$,*

$$|\mathsf{LSC}_\Phi(\theta\Phi)(x) - \theta\Phi(x)| \leq |\theta| \left(\sup_{x' \in I} |\Phi(x') - \Phi(x'-)|\right).$$

PROOF. First, if $\theta > 0$, then the result follows from the fact that, by (SA.5),

$$\mathsf{LSC}_\Phi(\theta\Phi)(x) = \theta\Phi(\Phi^-(\Phi(x))-) \leq \theta\Phi(\Phi^-(\Phi(x))) = \theta\Phi(x),$$

where

$$\Phi(x) = \Phi(\Phi^-(\Phi(x))) \leq \Phi(\Phi^-(\Phi(x))-) + \sup_{x' \in I} |\Phi(x') - \Phi(x'-)|.$$

Next, if $\theta < 0$, then the result follows from the fact that, by (SA.5),

$$\mathsf{LSC}_\Phi(\theta\Phi)(x) = \theta\Phi(\Phi^-(\Phi(x)+)) \leq \theta\Phi(\Phi^-(\Phi(x))) = \theta\Phi(x),$$

where

$$\begin{aligned} \Phi(\Phi^-(\Phi(x)+)) &\leq \liminf_{\eta \downarrow 0} \Phi(\Phi^-(\Phi(x)+\eta)-) + \sup_{x' \in I} |\Phi(x') - \Phi(x'-)| \\ &\leq \Phi(x) + \sup_{x' \in I} |\Phi(x') - \Phi(x'-)|. \end{aligned}$$

$\square$

Next, we give a "local" approximation to $\Phi^- \circ \Phi$. That approximation will later be used in combination with (SA.5) to obtain "local" approximations to $\mathsf{LSC}_\Phi(f)$, but the approximation is also useful in its own right and we therefore state it as a separate lemma.

LEMMA SA-4. *Suppose Assumption A holds and suppose Φ is non-decreasing and right-continuous on I. Also, suppose*

$$\frac{2\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|}{\inf_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}\partial\Phi_0(x')}<\epsilon$$

for some $\delta,\epsilon>0$ with $I_{\mathsf{x}}^{\delta+\epsilon}\subseteq I$. Then, for every $x\in I_{\mathsf{x}}^{\delta}$,

$$|\Phi^-(\Phi(x))-x|<\epsilon \qquad \textit{and} \qquad |\Phi^-(\Phi(x)+)-x|<\epsilon.$$

PROOF. First, suppose $|\Phi^-(\Phi(x))-x|\geq\epsilon$. Then $\Phi^-(\Phi(x))\leq x-\epsilon$, implying in particular that $\Phi(x-\epsilon)=\Phi(x)$ and therefore also

$$[\Phi(x-\epsilon)-\Phi_0(x-\epsilon)]-[\Phi(x)-\Phi_0(x)]=\Phi_0(x)-\Phi_0(x-\epsilon).$$

Now, if $x\in I_{\mathsf{x}}^{\delta}$, then $[x-\epsilon,x]\subseteq I_{\mathsf{x}}^{\delta+\epsilon}$, so

$$\Phi_0(x)-\Phi_0(x-\epsilon)\geq\epsilon\left(\inf_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}\partial\Phi_0(x')\right)>2\left(\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|\right),$$

whereas

$$\Big|[\Phi(x-\epsilon)-\Phi_0(x-\epsilon)]-[\Phi(x)-\Phi_0(x)]\Big|\leq 2\left(\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|\right).$$

In other words, $x\notin I_{\mathsf{x}}^{\delta}$.

Next, suppose $|\Phi^-(\Phi(x)+)-x|\geq\epsilon$. Then, for every $\eta,\eta'>0$, $\Phi^-(\Phi(x)+\eta)\geq x+\epsilon$ and therefore

$$\Phi(x+\epsilon-\eta')-\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|-\eta<\Phi(x)-\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|,$$

where, for $x\in I_{\mathsf{x}}^{\delta}$,

$$\Phi(x)-\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|\leq\Phi_0(x),$$

whereas

$$\begin{aligned}
&\liminf_{\eta,\eta'\downarrow 0}\left[\Phi(x+\epsilon-\eta')-\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|-\eta\right]\\
&\geq\liminf_{\eta'\downarrow 0}\left[\Phi_0(x+\epsilon-\eta')-2\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|\right]\\
&\geq\Phi_0(x)+\left(\inf_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}\partial\Phi_0(x')\right)\liminf_{\eta'\downarrow 0}\left[\epsilon-\frac{2\sup_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}|\Phi(x')-\Phi_0(x')|}{\inf_{x'\in I_{\mathsf{x}}^{\delta+\epsilon}}\partial\Phi_0x')}-\eta'\right]>\Phi_0(x).
\end{aligned}$$

In other words, $x\notin I_{\mathsf{x}}^{\delta}$. $\square$

Next, we obtain two "local" approximations to $\mathsf{LSC}_\Phi(f)$. The first of these is a generic approximation obtained by simply combining (SA.5) and Lemma SA-4, but for later reference we state the result as a separate lemma.

LEMMA SA-5. *Suppose the assumptions of Lemma SA-4 hold. Then, for every $x \in I_{\mathsf{x}}^\delta$ and every $f : I \to \mathbb{R}$,*

$$\inf_{|x'-x|\le\epsilon} f(x') \le \mathsf{LSC}_\Phi(f)(x) \le \sup_{|x'-x|\le\epsilon} f(x')$$

and

$$|\mathsf{LSC}_\Phi(f)(x) - f(x)| \le \sup_{|x'-x|\le\epsilon} |f(x') - f(x)|.$$

The final lemma is concerned with the special case where f is proportional to Φ. In that case, the following "local" analog of Lemma SA-3 shows that the bound(s) obtained in Lemma SA-5 can be improved under mild conditions on Φ.

LEMMA SA-6. *Suppose the assumptions of Lemma SA-4 hold. Then, for every $x \in I_{\mathsf{x}}^\delta$ and every $\theta \in \mathbb{R}$,*

$$|\mathsf{LSC}_\Phi\big(\theta\Phi\big)(x) - \theta\Phi(x)| \le |\theta| \left(\sup_{x' \in I_{\mathsf{x}}^{\delta+\epsilon}} |\Phi(x') - \Phi(x'-)| \right).$$

PROOF. First, if $\theta > 0$, then the result follows from the fact that, by (SA.5),

$$\mathsf{LSC}_\Phi\big(\theta\Phi\big)(x) = \theta\Phi(\Phi^-(\Phi(x))-) \le \theta\Phi(\Phi^-(\Phi(x))) = \theta\Phi(x),$$

where, if $x \in I_{\mathsf{x}}^\delta$, then $\Phi^-(\Phi(x)) \in I_{\mathsf{x}}^{\delta+\epsilon}$ by Lemma SA-4, and therefore

$$\Phi(x) = \Phi(\Phi^-(\Phi(x))) \le \Phi(\Phi^-(\Phi(x))-) + \sup_{x' \in I_{\mathsf{x}}^{\delta+\epsilon}} |\Phi(x') - \Phi(x'-)|.$$

Next, if $\theta < 0$, then the result follows from the fact that, by (SA.5),

$$\mathsf{LSC}_\Phi\big(\theta\Phi\big)(x) = \theta\Phi(\Phi^-(\Phi(x)+)) \le \theta\Phi(\Phi^-(\Phi(x))) = \theta\Phi(x),$$

where, if $x \in I_{\mathsf{x}}^\delta$, then $\Phi^-(\Phi(x)+) \in I_{\mathsf{x}}^{\delta+\epsilon}$ by Lemma SA-4, and therefore

$$\begin{aligned}\Phi(\Phi^-(\Phi(x)+)) &\le \liminf_{\eta\downarrow 0} \Phi(\Phi^-(\Phi(x)+\eta)-) + \sup_{x' \in I_{\mathsf{x}}^{\delta+\epsilon}} |\Phi(x') - \Phi(x'-)| \\ &\le \Phi(x) + \sup_{x' \in I_{\mathsf{x}}^{\delta+\epsilon}} |\Phi(x') - \Phi(x'-)|.\end{aligned}$$

□

SA.1.4. *Proof of Theorem 1.*

Proof of (2). Let $t \in \mathbb{R}$ be given. By Lemma A.1 and change of variables,

$$\begin{aligned}&\mathbb{P}\left[r_n(\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})) > t\right] \\ &= \mathbb{P}\left[\max \underset{x \in \widehat{\Phi}_n^-([0,\widehat{u}_n])}{\operatorname{argmax}} \left\{[\theta_0(\mathsf{x}) + t r_n^{-1}]\widehat{\Phi}_n(x) - \mathsf{LSC}_{\widehat{\Phi}_n}\big(\widehat{\Gamma}_n\big)(x)\right\} < \widehat{\Phi}_n^- \circ \widehat{\Phi}_n(\mathsf{x})\right] \\ &= \mathbb{P}\left[\max \underset{v \in \widehat{V}_{\mathsf{x},n}^{\mathsf{q}}}{\operatorname{argmin}} \widehat{H}_{\mathsf{x},n}^{\mathsf{q}}(v;t) < \widehat{Z}_{\mathsf{x},n}^{\mathsf{q}}\right],\end{aligned}$$

where

$$\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n} = \left\{a_n(x - \mathsf{x}) : x \in \widehat{\Phi}^{-}_n([0, \widehat{u}_n])\right\},$$

$$\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t) = \mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n} + M^{\mathfrak{q}}_{\mathsf{x},n} + r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v) - [r_n\theta_0(\mathsf{x}) + t]\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)$$

and

$$\widehat{Z}^{\mathfrak{q}}_{\mathsf{x},n} = a_n\left[\widehat{\Phi}^{-}_n \circ \widehat{\Phi}_n(\mathsf{x}) - \mathsf{x}\right],$$

with

$$\begin{aligned}
\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v) &= \sqrt{na_n}\left[\widehat{\Gamma}_n(\mathsf{x} + va_n^{-1}) - \widehat{\Gamma}_n(\mathsf{x}) - \Gamma_0(\mathsf{x} + va_n^{-1}) + \Gamma_0(\mathsf{x})\right] \\
&\qquad - \theta_0(\mathsf{x})\sqrt{na_n}\left[\widehat{\Phi}_n(\mathsf{x} + va_n^{-1}) - \widehat{\Phi}_n(\mathsf{x}) - \Phi_0(\mathsf{x} + va_n^{-1}) + \Phi_0(\mathsf{x})\right], \\
M^{\mathfrak{q}}_{\mathsf{x},n}(v) &= \sqrt{na_n}\left[\Gamma_0(\mathsf{x} + va_n^{-1}) - \Gamma_0(\mathsf{x})\right] - \theta_0(\mathsf{x})\sqrt{na_n}\left[\Phi_0(\mathsf{x} + va_n^{-1}) - \Phi_0(\mathsf{x})\right], \\
\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v) &= a_n\left[\widehat{\Phi}_n(\mathsf{x} + va_n^{-1}) - \widehat{\Phi}_n(\mathsf{x})\right].
\end{aligned}$$

By (B4) and Lemma SA-4, $\widehat{Z}^{\mathfrak{q}}_{\mathsf{x},n} = o_{\mathbb{P}}(1)$. Suppose also that

$$\max \operatorname*{argmin}_{v \in \widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}} \widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t) \rightsquigarrow \operatorname*{argmin}_{v \in \mathbb{R}} \mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(v;t), \tag{SA.6}$$

where $\mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(v;t) = \mathcal{G}_{\mathsf{x}}(v) + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}(v) - t\partial\Phi_0(\mathsf{x})v$. Then

$$\begin{aligned}
\mathbb{P}\left[r_n(\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})) > t\right] &= \mathbb{P}\left[\max \operatorname*{argmin}_{v \in \widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}} \widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t) < \widehat{Z}^{\mathfrak{q}}_{\mathsf{x},n}\right] \\
&\to \mathbb{P}\left[\operatorname*{argmin}_{v \in \mathbb{R}} \mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(v;t) < 0\right] \\
&= \mathbb{P}\left[\frac{1}{\partial\Phi_0(\mathsf{x})}\partial_{-}\mathsf{GCM}_{\mathbb{R}}(\mathcal{G}_{\mathsf{x}} + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}(0) > t\right],
\end{aligned}$$

where the second line uses Lemma A.2 and where the last equality uses Lemma SA-1. The proof of (2) can therefore be completed by showing (SA.6).

We shall do so by means of the argmax continuous mapping theorem of Cox (2022). To be specific, using that theorem it can be shown that (SA.6) holds if

$$\widehat{v}_n(t) = \max \operatorname*{argmin}_{v \in \widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}} \widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t) = O_{\mathbb{P}}(1) \tag{SA.7}$$

and if

$$\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(\cdot;t) \rightsquigarrow \mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(\cdot;t). \tag{SA.8}$$

We begin by showing (SA.8). First, by (B1), $\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n} \rightsquigarrow \mathcal{G}_{\mathsf{x}}$. Also, by (A2) and (A3), as $u \to 0$,

$$\begin{aligned}
\frac{\Gamma_0(\mathsf{x}+u) - \theta_0(\mathsf{x})\Phi_0(\mathsf{x}+u) - \Gamma_0(\mathsf{x}) + \theta_0(\mathsf{x})\Phi_0(\mathsf{x})}{u^{\mathfrak{q}+1}} &\to \lim_{u \to 0} \frac{[\theta_0(\mathsf{x}+u) - \theta_0(\mathsf{x})]\,\partial\Phi_0(\mathsf{x}+u)}{(\mathfrak{q}+1)u^{\mathfrak{q}}} \\
&= \frac{\partial^{\mathfrak{q}}\theta_0(\mathsf{x})\partial\Phi_0(\mathsf{x})}{(\mathfrak{q}+1)!},
\end{aligned}$$

where the first equality uses L'Hôpital's rule and

$$\partial\left[\Gamma_0(\mathsf{x}+u)-\theta_0(\mathsf{x})\Phi_0(\mathsf{x}+u)-\Gamma_0(\mathsf{x})+\theta_0(\mathsf{x})\Phi_0(\mathsf{x})\right]=\left[\theta_0(\mathsf{x}+u)-\theta_0(\mathsf{x})\right]\partial\Phi_0(\mathsf{x}+u).$$

As a consequence, $M_{\mathsf{x},n}^{\mathtt{q}} \rightsquigarrow \mathcal{M}_{\mathsf{x}}^{\mathtt{q}}$. Moreover, $\widehat{L}_{\mathsf{x},n}^{\mathtt{q}} - L_{\mathsf{x},n}^{\mathtt{q}} \rightsquigarrow 0$ by (B4) and $L_{\mathsf{x},n}^{\mathtt{q}} \rightsquigarrow \mathcal{L}_{\mathsf{x}}$ by (A3), where $L_{\mathsf{x},n}^{\mathtt{q}}(v) = a_n\left[\Phi_0(\mathsf{x}+va_n^{-1})-\Phi_0(\mathsf{x})\right]$. In particular, $\widehat{L}_{\mathsf{x},n}^{\mathtt{q}} \rightsquigarrow \mathcal{L}_{\mathsf{x}}$ and therefore

$$(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}, M_{\mathsf{x},n}^{\mathtt{q}}, \widehat{L}_{\mathsf{x},n}^{\mathtt{q}}) \rightsquigarrow (\mathcal{G}_{\mathsf{x}}, \mathcal{M}_{\mathsf{x}}^{\mathtt{q}}, \mathcal{L}_{\mathsf{x}}).$$

Because $\widehat{G}_{\mathsf{x},n}^{\mathtt{q}} + M_{\mathsf{x},n}^{\mathtt{q}}$ is asymptotically equicontinuous,

$$\mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}}\left(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}+M_{\mathsf{x},n}^{\mathtt{q}}\right)-\left(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}+M_{\mathsf{x},n}^{\mathtt{q}}\right)\rightsquigarrow 0$$

by (B4) and Lemma SA-5. Also, by (B8) and Lemma SA-6,

$$\mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}}\left(r_n\theta_0(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}\right)-\left(r_n\theta_0(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}\right)\rightsquigarrow 0.$$

The result (SA.8) follows from the three preceding displays and the fact that, on $\widehat{V}_{\mathsf{x},n}^{\mathtt{q}}$,

$$\begin{aligned}0 &\geq \mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}}\left(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}+M_{\mathsf{x},n}^{\mathtt{q}}+r_n\theta_0(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}\right)-\left(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}+M_{\mathsf{x},n}^{\mathtt{q}}+r_n\theta_0(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}\right)\\ &\geq \mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}}\left(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}+M_{\mathsf{x},n}^{\mathtt{q}}\right)-\left(\widehat{G}_{\mathsf{x},n}^{\mathtt{q}}+M_{\mathsf{x},n}^{\mathtt{q}}\right)+\mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}}\left(r_n\theta_0(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}\right)-\left(r_n\theta_0(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathtt{q}}\right).\end{aligned}$$

Next, to show (SA.7), we first define $\theta_n(\mathsf{x};t)=\theta_0(\mathsf{x})+tr_n^{-1}$ and note that

$$\widehat{v}_n(t)=a_n\left[\max\ \underset{x\in\widehat{\Phi}_n^{-}([0,\widehat{u}_n])}{\operatorname{argmax}}\left\{\theta_n(\mathsf{x};t)\widehat{\Phi}_n(x)-\mathsf{LSC}_{\widehat{\Phi}_n}\left(\widehat{\Gamma}_n\right)(x)\right\}-\mathsf{x}\right].$$

Now, if $|\widehat{v}_n(t)|>a_n\delta>0$, then

$$\sup_{x\notin I_{\mathsf{x}}^{\delta}}\left\{\theta_n(\mathsf{x};t)\widehat{\Phi}_n(x)-\mathsf{LSC}_{\widehat{\Phi}_n}\left(\widehat{\Gamma}_n\right)(x)\right\}\geq\theta_n(\mathsf{x};t)\widehat{\Phi}_n(\mathsf{x})-\mathsf{LSC}_{\widehat{\Phi}_n}\left(\widehat{\Gamma}_n\right)(\mathsf{x}),$$

where $|\theta_n(\mathsf{x};t)-\theta_0(\mathsf{x})|=O(r_n^{-1})=o(1)$, and, by (B4),

$$\sup_{x\in I}\left|\widehat{\Phi}_n(x)-\Phi_0(x)\right|=o_{\mathbb{P}}(1).$$

Also, using (B3), (B4), and Lemma SA-2,

$$\begin{aligned}&\sup_{x\in I}\left|\mathsf{LSC}_{\widehat{\Phi}_n}\left(\widehat{\Gamma}_n\right)(x)-\Gamma_0(x)\right|\\ &\leq\sup_{x\in I}\left|\mathsf{LSC}_{\widehat{\Phi}_n}\left(\widehat{\Gamma}_n\right)(x)-\mathsf{LSC}_{\widehat{\Phi}_n}\left(\Gamma_0\right)(x)\right|+\sup_{x\in I}\left|\mathsf{LSC}_{\widehat{\Phi}_n}\left(\Gamma_0\right)(x)-\Gamma_0(x)\right|\\ &\leq\sup_{x\in I}\left|\widehat{\Gamma}_n(x)-\Gamma_0(x)\right|+2\left(\sup_{x\in I}|\theta_0(x)|\right)\left(\sup_{x\in I}|\widehat{\Phi}_n(x)-\Phi_0(x)|\right)=o_{\mathbb{P}}(1).\end{aligned}$$

As a consequence, $\widehat{v}_n(t)=o_{\mathbb{P}}(a_n)$: For any $\delta>0$,

$$\begin{aligned}&\mathbb{P}\left[|\widehat{v}_n(t)|>a_n\delta\right]\\ &\leq\mathbb{P}\left[\sup_{x\notin I_{\mathsf{x}}^{\delta}}\left\{\theta_0(\mathsf{x})\Phi_0(x)-\Gamma_0(x)\right\}\geq\theta_0(\mathsf{x})\Phi_0(\mathsf{x})-\Gamma_0(\mathsf{x})+o_{\mathbb{P}}(1)\right]=o(1),\end{aligned}$$

where the equality uses the fact, noted by Westling and Carone (2020, Supplement, proof of Lemma 3), that the function $v\mapsto\theta_0(\mathsf{x})\Phi_0(v)-\Gamma_0(v)$ is unimodal and maximized at $v=\mathsf{x}$.

Next, defining $\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}(j)=\{v\in\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}:2^j<|v|\leq 2^{j+1}\}$ and using $\widehat{v}_n(t)=o_{\mathbb{P}}(a_n)$, we have, for any K, any positive δ', and any sequence of events $\{\mathcal{A}'_n\}$ with $\lim_{n\to\infty}\mathbb{P}[\mathcal{A}'_n]=1$,

$$\begin{aligned}\limsup_{n\to\infty}\mathbb{P}\left[|\widehat{v}_n(t)|>2^K\right]&\leq\limsup_{n\to\infty}\sum_{j\geq K:2^j\leq a_n\delta'}\mathbb{P}\left[2^j<|\widehat{v}_n(t)|\leq 2^{j+1}\cap\mathcal{A}'_n\right]\\&\leq\limsup_{n\to\infty}\sum_{j\geq K:2^j\leq a_n\delta'}\mathbb{P}\left[\inf_{v\in\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}(j)}\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t)\leq\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t)\cap\mathcal{A}'_n\right].\end{aligned}$$

The proof of (SA.7) can therefore be completed by showing that the majorant side in the display can be made arbitrarily small by choice of K, δ', and $\{\mathcal{A}'_n\}$.

To do so, we begin by analyzing each term in the basic bound

$$\begin{aligned}\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t)-\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t)\geq{}&\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)+\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(M^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)-t\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)\\&+\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)-r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)-\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t).\end{aligned}$$

Because $\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t)\rightsquigarrow\mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(0;t)=0$ and because, by (B8) and Lemma SA-6, there is a positive δ' such that

$$\sup_{|v|\leq a_n\delta'}\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)-r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|=o_{\mathbb{P}}(1),$$

we may assume that, on $\{\mathcal{A}'_n\}$ and for some C_0,

$$\sup_{|v|\leq a_n\delta'}\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)-r_n\theta_0(\mathsf{x})\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)-\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t)\right|\leq C_0.$$

Also, because, by (B4) and (A3), there is a positive δ' such that

$$\sup_{|v|\leq a_n\delta'}\left|\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)-L^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|=o_{\mathbb{P}}(1)\qquad\text{and}\qquad\sup_{1\leq|v|\leq a_n\delta'}\left|\frac{L^{\mathfrak{q}}_{\mathsf{x},n}(v)}{\mathcal{L}_{\mathsf{x}}(v)}\right|<\infty,$$

we may assume that, on $\{\mathcal{A}'_n\}$ and for some C_L,

$$\sup_{1\leq|v|\leq a_n\delta'}\left|\frac{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}(v)}{v}\right|\leq C_L.$$

Next, by (B4) and Lemma SA-5, with probability approaching one,

$$\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(M^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)\geq\inf_{|v|/2\leq|v'|\leq 2|v|}M^{\mathfrak{q}}_{\mathsf{x},n}(v')\qquad\text{for every }|v|\geq 2,$$

while, by (A2) and (A3), there is a positive δ' such that

$$\inf_{1\leq|v|\leq\eta a_n\delta'}\frac{M^{\mathfrak{q}}_{\mathsf{x},n}(v)}{\mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}(v)}>0.$$

We may therefore assume that, on $\{\mathcal{A}'_n\}$ and for some positive C_M,

$$\inf_{2\leq|v|\leq a_n\delta'}\frac{\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(M^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)}{v^{\mathfrak{q}+1}}\geq C_M.$$

Finally, by (B4) and Lemma SA-5, with probability approaching one,

$$\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)\right|\leq\sup_{|v|/2\leq|v'|\leq 2|v|}\left|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v')\right|\qquad\text{for every }|v|\geq 2,$$

and we may therefore assume that, on $\{\mathcal{A}_n'\}$,

$$\sup_{v\in\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}(j)}\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)\right|\leq\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|\qquad\text{for every } j\geq 2 \text{ with } 2^j\leq a_n\delta',$$

where $V_\eta(j)=\{v\in\mathbb{R}:\eta^{-1}2^j\leq|v|\leq\eta 2^{j+1}\}$.

As a consequence, by the Markov inequality,

$$\begin{aligned}
&\mathbb{P}\left[\inf_{v\in\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}(j)}\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t)\leq\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t)\cap\mathcal{A}_n'\right]\\
&\leq\mathbb{P}\left[\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|\geq\inf_{v\in V_1(j)}\left[C_Mv^{\mathfrak{q}+1}-C_L|v|-C_0\right]\cap\mathcal{A}_n'\right]\\
&\leq\frac{\mathbb{E}\left[\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|\mathbb{1}_{\mathcal{A}_n'}\right]}{\inf_{v\in V_1(j)}\left[C_Mv^{\mathfrak{q}+1}-C_L|v|-C_0\right]}\qquad\text{for every } j\geq 2\text{ with } 2^j\leq a_n\delta',
\end{aligned}$$

where, by (B4), we may assume that, for some C_G,

$$\mathbb{E}\left[\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|\mathbb{1}_{\mathcal{A}_n'}\right]\leq\sum_{j-1\leq j'\leq j+1}\mathbb{E}\left[\sup_{v\in V_1(j)}\left|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|\mathbb{1}_{\mathcal{A}_n'}\right]\leq C_G2^{j\beta},$$

and where, for all sufficiently large j,

$$\inf_{v\in V_1(j)}\left[C_Mv^{\mathfrak{q}+1}-C_L|v|-C_0\right]\geq\frac{1}{2}C_M2^{j(\mathfrak{q}+1)}.$$

In other words, for large K,

$$\limsup_{n\to\infty}\sum_{j\geq K:2^j\leq a_n\delta'}\mathbb{P}\left[\inf_{v\in\widehat{V}^{\mathfrak{q}}_{\mathsf{x},n}(j)}\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(v;t)\leq\widehat{H}^{\mathfrak{q}}_{\mathsf{x},n}(0;t)\cap\mathcal{A}_n'\right]\leq\frac{2C_G}{C_M}\sum_{j\geq K}2^{j[\beta-(\mathfrak{q}+1)]},$$

which can be made arbitrarily small by choice of K.

Proof of (7). We proceed as in the proof of (2). Let $t\in\mathbb{R}$ be given. By Lemma A.1 and change of variables,

$$\begin{aligned}
&\mathbb{P}_n^*\left[r_n(\widetilde{\theta}_n^*(\mathsf{x})-\widehat{\theta}_n(\mathsf{x}))>t\right]\\
&=\mathbb{P}_n^*\left[\max\operatorname*{argmax}_{x\in\widehat{\Phi}_n^{*-}([0,\widehat{u}_n^*])}\left\{[\widehat{\theta}_n(\mathsf{x})+tr_n^{-1}]\widehat{\Phi}_n^*(x)-\mathsf{LSC}_{\widehat{\Phi}_n^*}\left(\widetilde{\Gamma}_n^*\right)(x)\right\}<\widehat{\Phi}_n^{*-}\circ\widehat{\Phi}_n^*(\mathsf{x})\right]\\
&=\mathbb{P}\left[\max\operatorname*{argmin}_{v\in\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}}\widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(v;t)<\widehat{Z}^{\mathfrak{q},*}_{\mathsf{x},n}\right],
\end{aligned}$$

where

$$\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}=\left\{a_n(x-\mathsf{x}):x\in\widehat{\Phi}_n^{*-}([0,\widehat{u}_n^*])\right\},$$

$$\widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(v;t)=\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}+\widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}+r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right)(v)-[r_n\widehat{\theta}_n(\mathsf{x})+t]\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}(v)$$

and

$$\widehat{Z}^{\mathfrak{q},*}_{\mathsf{x},n}=a_n\left[\widehat{\Phi}_n^{*-}\circ\widehat{\Phi}_n^*(\mathsf{x})-\mathsf{x}\right],$$

with

$$\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v) = \sqrt{na_n}\left[\widehat{\Gamma}^*_n(\mathsf{x}+va_n^{-1}) - \widehat{\Gamma}^*_n(\mathsf{x}) - \widehat{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \widehat{\Gamma}_n(\mathsf{x})\right]$$
$$- \widehat{\theta}_n(\mathsf{x})\sqrt{na_n}\left[\widehat{\Phi}^*_n(\mathsf{x}+va_n^{-1}) - \widehat{\Phi}^*_n(\mathsf{x}) - \widehat{\Phi}_n(\mathsf{x}+va_n^{-1}) + \widehat{\Phi}_n(\mathsf{x})\right],$$
$$\widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}(v) = \sqrt{na_n}\widetilde{M}_{\mathsf{x},n}(va_n^{-1}), \qquad \widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}(v) = a_n\left[\widehat{\Phi}^*_n(\mathsf{x}+va_n^{-1}) - \widehat{\Phi}^*_n(\mathsf{x})\right].$$

By (B4) and Lemma SA-4, $\widehat{Z}^{\mathfrak{q},*}_{\mathsf{x},n} = o_{\mathbb{P}}(1)$. Suppose also that

$$\max \operatorname*{argmin}_{v\in\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}} \widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(v;t) \rightsquigarrow_{\mathbb{P}} \operatorname*{argmin}_{v\in\mathbb{R}} \mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(v;t). \tag{SA.9}$$

Then, as in the proof of (2),

$$\mathbb{P}^*_n\left[r_n(\widetilde{\theta}^*_n(\mathsf{x}) - \widehat{\theta}_n(\mathsf{x})) > t\right] \to_{\mathbb{P}} \mathbb{P}\left[\frac{1}{\partial\Phi_0(\mathsf{x})}\partial_-\mathsf{GCM}_{\mathbb{R}}(\mathcal{G}_{\mathsf{x}} + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}(0) > t\right].$$

The proof of (7) can therefore be completed by showing (SA.9).

We shall do so by showing that

$$\widehat{v}^*_n(t) = \max \operatorname*{argmin}_{v\in\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}} \widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(v;t) = O_{\mathbb{P}}(1) \tag{SA.10}$$

and

$$\widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(\cdot;t) \rightsquigarrow_{\mathbb{P}} \mathcal{H}^{\mathfrak{q}}_{\mathsf{x}}(\cdot;t). \tag{SA.11}$$

We begin by showing (SA.11). First, by (B1), $\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} \rightsquigarrow_{\mathbb{P}} \mathcal{G}_{\mathsf{x}}$. Also, by Assumption C, $\widetilde{M}^{\mathfrak{q},*}_{\mathsf{x},n} \rightsquigarrow_{\mathbb{P}} \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}$. Moreover, $\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n} - \widehat{L}^{\mathfrak{q}}_{\mathsf{x},n} \rightsquigarrow_{\mathbb{P}} 0$ by (B4), where, as shown in the proof of (SA.6), $\widehat{L}^{\mathfrak{q}}_{\mathsf{x},n} \rightsquigarrow \mathcal{L}_{\mathsf{x}}$. In particular, $\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n} \rightsquigarrow_{\mathbb{P}} \mathcal{L}_{\mathsf{x}}$ and therefore

$$(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}, \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}, \widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}) \rightsquigarrow_{\mathbb{P}} (\mathcal{G}^{\mathfrak{q}}_{\mathsf{x}}, \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}, \mathcal{L}_{\mathsf{x}}).$$

Because $\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}$ is asymptotically equicontinuous,

$$\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}\right) - \left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}\right) \rightsquigarrow_{\mathbb{P}} 0$$

by (B4) and Lemma SA-5. Also, by (B8) and Lemma SA-6,

$$\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right) - \left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right) \rightsquigarrow_{\mathbb{P}} 0.$$

The result (SA.11) follows from the three preceding displays and the fact that, on $\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}$,

$$0 \geq \mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n} + r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right) - \left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n} + r_n\widehat{\theta}_n(\mathsf{x}), \widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right)$$
$$\geq \mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}\right) - \left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n} + \widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}\right)$$
$$+ \mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right) - \left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right).$$

Next, to show (SA.10), we first define $\widehat{\theta}_n(\mathsf{x};t) = \widehat{\theta}_n(\mathsf{x}) + tr_n^{-1}$ and note that

$$\widehat{v}^*_n(t) = a_n\left[\max \operatorname*{argmax}_{x\in\widehat{\Phi}^{*-}_n([0,\widehat{u}^*_n])} \left\{\widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}^*_n(x) - \mathsf{LSC}_{\widehat{\Phi}^*_n}\left(\widetilde{\Gamma}^*_n\right)(x)\right\} - \mathsf{x}\right].$$

Now, if $|\widehat{v}_n^*(t)| > a_n\delta > 0$, then

$$\sup_{x\notin I_{\mathsf{x}}^\delta}\left\{\widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}_n^*(x) - \mathsf{LSC}_{\widehat{\Phi}_n^*}\left(\widetilde{\Gamma}_n^*\right)(x)\right\} \geq \widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}_n^*(\mathsf{x}) - \mathsf{LSC}_{\widehat{\Phi}_n^*}\left(\widetilde{\Gamma}_n^*\right)(\mathsf{x}),$$

where $|\widehat{\theta}_n(\mathsf{x};t) - \theta_0(\mathsf{x})| = O_{\mathbb{P}}(r_n^{-1}) = o_{\mathbb{P}}(1)$, and, by (B4),

$$\sup_{x\in I}\left|\widehat{\Phi}_n^*(x) - \Phi_0(x)\right| = o_{\mathbb{P}}(1).$$

Therefore, defining $\widehat{\mathsf{x}}_n^* = \widehat{\Phi}_n^{*-}(\widehat{\Phi}_n^*(\mathsf{x})) = \mathsf{x} + o_{\mathbb{P}}(1)$ and using (SA.5),

$$\begin{aligned}
&\widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}_n^*(\mathsf{x}) - \mathsf{LSC}_{\widehat{\Phi}_n^*}\left(\widetilde{\Gamma}_n^*\right)(\mathsf{x})\\
&\leq \widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}_n^*(\mathsf{x}) - \widetilde{\Gamma}_n^*(\widehat{\mathsf{x}}_n^*)\\
&= \widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}_n^*(\mathsf{x}) - \widehat{\theta}_n(\mathsf{x})\widehat{\Phi}_n(\widehat{\mathsf{x}}_n^*) - \widehat{\Gamma}_n^*(\widehat{\mathsf{x}}_n^*) + \widehat{\Gamma}_n(\widehat{\mathsf{x}}_n^*) - \widetilde{M}_{\mathsf{x},n}^{\mathfrak{q}}(\widehat{Z}_n^{\mathfrak{q},*})/\sqrt{na_n} = o_{\mathbb{P}}(1),
\end{aligned}$$

where the last equality uses (B3), $\widehat{Z}_n^{\mathfrak{q},*} = o_{\mathbb{P}}(1)$, and Assumption C. Also, using (B3), (B4), and Assumption C, we have, uniformly in $x \notin I_{\mathsf{x}}^\delta$ and for some $c > 0$,

$$\begin{aligned}
\widetilde{\Gamma}_n^*(x) - \widehat{\theta}_n(\mathsf{x})\widehat{\Phi}_n^*(x) &= \widehat{\Gamma}_n^*(x) - \widehat{\Gamma}_n(x) - \widehat{\theta}_n(\mathsf{x})\left(\widehat{\Phi}_n^*(x) - \widehat{\Phi}_n(x)\right) + \widetilde{M}_n(x-\mathsf{x})\\
&\geq c\delta^{\mathfrak{q}+1} + o_{\mathbb{P}}(1),
\end{aligned}$$

and therefore, by Lemma SA-3,

$$\sup_{x\notin I_{\mathsf{x}}^\delta}\left\{\widehat{\theta}_n(\mathsf{x};t)\widehat{\Phi}_n^*(x) - \mathsf{LSC}_{\widehat{\Phi}_n^*}\left(\widetilde{\Gamma}_n^*\right)(x)\right\} \leq -c\delta^{\mathfrak{q}+1} + o_{\mathbb{P}}(1).$$

As a consequence, $\widehat{v}_n^*(t) = o_{\mathbb{P}}(a_n)$: For any $\delta > 0$,

$$\mathbb{P}\left[|\widehat{v}_n^*(t)| > a_n\delta\right] \leq \mathbb{P}\left[-c\delta^{\mathfrak{q}+1} \geq o_{\mathbb{P}}(1)\right] = o(1).$$

Next, defining $\widehat{V}_{\mathsf{x},n}^{\mathfrak{q},*}(j) = \{v \in \widehat{V}_{\mathsf{x},n}^{\mathfrak{q},*} : 2^j < |v| \leq 2^{j+1}\}$ and using $\widehat{v}_n^*(t) = o_{\mathbb{P}}(a_n)$, we have, for any K, any positive δ', and any sequence of events $\{\mathcal{A}_n'\}$ with $\lim_{n\to\infty}\mathbb{P}[\mathcal{A}_n'] = 1$,

$$\begin{aligned}
\limsup_{n\to\infty}\mathbb{P}\left[|\widehat{v}_n^*(t)| > 2^K\right] &\leq \limsup_{n\to\infty}\sum_{j\geq K:2^j\leq a_n\delta'}\mathbb{P}\left[2^j < |\widehat{v}_n^*(t)| \leq 2^{j+1} \cap \mathcal{A}_n'\right]\\
&\leq \limsup_{n\to\infty}\sum_{j\geq K:2^j\leq a_n\delta'}\mathbb{P}\left[\inf_{v\in\widehat{V}_{\mathsf{x},n}^{\mathfrak{q},*}(j)}\widehat{H}_{\mathsf{x},n}^{\mathfrak{q},*}(v;t) \leq \widehat{H}_{\mathsf{x},n}^{\mathfrak{q},*}(0;t) \cap \mathcal{A}_n'\right].
\end{aligned}$$

The proof of (SA.10) can therefore be completed by showing that the majorant side in the display can be made arbitrarily small by choice of K, δ', and $\{\mathcal{A}_n'\}$.

To do so, we begin by analyzing each term in the basic bound

$$\begin{aligned}
\widehat{H}_{\mathsf{x},n}^{\mathfrak{q},*}(v;t) - \widehat{H}_{\mathsf{x},n}^{\mathfrak{q},*}(0;t) \geq\; &\mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}}\left(\widehat{G}_{\mathsf{x},n}^{\mathfrak{q},*}\right)(v) + \mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}}\left(\widetilde{M}_{\mathsf{x},n}^{\mathfrak{q}}\right)(v) - t\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}(v)\\
&+ \mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}}\left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}\right)(v) - r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}(v) - \widehat{H}_{\mathsf{x},n}^{\mathfrak{q},*}(0;t).
\end{aligned}$$

Because $\widehat{H}_{\mathsf{x},n}^{\mathfrak{q},*}(0;t) \rightsquigarrow_{\mathbb{P}} \mathcal{H}_{\mathsf{x}}^{\mathfrak{q}}(0;t) = 0$ and because, by (B8) and Lemma SA-6, there is a positive δ' such that

$$\sup_{|v|\leq a_n\delta'}\left|\mathsf{LSC}_{\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}}\left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}\right)(v) - r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}_{\mathsf{x},n}^{\mathfrak{q},*}(v)\right| = o_{\mathbb{P}}(1),$$

we may assume that, on $\{\mathcal{A}'_n\}$ and for some C_0,

$$\sup_{|v|\leq a_n\delta'}\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}\right)(v)-r_n\widehat{\theta}_n(\mathsf{x})\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}(v)-\widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(0;t)\right|\leq C_0.$$

Also, because, by (B4) and (A3), there is a positive δ' such that

$$\sup_{|v|\leq a_n\delta'}\left|\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}(v)-L^{\mathfrak{q}}_{\mathsf{x},n}(v)\right|=o_{\mathbb{P}}(1)\qquad\text{and}\qquad\sup_{1\leq|v|\leq a_n\delta'}\left|\frac{L^{\mathfrak{q}}_{\mathsf{x},n}(v)}{\mathcal{L}_{\mathsf{x}}(v)}\right|<\infty,$$

we may assume that, on $\{\mathcal{A}'_n\}$ and for some C_L,

$$\sup_{1\leq|v|\leq a_n\delta'}\left|\frac{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}(v)}{v}\right|\leq C_L.$$

Next, by (B4) and Lemma SA-5, with probability approaching one,

$$\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)\geq\inf_{|v|/2\leq|v'|\leq2|v|}\widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}(v')\qquad\text{for every }|v|\geq2,$$

while, by Assumption C, there is a positive c such that, with probability approaching one,

$$\inf_{|v|\geq1}\frac{\widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}(v)}{v^{\mathfrak{q}+1}}>c.$$

We may therefore assume that, on $\{\mathcal{A}'_n\}$ and for some positive C_M,

$$\inf_{2\leq|v|\leq a_n\delta'}\frac{\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widetilde{M}^{\mathfrak{q}}_{\mathsf{x},n}\right)(v)}{v^{\mathfrak{q}+1}}\geq C_M.$$

Finally, by (B4) and Lemma SA-5, with probability approaching one,

$$\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}\right)(v)\right|\leq\sup_{|v|/2\leq|v'|\leq2|v|}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v')\right|\qquad\text{for every }|v|\geq2,$$

and we may therefore assume that, on $\{\mathcal{A}'_n\}$,

$$\sup_{v\in\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}(j)}\left|\mathsf{LSC}_{\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}}\left(\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}\right)(v)\right|\leq\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v)\right|\qquad\text{for every }j\geq2\text{ with }2^j\leq a_n\delta'.$$

As a consequence, by the Markov inequality,

$$\begin{aligned}&\mathbb{P}\left[\inf_{v\in\widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}(j)}\widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(v;t)\leq\widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(0;t)\cap\mathcal{A}'_n\right]\\&\leq\mathbb{P}\left[\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v)\right|\geq\inf_{v\in V_1(j)}\left[C_Mv^{\mathfrak{q}+1}-C_L|v|-C_0\right]\cap\mathcal{A}'_n\right]\\&\leq\frac{\mathbb{E}\left[\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v)\right|\mathbb{1}_{\mathcal{A}'_n}\right]}{\inf_{v\in V_1(j)}\left[C_Mv^{\mathfrak{q}+1}-C_L|v|-C_0\right]}\qquad\text{for every }j\geq2\text{ with }2^j\leq a_n\delta',\end{aligned}$$

where, by (B4), we may assume that, for some C_G,

$$\mathbb{E}\left[\sup_{v\in V_2(j)}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v)\right|\mathbb{1}_{\mathcal{A}'_n}\right]\leq\sum_{j-1\leq j'\leq j+1}\mathbb{E}\left[\sup_{v\in V_1(j)}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v)\right|\mathbb{1}_{\mathcal{A}'_n}\right]\leq C_G2^{j\beta},$$

and where, for all sufficiently large j,

$$\inf_{v\in V_1(j)} \left[C_M v^{\mathfrak{q}+1} - C_L|v| - C_0\right] \geq \frac{1}{2} C_M 2^{j(\mathfrak{q}+1)}.$$

In other words, for large K,

$$\limsup_{n\to\infty} \sum_{j\geq K:2^j\leq a_n\delta'} \mathbb{P}\left[\inf_{v\in \widehat{V}^{\mathfrak{q},*}_{\mathsf{x},n}(j)} \widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(v;t) \leq \widehat{H}^{\mathfrak{q},*}_{\mathsf{x},n}(0;t) \cap \mathcal{A}'_n\right] \leq \frac{2C_G}{C_M} \sum_{j\geq K} 2^{j[\beta-(\mathfrak{q}+1)]},$$

which can be made arbitrarily small by choice of K.

Proof of (8). The bootstrap consistency result (8) follows from (2), (7), Polya's theorem, and the fact that, by Lemma A.2, the limiting distribution in (2) and (7) has a continuous cdf. □

SA.1.5. *Proof of Lemma 1.* For the monomial approximation estimator, we have

$$\begin{aligned}
\widetilde{\mathcal{D}}^{\mathtt{MA}}_{\mathfrak{q},n}(\mathsf{x}) &= \epsilon_n^{-(\mathfrak{q}+1)}[\Gamma_0(\mathsf{x}+\epsilon_n) - \Gamma_0(\mathsf{x}) - \theta_0(\mathsf{x})\{\Phi_0(\mathsf{x}+\epsilon_n) - \Phi_0(\mathsf{x})\}] \\
&\quad + \epsilon_n^{-(\mathfrak{q}+1/2)} n^{-1/2} \widehat{G}_{\mathsf{x},n}(1;\epsilon_n) \\
&\quad - \epsilon_n^{-\mathfrak{q}}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})]\widehat{R}_{\mathsf{x},n}(1;\epsilon_n) \\
&\quad - \epsilon_n^{-(\mathfrak{q}+1)}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})][\Phi_0(\mathsf{x}+\epsilon_n) - \Phi_0(\mathsf{x})] \\
&= \frac{\partial^{\mathfrak{q}}\theta_0(\mathsf{x})\partial\Phi_0(\mathsf{x})}{(\mathfrak{q}+1)!} + o(1) + O_{\mathbb{P}}[(n\epsilon_n^{1+2\mathfrak{q}})^{-1/2} + (n\epsilon_n^{1+2\mathfrak{q}})^{-\mathfrak{q}/(1+2\mathfrak{q})}] \\
&= \frac{\partial^{\mathfrak{q}}\theta_0(\mathsf{x})\partial\Phi_0(\mathsf{x})}{(\mathfrak{q}+1)!} + o_{\mathbb{P}}(1),
\end{aligned}$$

where the second equality uses $\epsilon_n \to 0$ and the last equality uses $n\epsilon_n^{1+2\mathfrak{q}} \to \infty$.

Similarly, for the forward difference estimator, we have

$$\begin{aligned}
\widetilde{\mathcal{D}}^{\mathtt{FD}}_{\mathfrak{q},n}(\mathsf{x}) &= \epsilon_n^{-(\mathfrak{q}+1)} \sum_{k=1}^{\mathfrak{q}+1} (-1)^{k+\mathfrak{q}+1} \binom{\mathfrak{q}+1}{k} [\Upsilon_0(\mathsf{x}+k\epsilon_n) - \Upsilon_0(\mathsf{x})] \\
&\quad + \epsilon_n^{-(\mathfrak{q}+1/2)} n^{-1/2} \sum_{k=1}^{\mathfrak{q}+1} (-1)^{k+\mathfrak{q}+1} \binom{\mathfrak{q}+1}{k} \widehat{G}_{\mathsf{x},n}(k;\epsilon_n) \\
&\quad - \epsilon_n^{-\mathfrak{q}}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})] \sum_{k=1}^{\mathfrak{q}+1} (-1)^{k+\mathfrak{q}+1} \binom{\mathfrak{q}+1}{k} \widehat{R}_{\mathsf{x},n}(k;\epsilon_n) \\
&\quad - \epsilon_n^{-(\mathfrak{q}+1)}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})] \sum_{k=1}^{\mathfrak{q}+1} (-1)^{k+\mathfrak{q}+1} \binom{\mathfrak{q}+1}{k} [\Phi_0(\mathsf{x}+k\epsilon_n) - \Phi_0(\mathsf{x})] \\
&= \frac{\partial^{\mathfrak{q}}\theta_0(\mathsf{x})\partial\Phi_0(\mathsf{x})}{(\mathfrak{q}+1)!} + o(1) + O_{\mathbb{P}}[(n\epsilon_n^{1+2\mathfrak{q}})^{-1/2} + (n\epsilon_n^{1+2\mathfrak{q}})^{-\mathfrak{q}/(1+2\mathfrak{q})}] \\
&= \frac{\partial^{\mathfrak{q}}\theta_0(\mathsf{x})\partial\Phi_0(\mathsf{x})}{(\mathfrak{q}+1)!} + o_{\mathbb{P}}(1). \quad \square
\end{aligned}$$

SA.1.6. *Proof of Lemma 2*. Proceeding as in the proof of Lemma 1, we have

$$\begin{aligned}
\widetilde{\mathcal{D}}^{\mathtt{BR}}_{j,n}(\mathsf{x}) = \epsilon_n^{-(j+1)} \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k)[\Upsilon_0(\mathsf{x}+c_k\epsilon_n) - \Upsilon_0(\mathsf{x})] \\
+ \epsilon_n^{-(j+1/2)} n^{-1/2} \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k) \widehat{G}_{\mathsf{x},n}(c_k;\epsilon_n) \\
- \epsilon_n^{-j}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})] \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k) \widehat{R}_{\mathsf{x},n}(c_k;\epsilon_n) \\
- \epsilon_n^{-(j+1)}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})] \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k)[\Phi_0(\mathsf{x}+c_k\epsilon_n) - \Phi_0(\mathsf{x})] \\
= \frac{\partial^{j+1}\Upsilon_0(\mathsf{x})}{(j+1)!} + O(\epsilon_n^{\min\{\mathfrak{s},\underline{\mathfrak{s}}+1\}-j}) + O_{\mathbb{P}}(\epsilon_n^{-(j+1/2)} n^{-1/2} + \epsilon_n^{-j} a_n^{-\mathfrak{q}}) \\
= \mathcal{D}_j(\mathsf{x}) + O(\epsilon_n^{\min\{\mathfrak{s},\underline{\mathfrak{s}}+1\}-j}) + a_n^{j-\mathfrak{q}} O_{\mathbb{P}}[(a_n\epsilon_n)^{-(j+1/2)} + (a_n\epsilon_n)^{-j}],
\end{aligned}$$

where the second equality uses $\epsilon_n \to 0$ and the defining property of $\{\lambda_j^{\mathtt{BR}}(k) : k = 1, \ldots, \underline{\mathfrak{s}}\}$.

The second part of the lemma follows from the fact that if

$$n\epsilon_n^{(1+2\bar{\mathfrak{q}})\min(\underline{\mathfrak{s}},\mathfrak{s}-1)/(\bar{\mathfrak{q}}-1)} \to 0 \quad \text{and} \quad n\epsilon_n^{1+2\bar{\mathfrak{q}}} \to \infty,$$

then

$$a_n^{\mathfrak{q}-j} \epsilon_n^{\min\{\mathfrak{s},\underline{\mathfrak{s}}+1\}-j} \to 0 \quad \text{and} \quad a_n\epsilon_n \to \infty.$$

□

SA.1.7. *Higher-order expansion of the bias-reduced estimator* . In addition to the assumptions of Lemma 2, suppose that $\widehat{R}_{\mathsf{x},n}(1;\eta_n) = O_{\mathbb{P}}(a_n^{-1/2})$ for $a_n^{-1}\eta_n^{-1} = O(1)$ and that, for some $\delta > 0$, θ_0 is $(\underline{\mathfrak{s}}+1)$-times continuously differentiable and Φ_0 is $(\underline{\mathfrak{s}}+2)$-times continuously differentiable on I_{x}^{δ}. Then, the first term in the stochastic expansion of $\widetilde{\mathcal{D}}^{\mathtt{BR}}_{j,n}(\mathsf{x})$ satisfies

$$\begin{aligned}
\epsilon_n^{-(j+1)} \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k)[\Upsilon_0(\mathsf{x}+c_k\epsilon_n) - \Upsilon_0(\mathsf{x})] - \frac{\partial^{j+1}\Upsilon_0(\mathsf{x})}{(j+1)!} \\
= \epsilon_n^{\underline{\mathfrak{s}}+1-j} \frac{\partial^{\underline{\mathfrak{s}}+2}\Upsilon_0(\mathsf{x})}{(\underline{\mathfrak{s}}+2)!} \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k) c_k^{\underline{\mathfrak{s}}+2} + o(\epsilon_n^{\underline{\mathfrak{s}}+1-j}) = \epsilon_n^{\underline{\mathfrak{s}}+1-j} \mathsf{B}_j^{\mathtt{BR}}(\mathsf{x}) + o(\epsilon_n^{\underline{\mathfrak{s}}+1-j}).
\end{aligned}$$

Also, the approximate variance of

$$\epsilon_n^{-(j+1/2)} n^{-1/2} \sum_{k=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k) \widehat{G}_{\mathsf{x},n}(c_k;\epsilon_n)$$

is

$$\frac{1}{n\epsilon_n^{1+2j}} \sum_{k=1}^{\underline{\mathfrak{s}}+1} \sum_{l=1}^{\underline{\mathfrak{s}}+1} \lambda_j^{\mathtt{BR}}(k) \lambda_j^{\mathtt{BR}}(l) \mathcal{C}_{\mathsf{x}}(c_k, c_l) = \frac{1}{n\epsilon_n^{1+2j}} \mathsf{V}_j^{\mathtt{BR}}(\mathsf{x}).$$

Finally, the third term in the stochastic expansion of $\widetilde{\mathcal{D}}^{\mathtt{BR}}_{j,n}(\mathsf{x})$ is asymptotically negligible under the condition that $\widehat{R}_{\mathsf{x},n}(1;\eta_n) = O_{\mathbb{P}}(a_n^{-1/2})$ for $a_n^{-1}\eta_n^{-1} = O(1)$, while the fourth term exhibits only a higher-order dependence on ϵ_n (relative to the dependence exhibited by the first two terms).

SA.1.8. *Proof of Lemma 3.* We verify that Assumptions A and E imply Assumptions (B1)-(B4). Define

$$\bar{\Gamma}^*_n(x) = \frac{1}{n}\sum_{i=1}^{n} W_{i,n}\gamma_0(x;\mathbf{Z}_i) \qquad \text{and} \qquad \bar{\Phi}^*_n(x) = \frac{1}{n}\sum_{i=1}^{n} W_{i,n}\phi_0(x;\mathbf{Z}_i).$$

Verifying Assumption (B1).

Non-bootstrap weak convergence. We first prove $\widehat{G}^{\mathtt{q}}_{\mathsf{x},n} \rightsquigarrow \mathcal{G}_{\mathsf{x}}$. By Assumption (E3)-(E4),

$$\sqrt{na_n}\sup_{|v|\le K}\left|\widehat{\Gamma}_n(\mathsf{x}+va_n^{-1}) - \widehat{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \bar{\Gamma}_n(\mathsf{x})\right| = o_{\mathbb{P}}(1),$$

$$\sqrt{na_n}\sup_{|v|\le K}\left|\widehat{\Phi}_n(\mathsf{x}+va_n^{-1}) - \widehat{\Phi}_n(\mathsf{x}) - \bar{\Phi}_n(\mathsf{x}+va_n^{-1}) + \bar{\Phi}_n(\mathsf{x})\right| = o_{\mathbb{P}}(1),$$

for each $K > 0$, and thus,

$$\sup_{|v|\le K}\left|\widehat{G}^{\mathtt{q}}_{\mathsf{x},n}(v) - \sqrt{\frac{a_n}{n}}\sum_{i=1}^{n}\left\{\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_i) - \mathbb{E}[\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z})]\right\}\right| = o_{\mathbb{P}}(1).$$

Letting $\bar{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i) = \sqrt{a_n}\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_i)$, we want to prove that the empirical process of $\{\bar{\psi}_{\mathsf{x},n}(v;\cdot) : |v| \le K\}$ weakly converges to $\mathcal{G}_{\mathsf{x}}$. We verify finite-dimensional weak convergence and stochastic equicontinuity.

Letting $\eta_n = Ka_n^{-1}$,

$$n^{-1}\sup_{|v|\le K}\mathbb{E}[|\bar{\psi}_{\mathsf{x},n}(v;\mathbf{Z})|^4] \le (1+|\theta_0(\mathsf{x})|)^4 n^{-1}a_n^2\mathbb{E}[\bar{D}^{\eta_n}_{\gamma}(\mathbf{Z})^4 + \bar{D}^{\eta_n}_{\phi}(\mathbf{Z})^4] = o(1).$$

Also, convergence of the covariance kernel is imposed in Assumption (E5). Thus, the Lyapunov central limit theorem implies the finite-dimensional convergence.

For stochastic equicontinuity, following the argument of Kim and Pollard (1990, Lemma 4.6) and using $a_n\mathbb{E}[\bar{D}^{\eta_n}_{\gamma}(\mathbf{Z})^2 + \bar{D}^{\eta_n}_{\phi}(\mathbf{Z})^2] = O(1)$, it suffices to show $\sup_{|v-s|\le\epsilon_n,|v|\vee|s|\le K}\frac{1}{n}\sum_{i=1}^{n}|\bar{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i) - \bar{\psi}_{\mathsf{x},n}(s;\mathbf{Z}_i)|^2 = o_{\mathbb{P}}(1)$ for any $\epsilon_n = o(1)$. For a constant $M > 0$,

$$\begin{aligned}
&\sup_{|v-s|\le\epsilon_n,|v|\vee|s|\le K}\frac{1}{n}\sum_{i=1}^{n}|\bar{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i) - \bar{\psi}_{\mathsf{x},n}(s;\mathbf{Z}_i)|^2 \\
&\le 4(1+|\theta_0(\mathsf{x})|)^2\frac{1}{n}\sum_{i=1}^{n}a_n[\bar{D}^{\eta_n}_{\gamma}(\mathbf{Z}_i)^2 + \bar{D}^{\eta_n}_{\phi}(\mathbf{Z}_i)^2]\mathbb{1}\{\bar{D}^{\eta_n}_{\gamma}(\mathbf{Z}_i) + \bar{D}^{\eta_n}_{\phi}(\mathbf{Z}_i) > M\} \\
&+ a_nM\sup_{|v-s|\le\epsilon_n,|v|\vee|s|\le K}\mathbb{E}\left[|\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}) - \psi_{\mathsf{x}}(sa_n^{-1};\mathbf{Z})|\right] \\
&+ Ma_n\sup_{\substack{|v-s|\le\epsilon_n\\|v|\vee|s|\le K}}\frac{1}{n}\sum_{i=1}^{n}\left\{|\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_i) - \psi_{\mathsf{x}}(sa_n^{-1};\mathbf{Z}_i)| - \mathbb{E}\left[|\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}) - \psi_{\mathsf{x}}(sa_n^{-1};\mathbf{Z})|\right]\right\}
\end{aligned}$$

where the first term after the inequality can be made arbitrarily small by making M large using $a_n\mathbb{E}[\bar{D}^{\eta_n}_{\gamma}(\mathbf{Z})^4 + \bar{D}^{\eta_n}_{\phi}(\mathbf{Z})^4] = O(1)$. The second term is $o_{\mathbb{P}}(1)$ by Assumption (E5). Finally, the third term is $O_{\mathbb{P}}(\sqrt{a_n/n})$ using Theorem 4.2 of Pollard (1989).

Bootstrap weak convergence. We next prove $\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n} \rightsquigarrow_{\mathbb{P}} \mathcal{G}_{\mathsf{x}}$. First we posit

$$\sqrt{na_n}\sup_{|v|\le K}|\widehat{\Gamma}^*_n(\mathsf{x}+va_n^{-1})-\widehat{\Gamma}^*_n(\mathsf{x})-\bar{\Gamma}^*_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}^*_n(\mathsf{x})|=o_{\mathbb{P}}(1),$$

$$\sqrt{na_n}\sup_{|v|\le K}|\widehat{\Phi}^*_n(\mathsf{x}+va_n^{-1})-\widehat{\Phi}^*_n(\mathsf{x})-\bar{\Phi}^*_n(\mathsf{x}+va_n^{-1})+\bar{\Phi}^*_n(\mathsf{x})|=o_{\mathbb{P}}(1), \tag{SA.12}$$

which follows from the hypothesis of the lemma as shown below. By the above display,

$$\sup_{|v|\le K}\left|\widehat{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v)-\sqrt{\frac{a_n}{n}}\sum_{i=1}^n W_{i,n}\Big\{\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_i)-\frac{1}{n}\sum_{j=1}^n\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_j)\Big\}\right|=o_{\mathbb{P}}(1)$$

where we use $\sqrt{\frac{a_n}{n}}\sum_{i=1}^n W_{i,n}\{\phi_0(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)-\phi_0(\mathsf{x};\mathbf{Z}_i)-\frac{1}{n}\sum_{j=1}^n[\phi_0(\mathsf{x}+va_n^{-1};\mathbf{Z}_j)-\phi_0(\mathsf{x};\mathbf{Z}_j)]\}=O_{\mathbb{P}}(1)$ uniformly over $|v|\le K$, and $\widehat{\theta}_n(\mathsf{x})\to_{\mathbb{P}}\theta_0(\mathsf{x})$. Let $\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z})=\sqrt{a_n}[\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z})-\frac{1}{n}\sum_{j=1}^n\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_j)]$ and to prove the finite-dimensional convergence, we apply Lemma 3.6.15 of van der Vaart and Wellner (1996). Assumption (E2) implies $\frac{1}{n}\sum_{i=1}^n(W_{i,n}-1)^2\to_{\mathbb{P}}1$ and $n^{-1}\max_{1\le i\le n}W_{i,n}^2=o_{\mathbb{P}}(1)$. Since

$$\frac{1}{n}\sum_{i=1}^n\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i)\widehat{\psi}_{\mathsf{x},n}(u;\mathbf{Z}_i)=a_n\Bigg[\frac{1}{n}\sum_{i=1}^n\psi_{\mathsf{x}}(v;\mathbf{Z}_i)\psi_{\mathsf{x}}(u;\mathbf{Z}_i)-\frac{1}{n}\sum_{i=1}^n\psi_{\mathsf{x}}(v;\mathbf{Z}_i)\frac{1}{n}\sum_{i=1}^n\psi_{\mathsf{x}}(u;\mathbf{Z}_i)\Bigg]$$

and $\sup_{|v|\le\eta}|\psi_{\mathsf{x}}(v;\mathbf{Z})|\le\bar{D}^\eta_\gamma(\mathbf{Z})+|\theta_0(\mathsf{x})|\bar{D}^\eta_\phi(\mathbf{Z})$, for any $v,u\in\mathbb{R}$,

$$\frac{1}{n}\sum_{i=1}^n\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i)\widehat{\psi}_{\mathsf{x},n}(u;\mathbf{Z}_i)-a_n\mathbb{E}[\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z})\psi_{\mathsf{x}}(ua_n^{-1};\mathbf{Z})]=o_{\mathbb{P}}(1).$$

Also, $\frac{1}{n}\sum_{i=1}^n\widehat{\psi}^4_{\mathsf{x},n}(v;\mathbf{Z}_i)=O_{\mathbb{P}}(1)$ and we verified the hypothesis of the lemma.

For stochastic equicontinuity, let $\epsilon_n=o(1)$ and $\eta_n=Ka_n^{-1}$. Lemma 3.6.7 of van der Vaart and Wellner (1996) implies that for any $n_0\in\{1,\dots,n\}$, there is a fixed constant $C>0$ such that

$$\begin{aligned}&\mathbb{E}\Bigg[\sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K}\Bigg|\frac{1}{\sqrt{n}}\sum_{i=1}^n W_{i,n}\big[\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i)-\widehat{\psi}_{\mathsf{x},n}(u;\mathbf{Z}_i)\big]\Bigg|\Bigg|\{\mathbf{Z}_i\}_{i=1}^n\Bigg]\\
&\le C\frac{\sqrt{a_n}}{n}\sum_{i=1}^n\big[\bar{D}^{\eta_n}_\gamma(\mathbf{Z}_i)+\bar{D}^{\eta_n}_\phi(\mathbf{Z}_i)\big](n_0-1)\mathbb{E}\max_{1\le i\le n}|W_{i,n}|n^{-1/2}\\
&+C\max_{n_0\le k\le n}\mathbb{E}\Bigg[\sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K}\Bigg|\frac{1}{\sqrt{k}}\sum_{i=n_0}^k\big[\widehat{\psi}_{\mathsf{x},n}(va_n^{-1};\mathbf{Z}_{R_i})-\widehat{\psi}_{\mathsf{x},n}(ua_n^{-1};\mathbf{Z}_{R_i})\big]\Bigg|\Bigg|\{\mathbf{Z}_i\}_{i=1}^n\Bigg]\end{aligned}$$

where $(R_1,\dots,R_n)$ is uniformly distributed on the set of all permutations of $\{1,\dots,n\}$, independent of $\{\mathbf{Z}_i\}_{i=1}^n$. Choose n_0 such that $n^{1/2-1/\mathfrak{r}}/n_0\to\infty$ and $n_0/a_n\to\infty$ (which is possible by $\mathfrak{r}>(4\mathfrak{q}+2)/(2\mathfrak{q}-1)$), and the first term after the inequality in the above display is $o_{\mathbb{P}}(1)$. For the second term, following the argument of van der Vaart and Wellner (1996, Theorem 3.6.13), it suffices to bound

$$\max_{n_0\le k\le n}\mathbb{E}^*\sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K}\Bigg|\frac{1}{\sqrt{k}}\sum_{i=1}^k\big[\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z}^*_i)-\widehat{\psi}_{\mathsf{x},n}(u;\mathbf{Z}^*_i)\big]\Bigg|$$

where $\{\mathbf{Z}_i^*\}_{i=1}^k$ denotes a random sample from the empirical cdf and $\mathbb{E}^*$ is the expectation under this empirical bootstrap law. Following the argument of Kim and Pollard (1990, Lemma 4.6), it suffices to show

$$\max_{n_0\le k\le n} \mathbb{E}^* \sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K} \frac{1}{k}\sum_{i=1}^k \left|\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i^*)-\widehat{\psi}_{\mathsf{x},n}(u;\mathbf{Z}_i^*)\right|^2 = o_{\mathbb{P}}(1).$$

For $k\in\{n_0,\dots,n\}$ and $M>0$,

$$\begin{aligned}
&\mathbb{E}^* \sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K} \frac{1}{k}\sum_{i=1}^k \left|\widehat{\psi}_{\mathsf{x},n}(v;\mathbf{Z}_i^*)-\widehat{\psi}_{\mathsf{x},n}(u;\mathbf{Z}_i^*)\right|^2\\
&\le 2(1+|\theta_0(\mathsf{x})|)^2 a_n \frac{1}{n}\sum_{i=1}^n \left[\bar{D}_\gamma^{\eta_n}(\mathbf{Z}_i)^2+\bar{D}_\phi^{\eta_n}(\mathbf{Z}_i)^2\right]\mathbb{1}\{\bar{D}_\gamma^{\eta_n}(\mathbf{Z}_i)+\bar{D}_\phi^{\eta_n}(\mathbf{Z}_i)>M-o_{\mathbb{P}}(1)\}\\
&\quad +2Ma_n \sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K} \frac{1}{n}\sum_{i=1}^n \left|\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_i)-\psi_{\mathsf{x}}(ua_n^{-1};\mathbf{Z}_i)\right|\\
&\quad +2Ma_n\mathbb{E}^* \sup_{|v-u|\le\epsilon_n,|v|\vee|u|\le K} \frac{1}{k}\sum_{i=1}^k \left|\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}_i^*)-\psi_{\mathsf{x}}(ua_n^{-1};\mathbf{Z}_i^*)\right|\\
&\qquad\qquad -\mathbb{E}^*\left[\left|\psi_{\mathsf{x}}(va_n^{-1};\mathbf{Z}^*)-\psi_{\mathsf{x}}(ua_n^{-1};\mathbf{Z}^*)\right|\right].
\end{aligned}$$

The first term after the inequality does not depend on k and its expectation can be made arbitrarily small by taking M sufficiently large. The second term is independent of k and we can handle this term by adding and subtracting the expectation inside the summation. For the third term, applying Theorem 4.2 of Pollard (1989) again, it is bounded by a constant multiple of

$$a_n k^{-1/2}\left(\frac{1}{n}\sum_{i=1}^n \left[\bar{D}_\gamma^{\eta_n}(\mathbf{Z}_i)^2+\bar{D}_\phi^{\eta_n}(\mathbf{Z}_i)^2\right]\right)^{1/2} = O_{\mathbb{P}}\big(\sqrt{a_n/k}\big),$$

which is $o_{\mathbb{P}}(1)$ by the choice of n_0.

Verifying (SA.12). We focus on the first display. By adding and subtracting the bootstrap means,

$$\begin{aligned}
&\widehat{\Gamma}_n^*(\mathsf{x}+va_n^{-1})-\widehat{\Gamma}_n^*(\mathsf{x})-\bar{\Gamma}_n^*(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n^*(\mathsf{x})\\
&=\frac{1}{n}\sum_{i=1}^n W_{i,n}\breve{\gamma}_n(v;\mathbf{Z})+\left[\check{\Gamma}_n(\mathsf{x}+va_n^{-1})-\check{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n(\mathsf{x})\right]
\end{aligned}$$

where

$$\begin{aligned}
\breve{\gamma}_n(v;\mathbf{Z}) &= \widehat{\gamma}(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)-\widehat{\gamma}(\mathsf{x};\mathbf{Z}_i)-\gamma_0(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)+\gamma_0(\mathsf{x};\mathbf{Z}_i)\\
&\quad -\check{\Gamma}_n(\mathsf{x}+va_n^{-1})+\check{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n(\mathsf{x}).
\end{aligned}$$

By Assumption (E3), $\sqrt{na_n}\sup_{|v|\le K}|\check{\Gamma}_n(\mathsf{x}+va_n^{-1})-\check{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n(\mathsf{x})| = o_{\mathbb{P}}(1)$. Identical to above, Lemma 3.6.7 and the argument in Theorem 3.6.13 of van der Vaart and Wellner (1996) imply that for some fixed $C>0$,

$$\sqrt{na_n}\mathbb{E}\left[\sup_{|v|\le K}\left|\frac{1}{n}\sum_{i=1}^n W_{i,n}\breve{\gamma}_n(v;\mathbf{Z}_i)\right|\middle|\{\mathbf{Z}_i\}_{i=1}^n\right]$$

(SA.13)

$$\leq C\frac{\sqrt{a_n}}{n}\sum_{i=1}^{n}\sup_{|v|\leq K}|\breve{\gamma}_n(v;\mathbf{Z}_i)|\frac{n_0 n^{\mathfrak{r}}}{\sqrt{n}}+C\sqrt{a_n}\max_{n_0\leq k\leq n}\mathbb{E}^*\sup_{|v|\leq K}\left|\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\breve{\gamma}_n(v;\mathbf{Z}_i^*)\right|.$$

For the first term in the last line,

$$\begin{aligned}
&\frac{\sqrt{a_n}}{n}\sum_{i=1}^{n}\sup_{|v|\leq K}|\breve{\gamma}_n(v;\mathbf{Z}_i)|\\
&\leq \frac{\sqrt{a_n}}{n}\sum_{i=1}^{n}\sup_{|v|\leq K}|\widehat{\gamma}(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)-\widehat{\gamma}(\mathsf{x};\mathbf{Z}_i)-\gamma_0(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)+\gamma_0(\mathsf{x};\mathbf{Z}_i)|\\
&\quad+\sqrt{a_n}\sup_{|v|\leq K}\left|\check{\Gamma}_n(\mathsf{x}+va_n^{-1})-\check{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n(\mathsf{x})\right|
\end{aligned}$$

and both terms are $o_{\mathbb{P}}(1)$ by Assumption (E3). For the second term in (SA.13), Corollary 4.3 of Pollard (1989) implies that for some fixed $C>0$,

$$\begin{aligned}
&\mathbb{E}^*\left[\sup_{|v|\leq K}\left|\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\breve{\gamma}_n(v;\mathbf{Z}_i^*)\right|\right]^2\\
&\leq C\frac{1}{n}\sum_{i=1}^{n}\sup_{|v|\leq K}|\widehat{\gamma}(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)-\widehat{\gamma}(\mathsf{x};\mathbf{Z}_i)-\gamma_0(\mathsf{x}+va_n^{-1};\mathbf{Z}_i)+\gamma_0(\mathsf{x};\mathbf{Z}_i)|^2
\end{aligned}$$

and this term is $o_{\mathbb{P}}(a_n^{-1})$ by Assumption (E3).

Verifying Assumption (B2). Let

$$\begin{aligned}
\bar{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)&=\sqrt{na_n}\big[\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})-\bar{\Gamma}_n(\mathsf{x})-\Gamma_0(\mathsf{x}+va_n^{-1})+\Gamma_0(\mathsf{x})\big]\\
&\quad-\theta_0(\mathsf{x})\sqrt{na_n}\big[\bar{\Phi}_n(\mathsf{x}+va_n^{-1})-\bar{\Phi}_n(\mathsf{x})-\Phi_0(\mathsf{x}+va_n^{-1})+\Phi_0(\mathsf{x})\big].
\end{aligned}$$

From the definition,

$$\begin{aligned}
\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)-\bar{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)&=\sqrt{na_n}\big[\widehat{\Gamma}_n(\mathsf{x}+va_n^{-1})-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n(\mathsf{x})\big]\\
&\quad-\theta_0(\mathsf{x})\sqrt{na_n}\big[\widehat{\Phi}_n(\mathsf{x}+va_n^{-1})-\widehat{\Phi}_n(\mathsf{x})-\bar{\Phi}_n(\mathsf{x}+va_n^{-1})+\bar{\Phi}_n(\mathsf{x})\big]
\end{aligned}$$

For the two terms after the equality, Assumption (E3) and (E4) imply that for $V\in[1,a_n\delta]$,

$$\sqrt{na_n}\sup_{|v|\in[V,2V]}|\widehat{\Gamma}_n(\mathsf{x}+va_n^{-1})-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n(\mathsf{x})|\leq(2V)^{\beta}a_n^{-\beta}B_n+A_n$$

and

$$\sqrt{na_n}\sup_{|v|\in[V,2V]}|\widehat{\Phi}_n(\mathsf{x}+va_n^{-1})-\widehat{\Phi}_n(\mathsf{x})-\bar{\Phi}_n(\mathsf{x}+va_n^{-1})+\bar{\Phi}_n(\mathsf{x})|\leq(2V)^{\beta}a_n^{-\beta}B_n+A_n,$$

where $\beta=\max(\beta_\gamma,\beta_\phi)$, $A_n=\max(A_{\gamma,n},A_{\phi,n})$, and $B_n=\max(B_{\gamma,n},B_{\phi,n})$. Since $A_n=o_{\mathbb{P}}(1)$, $a_n^{-\beta}B_n=o_{\mathbb{P}}(1)$, and A_n and B_n are independent of V, there exists $\eta_n'=o(1)$ such that

$$\lim_{n\to\infty}\mathbb{P}\left[\bigcap_{V\in[1,a_n\delta]}\left\{\sup_{|v|\in[V,2V]}|\widehat{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)-\bar{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)|\leq[V^{\beta}+1]\eta_n'\right\}\right]=1$$

holds, and we take the event in the display to be $\mathcal{A}_n$. Also, $\mathbb{E}[\sup_{|v|\leq V}|\bar{G}^{\mathfrak{q}}_{\mathsf{x},n}(v)|]\leq C\sqrt{V}$ for $V\in(0,a_n\delta]$ follows from Corollary 4.3 of Pollard (1989) and $\limsup_{\eta\downarrow 0}\mathbb{E}[\bar{D}^{\eta}_{\gamma}(\mathbf{Z})^2/\eta]+$

$\mathbb{E}[\bar{D}_{\phi}^{\eta}(\mathbf{Z})^2/\eta] < \infty$. Then, $\mathbb{E}[\sup_{|v|\in[V,2V]}|\widehat{G}_{\mathsf{x},n}^{\mathfrak{q}}(v)|\mathbb{1}_{\mathcal{A}_n}] \leq C\sqrt{V} + [V^{\beta}+1]\eta_n'$ holds, which implies

$$\sup_{V\in[1,a_n\delta]} \mathbb{E}\left[V^{-\beta}\sup_{|v|\in[V,2V]}|\widehat{G}_{\mathsf{x},n}^{\mathfrak{q}}(v)|\mathbb{1}_{\mathcal{A}_n}\right] = O(1).$$

For the bootstrap counterpart, let

$$\begin{aligned}\bar{G}_{\mathsf{x},n}^{\mathfrak{q},*}(v) &= \sqrt{na_n}\big[\bar{\Gamma}_n^*(\mathsf{x}+va_n^{-1}) - \bar{\Gamma}_n^*(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \bar{\Gamma}_n(\mathsf{x})\big]\\ &\quad - \theta_0(\mathsf{x})\sqrt{na_n}\big[\bar{\Phi}_n^*(\mathsf{x}+va_n^{-1}) - \bar{\Phi}_n^*(\mathsf{x}) - \bar{\Phi}_n(\mathsf{x}+va_n^{-1}) + \bar{\Phi}_n(\mathsf{x})\big]\end{aligned}$$

and we have

$$\begin{aligned}&\widehat{G}_{\mathsf{x},n}^{\mathfrak{q},*}(v) - \bar{G}_{\mathsf{x},n}^{\mathfrak{q},*}(v)\\ &= \sqrt{na_n}\big[\widehat{\Gamma}_n^*(\mathsf{x}+va_n^{-1}) - \widehat{\Gamma}_n^*(\mathsf{x}) - \bar{\Gamma}_n^*(\mathsf{x}+va_n^{-1}) + \bar{\Gamma}_n^*(\mathsf{x})\big]\\ &- \widehat{\theta}_n(\mathsf{x})\sqrt{na_n}\big[\widehat{\Phi}_n^*(\mathsf{x}+va_n^{-1}) - \widehat{\Phi}_n^*(\mathsf{x}) - \bar{\Phi}_n^*(\mathsf{x}+va_n^{-1}) + \bar{\Phi}_n^*(\mathsf{x})\big]\\ &+ \sqrt{na_n}\big[\bar{\Gamma}_n(\mathsf{x}+va_n^{-1}) - \bar{\Gamma}_n(\mathsf{x}) - \widehat{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \widehat{\Gamma}_n(\mathsf{x})\big]\\ &- \widehat{\theta}_n(\mathsf{x})\sqrt{na_n}\big[\bar{\Phi}_n(\mathsf{x}+va_n^{-1}) - \bar{\Phi}_n(\mathsf{x}) - \widehat{\Phi}_n(\mathsf{x}+va_n^{-1}) + \widehat{\Phi}_n(\mathsf{x})\big]\\ &+ \sqrt{a_n}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})]\sqrt{n}\big[\bar{\Phi}_n^*(\mathsf{x}+va_n^{-1}) - \bar{\Phi}_n^*(\mathsf{x}) - \bar{\Phi}_n(\mathsf{x}+va_n^{-1}) + \bar{\Phi}_n(\mathsf{x})\big].\end{aligned}\tag{SA.14}$$

The last term is $o_{\mathbb{P}}(1)$ uniformly over $|v| \leq a_n\delta$ as $\sqrt{a_n}[\widehat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})] = o_{\mathbb{P}}(1)$ and $\sqrt{n}\sup_{|v|\leq\delta}|\bar{\Phi}_n^*(\mathsf{x}+v) - \bar{\Phi}_n(\mathsf{x}+v)| = O_{\mathbb{P}}(1)$. For the first term after the equality in (SA.14),

$$\begin{aligned}&\sqrt{na_n}\big[\widehat{\Gamma}_n^*(\mathsf{x}+va_n^{-1}) - \widehat{\Gamma}_n^*(\mathsf{x}) - \bar{\Gamma}_n^*(\mathsf{x}+va_n^{-1}) + \bar{\Gamma}_n^*(\mathsf{x})\big]\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n W_{i,n}\breve{\gamma}_n(va_n^{-1};\mathbf{Z}_i) + \sqrt{na_n}\big[\check{\Gamma}_n(\mathsf{x}+va_n^{-1}) - \check{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \bar{\Gamma}_n(\mathsf{x})\big]\end{aligned}$$

where $\breve{\gamma}_n(v;\mathbf{Z}) = \sqrt{a_n}\{\widehat{\gamma}_n(\mathsf{x}+v;\mathbf{Z}) - \widehat{\gamma}_n(\mathsf{x};\mathbf{Z}) - \gamma_0(\mathsf{x}+v;\mathbf{Z}) + \gamma_0(\mathsf{x};\mathbf{Z}) - [\check{\Gamma}_n(\mathsf{x}+v) - \check{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+v) + \bar{\Gamma}_n(\mathsf{x})]\}$ and the second part satisfies $\sqrt{na_n}\sup_{|v|\leq V}|\check{\Gamma}_n(\mathsf{x}+va_n^{-1}) - \check{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \bar{\Gamma}_n(\mathsf{x})| = [1+V^{\beta}]o_{\mathbb{P}}(1)$ uniformly over $V\in[1,a_n\delta]$ by Assumption (E3). For $\frac{1}{\sqrt{n}}\sum_{i=1}^n W_{i,n}\breve{\gamma}_n(va_n^{-1};\mathbf{Z}_i)$, we apply Lemma 3.6.7 of van der Vaart and Wellner (1996) as we did above to verify Assumption (B1): for $n_0 \in \{1,\ldots,n\}$,

$$\begin{aligned}&\mathbb{E}\left[\sup_{|v|\leq\delta}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n W_{i,n}\breve{\gamma}_n(v;\mathbf{Z}_i)\right|\,\middle|\,\{\mathbf{Z}_i\}_{i=1}^n\right]\\ &\leq C\frac{1}{n}\sum_{i=1}^n\sup_{|v|\leq\delta}|\breve{\gamma}_n(v;\mathbf{Z}_i)|\frac{n_0}{\sqrt{n}}\mathbb{E}\Big[\max_{1\leq i\leq n}|W_{i,n}|\Big] + C\max_{n_0\leq k\leq n}\mathbb{E}^*\sup_{|v|\leq\delta}\left|\frac{1}{\sqrt{k}}\sum_{i=1}^k\breve{\gamma}_n(v;\mathbf{Z}_i^*)\right|\end{aligned}\tag{SA.15}$$

where $C>0$ is a fixed constant, $\mathbb{E}^*$ is the expectation under the empirical cdf measure, and $\{\mathbf{Z}_i^*\}_{i=1}^n$ is a random sample from the empirical cdf. Let n_0 be a diverging sequence (dependent on n) such that $n_0 n^{\mathfrak{r}}\sqrt{a_n/n} = o(1)$. The first term in (SA.15) is bounded by

$$C\frac{1}{n}\sum_{i=1}^n\sup_{|x-\mathsf{x}|\leq\delta}|\widehat{\gamma}_n(x;\mathbf{Z}_i) - \gamma_0(x;\mathbf{Z}_i)| + C\sup_{|v|\leq\delta}|\check{\Gamma}_n(\mathsf{x}+v) - \bar{\Gamma}_n(\mathsf{x}+v)| = o_{\mathbb{P}}(1).$$

By Corollary 4.3 of Pollard (1989), the second term in (SA.15) is bounded by (up to a constant)

$$\sqrt{a_n}\Big|\frac{1}{n}\sum_{i=1}^{n}\sup_{x\in I_{\mathsf{x}}^{\delta}}|\widehat{\gamma}_n(x;\mathbf{Z}_i)-\gamma_0(x;\mathbf{Z}_i)|^2\Big|^{1/2}$$

which is $o_{\mathbb{P}}(1)$ by Assumption (E3). Then, there exists a sequence of random variables $A_n'=o_{\mathbb{P}}(1)$ such that for $V\in[1,a_n\delta]$,

$$\sqrt{na_n}\sup_{|v|\le[V,2V]}\Big|\widehat{\Gamma}_n^*(\mathsf{x}+va_n^{-1})-\widehat{\Gamma}_n^*(\mathsf{x})-\bar{\Gamma}_n^*(\mathsf{x}+va_n^{-1})+\bar{\Gamma}_n^*(\mathsf{x})\Big|\le[V^{\beta}+1]A_n',$$

and by identical arguments, an analogous bound holds for the second term after the inequality in (SA.14). Then, there exists $\eta_n'=o(1)$ and events $\mathcal{A}_n$ such that $\lim_{n\to\infty}\mathbb{P}[\mathcal{A}_n]=1$ and

$$\mathbb{E}\left[\sup_{|v|\in[V,2V]}\big|\widehat{G}_{\mathsf{x},n}^{\mathsf{q},*}(v)-\bar{G}_{\mathsf{x},n}^{\mathsf{q},*}(v)\big|\mathbb{1}_{\mathcal{A}_n}\right]\le[V^{\beta}+1]\eta_n'$$

for any $V\in[1,a_n\delta]$, where bounds for the third and fourth terms after the equality in (SA.14) were derived in the non-bootstrap case. Finally, the bound $\mathbb{E}[\sup_{|v|\in[V,2V]}|\bar{G}_{\mathsf{x},n}^{\mathsf{q},*}(v)|]\le C\sqrt{V}$ holds by a similar argument to the stochastic equicontinuity, and the desired result follows.

Verifying Assumptions (B3)-(B4).

$$\sup_{x\in I}\big|\widehat{\Gamma}_n(x)-\Gamma_0(x)\big|\le\sup_{x\in I}\big|\widehat{\Gamma}_n(x)-\bar{\Gamma}_n(x)\big|+\sup_{x\in I}\big|\bar{\Gamma}_n(x)-\Gamma_0(x)\big|$$

where the first term after the inequality is assumed to be $o_{\mathbb{P}}(1)$ and the second term is $o_{\mathbb{P}}(1)$ by standard arguments. The identical argument implies $\sup_{x\in I}|\widehat{\Phi}_n(x)-\Phi_0(x)|=o_{\mathbb{P}}(1)$. By adding and subtracting,

$$\begin{aligned}\sup_{x\in I_{\mathsf{x}}^{\delta}}\big|\widehat{\Phi}_n(x)-\Phi_0(x)\big|&\le\sup_{x\in I_{\mathsf{x}}^{\delta}}\big|\widehat{\Phi}_n(x)-\widehat{\Phi}_n(\mathsf{x})-\bar{\Phi}_n(x)+\bar{\Phi}_n(\mathsf{x})\big|\\&\quad+\sup_{x\in I_{\mathsf{x}}^{\delta}}\big|\bar{\Phi}_n(x)-\Phi_0(x)\big|+\big|\widehat{\Phi}_n(\mathsf{x})-\bar{\Phi}_n(\mathsf{x})\big|\end{aligned}$$

where the last two terms are $o_{\mathbb{P}}(a_n^{-1})$. Assumption (E4) implies

$$\sup_{x\in I_{\mathsf{x}}^{\delta}}\big|\widehat{\Phi}_n(x)-\widehat{\Phi}_n(\mathsf{x})-\bar{\Phi}_n(x)+\bar{\Phi}_n(\mathsf{x})\big|\le\big(na_n\big)^{-1/2}[A_{\phi,n}+\delta^{\beta_\phi}B_{\phi,n}]=o_{\mathbb{P}}(a_n^{-1})$$

where the last equality uses $B_{\phi,n}=o_{\mathbb{P}}(a_n^{\beta_\phi})$ and $\beta_\phi\le\mathsf{q}$.

Now we look at the bootstrap objects. For $\widehat{\Gamma}_n^*$,

$$\sup_{x\in I}\big|\widehat{\Gamma}_n^*(x)-\widehat{\Gamma}_n(x)\big|\le\sup_{x\in I}\big|\widehat{\Gamma}_n^*(x)-\bar{\Gamma}_n^*(x)\big|+\sup_{x\in I}\big|\bar{\Gamma}_n^*(x)-\bar{\Gamma}_n(x)\big|+\sup_{x\in I}\big|\bar{\Gamma}_n(x)-\widehat{\Gamma}_n(x)\big|$$

where the last term is $o_{\mathbb{P}}(1)$ by the hypothesis. For $\widehat{\Gamma}_n^*(x)-\bar{\Gamma}_n^*(x)$,

$$\begin{aligned}\sup_{x\in I}\big|\widehat{\Gamma}_n^*(x)-\bar{\Gamma}_n^*(x)\big|&\le\frac{1}{n}\sum_{i=1}^{n}|W_{i,n}|\sup_{x\in I}\big|\widehat{\gamma}_n(x;\mathbf{Z}_i)-\gamma_0(x;\mathbf{Z}_i)\big|\\&\le\sqrt{\frac{1}{n}\sum_{i=1}^{n}|W_{i,n}|^2\frac{1}{n}\sum_{i=1}^{n}\sup_{x\in I}\big|\widehat{\gamma}_n(x;\mathbf{Z}_i)-\gamma_0(x;\mathbf{Z}_i)\big|^2}\end{aligned}$$

and the last term is $o_{\mathbb{P}}(1)$ by the hypothesis. For $\bar{\Gamma}^*_n(x) - \bar{\Gamma}_n(x)$, using Lemma 3.6.7 of van der Vaart and Wellner (1996) and the same argument as for verifying Assumption (B1), it suffices to show

$$n^{-1/2} \max_{\lfloor\sqrt{n}\rfloor \le k \le n} \mathbb{E}^* \sup_{x\in I} \left| \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \bar{\gamma}_n(x; \mathbf{Z}_i^*) \right| = o_{\mathbb{P}}(1)$$

where $\{\mathbf{Z}_i^*\}_{i=1}^k$ denotes a random sample from the empirical cdf and $\mathbb{E}^*$ is the expectation under this empirical bootstrap law. Corollary 4.3 of Pollard (1989) implies the desired result.

For $\widehat{\Phi}^*_n$, $\sup_{x\in I} |\widehat{\Phi}^*_n(x) - \widehat{\Phi}_n(x)| = o_{\mathbb{P}}(1)$ follows from the same argument as for $\widehat{\Gamma}^*_n$. For $a_n \sup_{x\in I_{\mathsf{x}}^\delta} |\widehat{\Phi}^*_n(x) - \widehat{\Phi}_n(x)| = o_{\mathbb{P}}(1)$,

$$\begin{aligned}\sup_{x\in I_{\mathsf{x}}^\delta} \left|\widehat{\Phi}^*_n(x) - \widehat{\Phi}_n(x)\right| \le& \sup_{x\in I_{\mathsf{x}}^\delta} \left|\widehat{\Phi}^*_n(x) - \bar{\Phi}^*_n(x)\right| + \sup_{x\in I_{\mathsf{x}}^\delta} \left|\bar{\Phi}^*_n(x) - \bar{\Phi}_n(x)\right| \\ &+ \sup_{x\in I_{\mathsf{x}}^\delta} \left|\bar{\Phi}_n(x) - \Phi_0(x)\right| + \sup_{x\in I_{\mathsf{x}}^\delta} \left|\widehat{\Phi}_n(x) - \Phi_0(x)\right|\end{aligned}$$

where the last term is $o_{\mathbb{P}}(a_n^{-1})$ as shown above and the second and third terms after the inequality are $O_{\mathbb{P}}(n^{-1/2})$ by standard arguments. For the remaining term,

$$\begin{aligned}\sup_{x\in I_{\mathsf{x}}^\delta} \left|\widehat{\Phi}^*_n(x) - \bar{\Phi}^*_n(x)\right| \le& \sup_{x\in I_{\mathsf{x}}^\delta} \left| \frac{1}{n} \sum_{i=1}^{n} W_{i,n} \breve{\phi}_n(x; \mathbf{Z}_i) \right| + \left|\breve{\Phi}_n(\mathsf{x}) - \bar{\Phi}_n(\mathsf{x})\right| \\ &+ \sup_{x\in I_{\mathsf{x}}^\delta} \left|\breve{\Phi}_n(x) - \breve{\Phi}_n(\mathsf{x}) - \bar{\Phi}_n(x) + \bar{\Phi}_n(\mathsf{x})\right|\end{aligned}$$

where $\breve{\phi}_n(x; \mathbf{Z}) = \widehat{\phi}_n(x; \mathbf{Z}) - \phi_0(x; \mathbf{Z}) - [\breve{\Phi}_n(x) - \bar{\Phi}_n(x)]$. The last two terms are $o_{\mathbb{P}}(a_n^{-1})$ by Assumption (E4). Using Lemma 3.6.7 of van der Vaart and Wellner (1996) and the argument similar to above, the remaining term is $o_{\mathbb{P}}(a_n^{-1})$.

SA.1.9. *Proof of Theorem A.1*. The proof is by contradiction and follows Kosorok (2008). We omit some details in cases where the arguments are almost identical to those for Theorem 1 and Lemma 3.

Suppose that the bootstrap approximation is consistent; that is, suppose

$$r_n\big(\widehat{\theta}^*_n(\mathsf{x}) - \widehat{\theta}_n(\mathsf{x})\big) \rightsquigarrow_{\mathbb{P}} Y, \qquad Y = (\partial\Phi_0(\mathsf{x}))^{-1}\partial_-\mathsf{GCM}_{\mathbb{R}}\{\mathcal{G}_{\mathsf{x}} + \mathcal{M}_{\mathsf{x}}^{\mathfrak{q}}\}(0).$$

Then, by Theorem 2.2 of Kosorok (2008), we have

$$r_n\big(\widehat{\theta}^*_n(\mathsf{x}) - \theta_0(\mathsf{x})\big) \rightsquigarrow Y_1 + Y_2 =_d \sqrt{2}Y \tag{SA.16}$$

where $=_d$ denotes the distributional equality, Y_1 and Y_2 are independent copies of Y, and where the convergence in distribution is unconditional.

Using the switching lemma, $\mathbb{P}[r_n\big(\widehat{\theta}^*_n(\mathsf{x}) - \theta_0(\mathsf{x})\big) > t]$ equals

$$\mathbb{P}\left[\max \operatorname*{argmax}_{x\in\widehat{\Phi}^{*-}_n([0,\widehat{u}^*_n])} \left\{[\theta_0(\mathsf{x}) + r_n^{-1}t]\widehat{\Phi}^*_n(x) - \mathsf{LSC}_{\widehat{\Phi}^*_n}\left(\widehat{\Gamma}^*_n\right)(x)\right\} < \widehat{\Phi}^{*-}_n \circ \widehat{\Phi}^*_n(\mathsf{x})\right].$$

By the arguments used in the proof of Theorem 1, to characterize the limiting distribution of $r_n(\widehat{\theta}^*_n(\mathsf{x}) - \theta_0(\mathsf{x}))$, it suffices to look at

$$-\breve{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v) - \breve{M}^{\mathfrak{q}}_{\mathsf{x},n}(v) + t\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n}(v)$$

where

$$\check{G}^{\mathfrak{q},*}_{\mathsf{x},n}(v) = \sqrt{na_n}\big[\widehat{\Gamma}^*_n(\mathsf{x}+va_n^{-1}) - \widehat{\Gamma}^*_n(\mathsf{x}) - \check{\Gamma}_n(\mathsf{x}+va_n^{-1}) + \check{\Gamma}_n(\mathsf{x})\big]$$
$$- \theta_0(\mathsf{x})\sqrt{na_n}\big[\widehat{\Phi}^*_n(\mathsf{x}+va_n^{-1}) - \widehat{\Phi}^*_n(\mathsf{x}) - \check{\Phi}_n(\mathsf{x}+va_n^{-1}) + \check{\Phi}_n(\mathsf{x})\big]$$

and

$$\check{M}^{\mathfrak{q}}_{\mathsf{x},n}(v) = \sqrt{na_n}\big[\check{\Gamma}_n(\mathsf{x}+va_n^{-1}) - \check{\Gamma}_n(\mathsf{x}) - \theta_0(\mathsf{x})\{\check{\Phi}_n(\mathsf{x}+va_n^{-1}) - \check{\Phi}_n(\mathsf{x})\}\big].$$

It can be shown that $\check{G}^{\mathfrak{q},*}_{\mathsf{x},n} \rightsquigarrow_{\mathbb{P}} \mathcal{G}_{\mathsf{x}}$, $\widehat{L}^{\mathfrak{q},*}_{\mathsf{x},n} \rightsquigarrow \mathcal{L}_{\mathsf{x}}$, and that $\check{M}^{\mathfrak{q}}_{\mathsf{x},n} \rightsquigarrow \mathcal{G}_{\mathsf{x}} + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}$. Thus,

$$\mathbb{P}\big[r_n\big(\widehat{\theta}^*_n(\mathsf{x}) - \theta_0(\mathsf{x})\big) > t\big] \to \mathbb{P}\Big[\operatorname*{argmin}_{v\in\mathbb{R}}\Big\{\mathcal{G}_{\mathsf{x},1}(v) + \mathcal{G}_{\mathsf{x},2}(v) + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}(v) - t\partial\Phi_0(\mathsf{x})v\Big\} < 0\Big]$$

where $\mathcal{G}_{\mathsf{x},1}$ and $\mathcal{G}_{\mathsf{x},2}$ are independent copies of $\mathcal{G}_{\mathsf{x}}$. Noting that $\mathcal{G}_{\mathsf{x}}(av) =_d \sqrt{|a|}\mathcal{G}_{\mathsf{x}}(v)$ and using the change of variable $v = u2^{\frac{1}{2\mathfrak{q}+1}}$, the limit distribution equals

$$\mathbb{P}\Big[2^{\frac{1}{2\mathfrak{q}+1}}\operatorname*{argmin}_{u\in\mathbb{R}}\Big\{\mathcal{G}_{\mathsf{x}}(u) + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}(u) - 2^{-\frac{\mathfrak{q}}{2\mathfrak{q}+1}}t\partial\Phi_0(\mathsf{x})u\Big\} < 0\Big]$$
$$= \mathbb{P}\Big[2^{\frac{\mathfrak{q}}{2\mathfrak{q}+1}}(\partial\Phi_0(\mathsf{x}))^{-1}\partial_-\mathsf{GCM}_{\mathbb{R}}\{\mathcal{G}_{\mathsf{x}} + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}\}(0) > t\Big].$$

As a consequence,

$$r_n\big(\widehat{\theta}^*_n(\mathsf{x}) - \theta_0(\mathsf{x})\big) \rightsquigarrow 2^{\frac{\mathfrak{q}}{2\mathfrak{q}+1}}(\partial\Phi_0(\mathsf{x}))^{-1}\partial_-\mathsf{GCM}_{\mathbb{R}}\{\mathcal{G}_{\mathsf{x}} + \mathcal{M}^{\mathfrak{q}}_{\mathsf{x}}\}(0),$$

contradicting (SA.16) because $2^{\frac{\mathfrak{q}}{2\mathfrak{q}+1}} \neq \sqrt{2}$.

In other words, the bootstrap estimator $\widehat{\theta}^*_n(\mathsf{x})$ fails to approximate the limit distribution. □

SA.1.10. *Remarks on verifying conditions in applications.* Below we verify the hypothesis of Theorem 1 for various examples. For this purpose, one should verify Assumptions A, (B5)-(B8), and E since Assumption E implies (B1)-(B4) by Lemma 3.

When γ_0 is known, it is natural to take $\widehat{\Gamma}_n = \check{\Gamma}_n = \bar{\Gamma}_n$, in which case (E3) reduces to the requirement that, for some $\rho_\gamma \in (0,2)$,

(SA.17)
$$\limsup_{\varepsilon\downarrow 0}\frac{\log N_U(\varepsilon, \mathfrak{F}_\gamma)}{\varepsilon^{-\rho_\gamma}} < \infty, \quad \mathbb{E}[\bar{F}_\gamma(\mathbf{Z})^2] < \infty, \quad \limsup_{\eta\downarrow 0}\frac{\mathbb{E}[\bar{D}^\eta_\gamma(\mathbf{Z})^2 + \bar{D}^\eta_\gamma(\mathbf{Z})^4]}{\eta} < \infty.$$

An identical remark applies to ϕ_0 and (E4).

In addition, as remarked in the main paper after Lemma 3, the second display of (B5) follows from the second display of (E5), and the first display of (B5) follows from

(SA.18)
$$\lim_{n\to\infty}\eta_n^{-1}\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})] = \mathcal{C}_{\mathsf{x}}(s,t)$$

for $a_n\eta_n = O(1)$ and any $\mathsf{x}_n \to \mathsf{x}$. To see the second claim,

$$\eta_n^{-1}\big\{\mathbb{E}[\psi_{\mathsf{x}}((s+t)\eta_n;\mathbf{Z})\psi_{\mathsf{x}}((s+t)\eta_n;\mathbf{Z})] - 2\mathbb{E}[\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}}((s+t)\eta_n;\mathbf{Z})]$$
$$+ \mathbb{E}[\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})]\big\} \to \mathcal{C}_{\mathsf{x}}(s+t,s+t) - \mathcal{C}_{\mathsf{x}}(s+t,s) - \mathcal{C}_{\mathsf{x}}(s,s+t) + \mathcal{C}_{\mathsf{x}}(s,s)$$

and at the same time, setting $\mathsf{x}_n = \mathsf{x} + s\eta_n$,

$$\eta_n^{-1}\big\{\mathbb{E}[\psi_{\mathsf{x}}((s+t)\eta_n;\mathbf{Z})\psi_{\mathsf{x}}((s+t)\eta_n;\mathbf{Z})] - 2\mathbb{E}[\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}}((s+t)\eta_n;\mathbf{Z})]$$
$$+ \mathbb{E}[\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})]\big\}$$
$$= \eta_n^{-1}\big\{\mathbb{E}[\{\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z}) - \psi_{\mathsf{x}_n}(-s\eta_n;\mathbf{Z})\}\{\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z}) - \psi_{\mathsf{x}_n}(-s\eta_n;\mathbf{Z})\}]$$
$$+ 2\mathbb{E}[\psi_{\mathsf{x}_n}(-s\eta_n;\mathbf{Z})\{\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z}) - \psi_{\mathsf{x}_n}(-s\eta_n;\mathbf{Z})\}] + \mathbb{E}[\psi_{\mathsf{x}_n}(-s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(-s\eta_n;\mathbf{Z})]\big\}$$
$$= \eta_n^{-1}\mathbb{E}[\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})] \to \mathcal{C}_{\mathsf{x}}(t,t)$$

and thus, $\mathcal{C}_{\mathsf{x}}(s+t,s+t)-\mathcal{C}_{\mathsf{x}}(s+t,s)-\mathcal{C}_{\mathsf{x}}(s,s+t)+\mathcal{C}_{\mathsf{x}}(s,s)=\mathcal{C}_{\mathsf{x}}(t,t)$ holds. Thus, for the two displays in (B5), it suffices to check (E5) and (SA.18).

SA.1.11. *Proof of Corollary 1.* Assumption A and (E1)-(E2) follow from the hypothesis.

(E3). In this example, $\gamma_0(x;\mathbf{Z})=\mathbb{1}\{X\leq x\}$ is known, so it suffices to verify (SA.17). The uniform covering number of $\{\mathbb{1}\{\cdot\leq x\}:x\in\mathbb{R}\}$ grows linearly, and an envelope function can be taken to be 1. For an envelope function of $\{\mathbb{1}\{\cdot\leq x\}-\mathbb{1}\{\cdot\leq \mathsf{x}\}:|x-\mathsf{x}|\leq\eta\}$, we can take $\mathbb{1}\{-\eta+\mathsf{x}\leq\cdot\leq\mathsf{x}+\eta\}$ and the moment bound is satisfied as $\mathbb{E}[\mathbb{1}\{-\eta+\mathsf{x}\leq X\leq \mathsf{x}+\eta\}]\leq C\eta$.

(E4) trivially holds as $\widehat{\Phi}_n(x)=\widehat{\Phi}_n^*(x)=x$.

(E5). Here $\psi_{\mathsf{x}}(v;\mathbf{Z})=\mathbb{1}\{X\leq\mathsf{x}+v\}-\mathbb{1}\{X\leq\mathsf{x}\}-\Phi_0(\mathsf{x})v$. Then,

$$\frac{\mathbb{E}[|\psi_{\mathsf{x}}(v;\mathbf{Z})-\psi_{\mathsf{x}}(v';\mathbf{Z})|]}{|v-v'|}\leq\frac{\mathbb{E}[\mathbb{1}\{\mathsf{x}+v\wedge v'<X\leq\mathsf{x}+v\vee v'\}]}{|v-v'|}+f_0(\mathsf{x})\leq C.$$

Also, $\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})=\mathbb{1}\{\mathsf{x}_n\wedge(\mathsf{x}_n+s\eta_n)\}<X\leq\mathsf{x}_n\vee(\mathsf{x}_n+s\eta_n)\}-f_0(\mathsf{x})s\eta_n$ and

$$\begin{aligned}\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})=&\ \mathbb{1}\{\mathsf{x}_n<X\leq\mathsf{x}_n+\eta_n(s\wedge t)\}\mathbb{1}\{s>0,t>0\}\\&+\mathbb{1}\{\mathsf{x}_n+(s\vee t)<X\leq\mathsf{x}_n\}\mathbb{1}\{s<0,t<0\}\\&-\mathbb{1}\{\mathsf{x}_n\wedge(\mathsf{x}_n+s\eta_n)<X\leq\mathsf{x}_n\vee(\mathsf{x}_n+s\eta_n)\}f_0(\mathsf{x}_n)t\eta_n\\&-\mathbb{1}\{\mathsf{x}_n\wedge(\mathsf{x}_n+t\eta_n)<X\leq\mathsf{x}_n\vee(\mathsf{x}_n+t\eta_n)\}f_0(\mathsf{x}_n)s\eta_n\\&+f_0(\mathsf{x}_n)^2st\eta_n^2.\end{aligned}$$

Then, for any $s,t\in\mathbb{R}$ and $\mathsf{x}_n\to\mathsf{x}$, using continuity of f_0 at x,

$$\begin{aligned}&\eta_n^{-1}\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})]\\&=\eta_n^{-1}\int_{\mathsf{x}_n}^{\mathsf{x}_n+\eta_n(s\wedge t)}f_0(u)du\mathbb{1}\{s>0,t>0\}\\&\quad+\eta_n^{-1}\int_{\mathsf{x}_n+\eta_n(s\vee t)}^{\mathsf{x}_n}f_0(u)du\mathbb{1}\{s<0,t<0\}+o(1)\\&=f_0(\mathsf{x})[(s\wedge t)\mathbb{1}\{s>0,t>0\}-(s\vee t)\mathbb{1}\{s<0,t<0\}]+o(1)\\&=f_0(\mathsf{x})(|s|\wedge|t|)\mathbb{1}\{\mathrm{sign}(s)=\mathrm{sign}(t)\}+o(1).\end{aligned}$$

(B5). It is clear that $\mathcal{C}_{\mathsf{x}}(1,1)>0$ from $f_0(\mathsf{x})>0$. Also, $\mathcal{C}_{\mathsf{x}}(1,\eta)/\sqrt{\eta}=f_0(\mathsf{x})\sqrt{\eta}\mathbb{1}\{\eta>0\}$ for $|\eta|<1$ and $\limsup_{\eta\downarrow 0}\mathcal{C}_{\mathsf{x}}(1,\eta)/\sqrt{\eta}=0$ holds. The remaining conditions follow from verifying (E5) above.

(B6) holds since $\widehat{u}_n=\widehat{u}_n^*$ converges in probability to u_0, the supremum of the support of X by i.i.d. assumption.

(B7) and (B8) hold trivially since $\widehat{\Phi}_n$ and $\widehat{\Phi}_n^*$ are the identity map.

Assumption D follows from (E3) and empirical process theory arguments.

SA.1.12. *Proof of Corollary 2.* Assumption A and (E1)-(E2) follow from the hypothesis.

(E3). We have $\widehat{\Gamma}_n = 1 - \widehat{S}_n$ with $\widehat{S}_n$ the Kaplan-Meier estimator. By Theorem 1 of Lo and Singh (1986),

$$\sup_{x \in I} \Big| \widehat{\Gamma}_n(x) - \frac{1}{n}\sum_{i=1}^n \gamma_0(x; \mathbf{Z}) \Big| = O_{\mathbb{P}}\Big(\Big| \frac{\log n}{n} \Big|^{3/4} \Big).$$

Since $\sqrt{na_n} \leq n^{2/3}$ for $\mathfrak{q} \geq 1$, $\sup_{x \in I} |\widehat{\Gamma}_n(x) - \Gamma_0(x)| = o_{\mathbb{P}}(1)$ and $\sqrt{na}\sup_{|v|\leq\delta} |\widehat{\Gamma}_n(\mathsf{x} + v) - \widehat{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+v) + \bar{\Gamma}_n(\mathsf{x})| = o_{\mathbb{P}}(1)$ hold.

We have

$$\begin{aligned}
&\widehat{\gamma}_n(x; \mathbf{Z}) - \gamma_0(x; \mathbf{Z}) \\
&= \widehat{F}_n(x) - F_0(x) + \big[\widehat{S}_n(x) - S_0(x)\big]\Bigg[\frac{\mathbb{1}\{\check{X} \leq x\}\Delta}{\widehat{S}_n(\check{X})\widehat{G}_n(\check{X})} - \int_0^{\check{X}\wedge x} \frac{\widehat{\Lambda}_n(du)}{\widehat{S}_n(u)\widehat{G}_n(u)}\Bigg] \\
&\quad + S_0(x)\mathbb{1}\{\check{X} \leq x\}\Delta \frac{S_0(\check{X})G_0(\check{X}) - \widehat{S}_n(\check{X})\widehat{G}_n(\check{X})}{S_0(\check{X})G_0(\check{X})\widehat{S}_n(\check{X})\widehat{G}_n(\check{X})} \\
&\quad - S_0(x)\int_0^{\check{X}\wedge x} \frac{S_0(u)G_0(u) - \widehat{S}_n(u)\widehat{G}_n(u)}{S_0(u)G_0(u)\widehat{S}_n(u)\widehat{G}_n(u)}\widehat{\Lambda}_n(du) - S_0(x)\int_0^{\check{X}\wedge x} \frac{[\widehat{\Lambda}_n - \Lambda_0](du)}{S_0(u)G_0(u)}.
\end{aligned}$$

By $S_0(u_0)G_0(u_0) > 0$, we have $\sqrt{n}\sup_{x\in I} |\widehat{S}_n(x) - S_0(x)| = O_{\mathbb{P}}(1)$, $\sqrt{n}\sup_{x \in I} |\widehat{G}_n(x) - G_0(x)| = O_{\mathbb{P}}(1)$, and $\sqrt{n}\sup_{x\in I} |\widehat{\Lambda}_n(x) - \Lambda_0(x)| = O_{\mathbb{P}}(1)$, which in turn implies

$$\frac{1}{n}\sum_{i=1}^n \sup_{x\in I} \big|\widehat{\gamma}_n(x; \mathbf{Z}_i) - \gamma_0(x; \mathbf{Z}_i)\big|^2 = o_{\mathbb{P}}(1), \; a_n \frac{1}{n}\sum_{i=1}^n \sup_{x \in I_{\mathsf{x}}^\delta} \big|\widehat{\gamma}_n(x; \mathbf{Z}_i) - \gamma_0(x; \mathbf{Z}_i)\big|^2 = o_{\mathbb{P}}(1).$$

Let $\delta = \min\{\mathsf{x}, (u_0 - \mathsf{x})\}/4$, $R_{1n}(v) = |\widehat{F}_n(\mathsf{x} + v) - \widehat{F}_n(\mathsf{x}) - F_0(\mathsf{x}+v) + F_0(\mathsf{x})|$, $R_{2n} = |\widehat{F}_n(\mathsf{x}) - F_0(\mathsf{x})|$, $R_{3n} = \sup_{x \in I} |[\widehat{S}_n(x)\widehat{G}_n(x)]^{-1} - [S_0(x)G_0(x)]^{-1}|$, and $R_{4n}(x_1, x_2) = |\int_{x_1}^{x_2} \frac{\widehat{\Lambda}_n(du)}{\widehat{S}_n(u)\widehat{G}_n(u)} - \int_{x_1}^{x_2} \frac{\Lambda_0(du)}{S_0(u)G_0(u)}|$. For $|v| \leq \delta$,

$$\begin{aligned}
&\big|\check{\Gamma}_n(\mathsf{x} + v) - \check{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+v) + \bar{\Gamma}_n(\mathsf{x})\big| \\
&\leq R_{1n}(v)\left[1 + \frac{1 + \widehat{\Lambda}_n(x)}{\widehat{S}_n(x)\widehat{G}_n(x)}\right] + R_{2n}\big[\widehat{S}_n(\mathsf{x} - \delta)\widehat{G}_n(\mathsf{x}-\delta)\big]^{-1}\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{|X_i - \mathsf{x}| \leq v\} \\
&+ R_{2n}\int_{\mathsf{x}-v}^{\mathsf{x}+v} \frac{\widehat{\Lambda}_n(du)}{\widehat{S}_n(u)\widehat{G}_n(u)} + \sup_{x \in I_{\mathsf{x}}^\delta} f_0(x)|v|\left[R_{3n} + \frac{1}{n}\sum_{i=1}^n R_{4n}(0, \check{X}_i \wedge (\mathsf{x}+v))\right] \\
&+ R_{3n}\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{|X_i - \mathsf{x}| \leq v\} + \frac{1}{n}\sum_{i=1}^n R_{4n}(\check{X}_i \wedge \mathsf{x}, \check{X}_i \wedge (\mathsf{x}+v)).
\end{aligned}$$

As noted above, $\sup_{|v|\leq\delta} R_{1n}(v) = o_{\mathbb{P}}((na_n)^{-1/2})$, $R_{2n} = O_{\mathbb{P}}(n^{-1/2})$, and $R_{3n} = O_{\mathbb{P}}(n^{-1/2})$ by $S_0(u_0)G_0(u_0) > 0$. Also, uniformly over $V \in (0, 2\delta]$, $\sup_{|v|\leq V} |\frac{1}{n}\sum_{i=1}^n (\mathbb{1}\{X_i \leq \mathsf{x}+v\} - \mathbb{1}\{X_i \leq \mathsf{x}\})| \leq CV + O_{\mathbb{P}}(n^{-1/2})$. If

$$\sup_{|v|\leq\delta} \frac{1}{n}\sum_{i=1}^n R_{4n}(0, \check{X}_i \wedge (\mathsf{x}+v)) = O_{\mathbb{P}}(n^{-1/2}) \tag{SA.19}$$

and

$$\frac{1}{n}\sum_{i=1}^n R_{4n}(\check{X}_i \wedge \mathsf{x}, \check{X}_i \wedge (\mathsf{x}+v)) \leq |v| O_{\mathbb{P}}(n^{-1/2}) + o_{\mathbb{P}}\Big((na_n)^{-1/2}\Big) \tag{SA.20}$$

uniformly over $|v| \le 2\delta$, then there exists random variables $A_n = o_{\mathbb{P}}(1)$ and $B_n = O_{\mathbb{P}}(\sqrt{a_n})$ independent of v such that for $V \in (0, 2\delta]$,

$$\sqrt{na_n} \sup_{|v|\le V} \left|\check{\Gamma}_n(\mathsf{x}+v) - \check{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+v) + \bar{\Gamma}_n(\mathsf{x})\right| \le A_n + VB_n$$

i.e., $\beta_\gamma = 1$ in the notation of (E3). To show (SA.19) and (SA.20), for $x_1, x_2 \in I$,

$$R_{4n}(x_1,x_2) \le R_{3n}\big|\widehat{\Lambda}_n(x_2) - \widehat{\Lambda}_n(x_1)\big| + \left|\int_{x_1}^{x_2} \frac{[\widehat{\Lambda}_n - \Lambda_0](du)}{S_0(u)G_0(u)}\right|.$$

Let $J_0(u) = 1/[S_0(u)G_0(u)]$ and integration by parts implies

$$\begin{aligned}\int_{x_1}^{x_2} \frac{[\widehat{\Lambda}_n - \Lambda_0](du)}{S_0(u)G_0(u)} &= \frac{\widehat{\Lambda}_n(x_2) - \Lambda_0(x_2)}{S_0(x_2)G_0(x_2)} - \frac{\widehat{\Lambda}_n(x_1) - \Lambda_0(x_1)}{S_0(x_1)G_0(x_1)} - \int_{x_1}^{x_2} [\widehat{\Lambda}_n(u) - \Lambda_0(u)]J_0(du)\\ &= [\widehat{\Lambda}_n(x_2) - \Lambda_0(x_2)]\left[\frac{1}{S_0(x_2)G_0(x_2)} - \frac{1}{S_0(x_1)G_0(x_1)}\right]\\ &\quad + \frac{\widehat{\Lambda}_n(x_2) - \Lambda_0(x_2) - \widehat{\Lambda}_n(x_1) + \Lambda_0(x_1)}{S_0(x_1)G_0(x_1)} - \int_{x_1}^{x_2} [\widehat{\Lambda}_n(u) - \Lambda_0(u)]J_0(du).\end{aligned}$$

The first term after the second equality is bounded by $O_{\mathbb{P}}(n^{-1/2})|x_2 - x_1|$. For the second term, Theorem 1 of Burke, Csörgő and Horváth (1988) implies that on a suitable probability space there exists a sequence of standard Brownian motion W_n such that $\sqrt{a_n}\sup_{|x_1-x_2|\le v} |\sqrt{n}[\widehat{\Lambda}_n(x_1) - \Lambda_0(x_2) - \widehat{\Lambda}_n(x_1) + \Lambda_0(x_2)] - W_n(d(x_1)) + W_n(d(x_2))| = o_{\mathbb{P}}(1)$, where $d(x) = \int_0^x \frac{F_0(du)}{S_0(u)^2G_0(u)}$. By Theorem 3.2 of Pollard (1989), there is some fixed constant $C > 0$ such that $\mathbb{E}[\sup_{|x_1-x_2|\le v} |W_n(d(x_1)) - W_n(d(x_2))|] \le Cv$. Finally, $|\int_{x_1}^{x_2}[\widehat{\Lambda}_n(u) - \Lambda_0(u)]J_0(du)| \le \sup_{x'\in[x_1,x_2]} |\widehat{\Lambda}_n(x') - \Lambda_0(x')|[J_0(x_2) - J_0(x_1)] \le O_{\mathbb{P}}(n^{-1/2})|x_2 - x_1|$. Thus, (SA.19) and (SA.20) hold.

For the function class $\mathfrak{F}_\gamma$, we can take $\bar{F}_\gamma(\mathbf{Z}) = 1 + [S_0(u_0)G_0(u_0)]^{-1}[1 + \Lambda_0(u_0)]$ as a constant envelope. For the function class $\{S_0(x) : x \in I\}$, given $m \in \mathbb{N}$, there exists $\{x_1, \dots, x_{m+1}\} \subset I$ such that $\sup_{x\in I} \min_{l=1,\dots,m+1} |S_0(x_l) - S_0(x)| \le 1/m$, which implies the uniform covering number is bounded by a linear function. The covering numbers of $\{\mathbb{1}\{\cdot \le s\} : s \in I\}$ and $\{\int_0^{\cdot\wedge s}[S_0(u)G_0(u)]^{-1}\Lambda_0(du) : s \in I\}$ are also bounded by a linear function. By Lemma 5.1 of van der Vaart and van der Laan (2006), there exists $\rho \in (0, 2)$ such that $\limsup_{\eta\downarrow 0} \log N_U(\eta, \mathfrak{F}_\gamma)\eta^\rho < \infty$ holds.

Now consider the uniform covering number of $\hat{\mathfrak{F}}_\gamma$. Given a realization of $(\widehat{S}_n, \widehat{G}_n)$, the mapping $x \mapsto \int_0^{x\wedge s}[\widehat{S}_n(u)\widehat{G}_n(u)]^{-1}\widehat{\Lambda}_n(du)$ is a composition of $x \mapsto x \wedge s$ and $x \mapsto \int_0^x[\widehat{S}_n(u)\widehat{G}_n(u)]^{-1}\widehat{\Lambda}_n(du)$. The latter mapping is monotone, and the first mapping is a VC-subgraph class, and Lemma 2.6.18 of van der Vaart and Wellner (1996) implies $\{\int_0^{\cdot\wedge s}[\widehat{S}_n(u)\widehat{G}_n(u)]^{-1}\widehat{\Lambda}_n(du) : s \in I\}$ is a VC-subgraph class. Note that since S_0, G_0 are bounded away from zero, $\widehat{S}_n, \widehat{G}_n$ are bounded away from zero with probability approaching one. Thus, for some $\rho \in (0, 2)$, $\limsup_{\eta\downarrow 0} \log N_U(\eta, \hat{\mathfrak{F}}_\gamma)\eta^\rho = O_{\mathbb{P}}(1)$ holds.

For $s \le t \in I$,

$$|\gamma_0(s;\mathbf{Z}) - \gamma_0(t;\mathbf{Z})| \le C|F_0(s) - F_0(t)| + C|\mathbb{1}\{\check{X} \le s\} - \mathbb{1}\{\check{X} \le t\}|\Delta + \int_{\check{X}\wedge s}^{\check{X}\wedge t} \frac{\Lambda_0(du)}{S_0(du)G_0(du)}$$

and we can take $D_\gamma^\eta(\mathbf{Z})$ to be a constant multiple of $\sup_{|s|\le\eta} |F_0(\mathsf{x}+s) - F_0(\mathsf{x})| + \Delta\mathbb{1}\{|\check{X} - \mathsf{x}| \le \eta\} + \int_{x-\eta}^{x+\eta} \Lambda_0(du)/S_0(u)G_0(u)$. For $\eta > 0$ small enough, there is some fixed $C > 0$ with

$$\mathbb{E}[D_\gamma^\eta(\mathbf{Z})^2 + D_\gamma^\eta(\mathbf{Z})^4] \le Cf_0(\mathsf{x}+\eta)\eta.$$

(E4) trivially holds as $\widehat{\Phi}_n(x) = \widehat{\Phi}_n^*(x) = x$.

(E5). We have

$$\psi_{\mathsf{x}}(v; \mathbf{Z}) = S_0(\mathsf{x}) \frac{(\mathbb{1}\{\check{X} \le \mathsf{x} + v\} - \mathbb{1}\{\check{X} \le \mathsf{x}\})\Delta}{S_0(\check{X}) G_0(\check{X})} + O(|v|)$$

where $O(|v|)$ is uniformly over small enough $|v|$. Since

$$\mathbb{E}[|\mathbb{1}\{\check{X} \le \mathsf{x} + v\} - \mathbb{1}\{\check{X} \le \mathsf{x} + v'\}|\Delta] = \int_{\mathsf{x}+v\wedge v'}^{\mathsf{x}+v\vee v'} G_0(u) f_0(u) du \le C|v - v'|,$$

the first display in (E5) is satisfied. For the covariance kernel,

$$\begin{aligned}
&\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n; \mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n; \mathbf{Z})] \\
&= S_0(\mathsf{x}_n)^2 \bigg(\int_{\mathsf{x}_n}^{\mathsf{x}_n + \eta_n(s\wedge t)} \frac{f_0(u)}{S_0(u)^2 G_0(u)} du \mathbb{1}\{s > 0, t > 0\} \\
&\qquad + \int_{\mathsf{x}_n + \eta_n(s\vee t)}^{\mathsf{x}_n} \frac{f_0(u)}{S_0(u)^2 G_0(u)} du \mathbb{1}\{s < 0, t < 0\} \bigg) + O(\eta_n^2) \\
&= \frac{S_0(\mathsf{x}_n)^2 f_0(\mathsf{x})}{S_0(\mathsf{x})^2 G_0(\mathsf{x})} (s \wedge t)\eta_n \mathbb{1}\{s > 0, t > 0\} \\
&\quad - \frac{S_0(\mathsf{x}_n)^2 f_0(\mathsf{x})}{S_0(\mathsf{x})^2 G_0(\mathsf{x})} (s \vee t)\eta_n \mathbb{1}\{s < 0, t < 0\} + o(\eta_n)
\end{aligned}$$

where the last equality uses continuity of (S_0, G_0, f_0) at x i.e., $\int_{\mathsf{x}_n}^{\mathsf{x}_n+\eta_n} [\frac{f_0(u)}{S_0(u)^2 G_0(u)} - \frac{f_0(\mathsf{x})}{S_0(\mathsf{x})^2 G_0(\mathsf{x})}] du = o(1)\eta_n$. Thus, (SA.18) holds.

(B5). $\mathcal{C}_{\mathsf{x}}(1,1) > 0$ follows from $f_0(\mathsf{x}) > 0$. $\lim_{\eta\downarrow 0} \mathcal{C}_{\mathsf{x}}(1,\eta)/\sqrt{\eta} = 0$ follows from the same computation as in the no censoring case. The remaining conditions follow from verifying (E5) above.

(B6), (B7), and (B8) hold since in this example, $\widehat{u}_n = \widehat{u}_n^* = u_0$ and $\widehat{\Phi}_n, \widehat{\Phi}_n^*$ are the identity map.

Assumption D. As noted when verifying (E3), $\widehat{G}_n(1; \eta_n) = o_{\mathbb{P}}(1)$ for any $\eta_n = o(1)$ with $a_n^{-1}\eta_n^{-1} = O(1)$. $\widehat{\Phi}_n = \Phi_0$ is the identity map and the desired result holds.

SA.1.13. *Proof of Corollary 3.* Assumption A and (E1)-(E2) follow from the hypothesis.

(E3). In this example, $\gamma_0(x; \mathbf{Z}) = Y\mathbb{1}\{X \le x\}$ is known, so it suffices to verify (SA.17). The uniform covering number bound is straightforward as $\{\mathbb{1}\{\cdot \le x\} : x \in \mathbb{R}\}$ is a VC-subgraph class. An envelope function is $|Y|$, whose second moment is finite. For $x \in I_{\mathsf{x}}^{\eta}$, $|\gamma_0(x; \mathbf{Z}) - \gamma_0(\mathsf{x}; \mathbf{Z})| \le |Y|\mathbb{1}\{\mathsf{x} - \eta \le X \le \mathsf{x} + \eta\}$, which we can take as $\bar{D}_{\gamma}^{\eta}(\mathbf{Z})$. Then, for $j = 2, 4$,

$$\mathbb{E}[\bar{D}_{\gamma}^{\eta}(\mathbf{Z})^j] \le 2^{j-1} \int_{\mathsf{x}-\eta}^{\mathsf{x}+\eta} \big(|\mu_0(x)|^j + \mathbb{E}[\varepsilon^j | X = x]\big) f_0(x) dx \le C\eta$$

and the desired bound holds.

(E4). $\phi_0(x; \mathbf{Z}) = \mathbb{1}\{X \le x\}$ is known, so it suffices to verify the analogue of (SA.17). The argument is the same as for checking (E3) in monotone density estimation with no censoring.

(E5). We have

$\psi_{\mathsf{x}}(v;\mathbf{Z}) = \varepsilon(\mathbb{1}\{X \le \mathsf{x}+v\} - \mathbb{1}\{X \le \mathsf{x}\}) + (\mu_0(X) - \mu_0(\mathsf{x}))(\mathbb{1}\{X \le \mathsf{x}+v\} - \mathbb{1}\{X \le \mathsf{x}\})$.

Then,

$$\mathbb{E}[|\psi_{\mathsf{x}}(v;\mathbf{Z}) - \psi_{\mathsf{x}}(v';\mathbf{Z})|] \le \int_{\mathsf{x}+v\wedge v'}^{\mathsf{x}+v\vee v'} [\sigma_0(x) + |\mu_0(x) - \mu_0(\mathsf{x})|] f_0(x)dx \le C|v - v'|$$

and the first display holds. For the covariance kernel, note $|(\mu_0(X) - \mu_0(\mathsf{x}_n))(\mathbb{1}\{X \le \mathsf{x}_n + v\} - \mathbb{1}\{X \le \mathsf{x}_n\}| \le |v| \sup_{|x-\mathsf{x}|\le 2\eta} |\partial\mu_0(x)|$ for $|x_n - \mathsf{x}| \vee |v| \le \eta$ for $\eta > 0$ small enough. Then,

$$\begin{aligned}
&\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})] \\
&= \mathbb{E}[\varepsilon^2(\mathbb{1}\{X \le \mathsf{x}_n + s\eta_n\} - \mathbb{1}\{X \le \mathsf{x}_n\})(\mathbb{1}\{X \le \mathsf{x}_n + t\eta_n\} - \mathbb{1}\{X \le \mathsf{x}_n\})] + O(\eta_n^2) \\
&= \int_{\mathsf{x}_n}^{\mathsf{x}_n+\eta_n(s\wedge t)} \sigma_0^2(x) f_0(x) dx \mathbb{1}\{s > 0, t > 0\} \\
&\quad + \int_{\mathsf{x}_n+\eta_n(s\vee t)}^{\mathsf{x}_n} \sigma_0^2(x) f_0(x) dx \mathbb{1}\{s < 0, t < 0\} + O(\eta_n^2)
\end{aligned}$$

and

$$\begin{aligned}
&\eta_n^{-1}\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})] \\
&\to \sigma_0^2(\mathsf{x}) f_0(\mathsf{x}) \big[(s\wedge t)\mathbb{1}\{s > 0, t > 0\} - (s \vee t)\mathbb{1}\{s < 0, t < 0\}\big] \\
&= \sigma_0^2(\mathsf{x}) f_0(\mathsf{x})(|s| \wedge |t|)\mathbb{1}\{\mathrm{sign}(s) = \mathrm{sign}(t)\},
\end{aligned}$$

as desired.

(B5). $\mathcal{C}_{\mathsf{x}}(1,1) > 0$ follows from $f_0(\mathsf{x})\sigma_0^2(\mathsf{x}) > 0$. $\lim_{\eta\downarrow 0} \mathcal{C}_{\mathsf{x}}(1,\eta)/\sqrt{\eta} = 0$ follows from the same computation as in the monotone density estimation. The remaining conditions follow from verifying (E5) above.

(B6) trivially holds since $\widehat{u}_n = \widehat{u}_n^* = 1$ in this example.

(B7). $\widehat{\Phi}_n, \widehat{\Phi}_n^*$ are (empirical) cdfs, so they are non-negative, non-decreasing, and right-continuous. $\{0, \widehat{u}_n\} \subset \widehat{\Phi}_n(I)$ and $\{0, \widehat{u}_n^*\} \subset \widehat{\Phi}_n^*(I)$ hold as $\widehat{u}_n = \widehat{u}_n^* = 1$, $\widehat{\Phi}_n(\min_i X_i-) = 0 = \widehat{\Phi}_n^*(\min_i X_i-)$, and $\widehat{\Phi}_n(\max_i X_i) = 1 = \widehat{\Phi}_n^*(\max_i X_i)$. The sets $\widehat{\Phi}_n(I), \widehat{\Phi}_n^*(I)$ are finite and thus closed.

(B8). With probability one, all X_i's are distinct. If x is one of X_i's, $\widehat{\Phi}_n^*(x) - \widehat{\Phi}_n^*(x-) = n^{-1}W_{j,n}$ for some j, and $\widehat{\Phi}_n^*(x) - \widehat{\Phi}_n^*(x-) = 0$ otherwise. Given $\mathbb{E}|W_{1,n}|^{\mathfrak{r}} < \infty$, $n^{-1}\max_{1\le i\le n} |W_{i,n}| = o_{\mathbb{P}}(n^{-5/6})$, which implies $\sqrt{na_n} \sup_{x\in I} |\widehat{\Phi}_n^*(x) - \widehat{\Phi}_n^*(x-)| = o_{\mathbb{P}}(1)$. The argument for $\widehat{\Phi}_n$ is similar, but simpler.

Assumption D follows from (E3)-(E4) and empirical process theory arguments.

SA.1.14. *Proof of Corollary 4.* Assumption A and (E1)-(E2) follow from the hypothesis.

(E3). In this example, $\check{\Gamma}_n = \widehat{\Gamma}_n$.

$$\begin{aligned}
|\widehat{\gamma}_n(x;\mathbf{Z}) - \gamma_0(x;\mathbf{Z})| \le \mathbb{1}\{X \le x\}\Bigg[&|\varepsilon|\Big|\widehat{g}_n(X,\mathbf{A})^{-1} - g_0(X,\mathbf{A})^{-1}\Big| + \frac{|\widehat{\mu}_n(X,\mathbf{A}) - \mu_0(X,\mathbf{A})|}{\widehat{g}_n(X,\mathbf{A})} \\
&+ \frac{1}{n}\sum_{j=1}^n |\widehat{\mu}_n(X,\mathbf{A}_j) - \mu_0(X,\mathbf{A}_j)| + \Bigg|\frac{1}{n}\sum_{j=1}^n \mu_0(X,\mathbf{A}_j) - \theta_0(X)\Bigg|\Bigg].
\end{aligned}$$

The last sum is bounded by $\sup_{x\in I}|\frac{1}{n}\sum_{j=1}^{n}\mu_0(x,\mathbf{A}_j)-\theta_0(x)|$, and this object is $O_{\mathbb{P}}(n^{-1/2})$: to see this claim, first note that Assumption MRC (iv) and Theorem 2.7.11 of van der Vaart and Wellner (1996) imply $\limsup_{\epsilon\downarrow 0}\log N_U(\epsilon,\{\mu(x,\cdot):x\in I\})\epsilon^V<\infty$ for some $V\in(0,2)$ and Theorem 4.2 of Pollard (1989) implies $\sup_{x\in I}|\frac{1}{n}\sum_{j=1}^{n}\mu_0(x,\mathbf{A}_j)-\theta_0(x)|=O_{\mathbb{P}}(n^{-1/2})$. Together with Assumption MRC (iii), $a_n\frac{1}{n}\sum_{i=1}^{n}\sup_{x\in I}|\widehat{\gamma}_n(x;\mathbf{Z})-\gamma_0(x;\mathbf{Z})|^2=o_{\mathbb{P}}(1)$ holds.

By Assumption MRC (iii), uniformly over $V\in(0,2\delta]$

$$\sqrt{na_n}\sup_{|v|\leq V}\left|\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+v)+\bar{\Gamma}_n(\mathsf{x})\right|\leq o_{\mathbb{P}}(1)+O_{\mathbb{P}}(\sqrt{a_n})V$$

and the desired inequality holds.

The uniform covering numbers of $\mathfrak{F}_\gamma,\hat{\mathfrak{F}}_{\gamma,n}$ are the same order as for $\{\mathbb{1}\{\cdot\leq x\}:x\in I\}$. For $x\in I_{\mathsf{x}}^{\eta}$, $|\gamma_0(x;\mathbf{Z})-\gamma_0(\mathsf{x};\mathbf{Z})|\leq\mathbb{1}\{\mathsf{x}-\eta\leq X\leq\mathsf{x}+\eta\}(|\varepsilon|c^{-1}+\theta_0(\mathsf{x}+\eta))$. Then, $\limsup_{\eta\downarrow 0}\mathbb{E}[\bar{D}_{\gamma}^{\eta}(\mathbf{Z})^j]\eta^{-1}<\infty$ holds for $j=2,4$.

(E4). $\phi_0(x;\mathbf{Z})=\mathbb{1}\{X\leq x\}$ is known and the same as in the classical case, so the same argument applies.

(E5). We have

$$\psi_{\mathsf{x}}(v;\mathbf{Z})=(\mathbb{1}\{X\leq\mathsf{x}+v\}-\mathbb{1}\{X\leq\mathsf{x}\})\left[\frac{\varepsilon}{g_0(X,\mathbf{A})}+\theta_0(X)-\theta_0(\mathsf{x})\right].$$

Then, for $v,v'\in[-\eta,\eta]$ with sufficiently small $\eta>0$,

$$|\psi_{\mathsf{x}}(v;\mathbf{Z})-\psi_{\mathsf{x}}(v';\mathbf{Z})|\leq|\mathbb{1}\{X\leq\mathsf{x}+v\}-\mathbb{1}\{X\leq\mathsf{x}+v'\}|\big(c^{-1}|\varepsilon|+|X-\mathsf{x}|\sup_{x\in I_{\mathsf{x}}^{\eta}}|\partial\theta_0(x)|\big)$$

and $\sup_{v\neq v'\in[-\eta_n,\eta_n]}\mathbb{E}[|\psi_{\mathsf{x}}(v;\mathbf{Z})-\psi_{\mathsf{x}}(v';\mathbf{Z})|]/|v-v'|=O(1)$ holds.

For $s\eta_n$ small enough, $\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})=(\mathbb{1}\{X\leq\mathsf{x}+s\eta_n\}-\mathbb{1}\{X\leq\mathsf{x}\})\varepsilon g_0(X,\mathbf{A})^{-1}+O(\eta_n)$ and

$$\begin{aligned}&\mathbb{E}[\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}}(t\eta_n;\mathbf{Z})]\\&=\mathbb{E}\left[\frac{\sigma_0^2(X,\mathbf{A})}{g_0(X,\mathbf{A})^2}(\mathbb{1}\{X\leq\mathsf{x}+s\eta_n\}-\mathbb{1}\{X\leq\mathsf{x}\})(\mathbb{1}\{X\leq\mathsf{x}+t\eta_n\}-\mathbb{1}\{X\leq\mathsf{x}\})\right]+O(\eta_n^2)\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left[\frac{\sigma_0^2(X,\mathbf{A})}{g_0(X,\mathbf{A})^2}(\mathbb{1}\{X\leq\mathsf{x}+s\eta_n\}-\mathbb{1}\{X\leq\mathsf{x}\})(\mathbb{1}\{X\leq\mathsf{x}+t\eta_n\}-\mathbb{1}\{X\leq\mathsf{x}\})\right]\\&=\mathbb{E}\left[\int_{\mathsf{x}}^{\mathsf{x}+\eta_n(s\wedge t)}\frac{\sigma_0^2(x,\mathbf{A})}{g_0(x,\mathbf{A})^2}f_{X|\mathbf{A}}(x|\mathbf{A})dx\right]\mathbb{1}\{s>0,t>0\}\\&\quad+\mathbb{E}\left[\int_{\mathsf{x}+\eta_n(s\vee t)}^{\mathsf{x}}\frac{\sigma_0^2(x,\mathbf{A})}{g_0(x,\mathbf{A})^2}f_{X|\mathbf{A}}(x|\mathbf{A})dx\right]\mathbb{1}\{s<0,t<0\}\\&=\mathbb{E}\left[\frac{\sigma_0^2(\mathsf{x},\mathbf{A})}{g_0(\mathsf{x},\mathbf{A})^2}f_{X|\mathbf{A}}(\mathsf{x}|\mathbf{A})\right]\Big(\eta_n(s\wedge t)\mathbb{1}\{s>0,t>0\}-\eta_n(s\vee t)\mathbb{1}\{s<0,t<0\}\Big)+o(\eta_n).\end{aligned}$$

Since $\frac{f_{X|\mathbf{A}}(x|\mathbf{A})}{g_0(x,\mathbf{A})^2}=\frac{f_0(x)}{g_0(x,\mathbf{A})}$, we have

$$\eta_n^{-1}\mathbb{E}[\psi_{\mathsf{x}}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}}(t\eta_n;\mathbf{Z})]\to f_0(\mathsf{x})\mathbb{E}\left[\frac{\sigma_0^2(\mathsf{x},\mathbf{A})}{g_0(\mathsf{x},\mathbf{A})}\right]\Big((s\wedge t)\mathbb{1}\{s,t>0\}-(s\vee t)\mathbb{1}\{s,t<0\}\Big),$$

as desired.

(B5). $\mathcal{C}_{\mathsf{x}}(1,1)>0$ follows from $f_0(\mathsf{x})\mathbb{E}[\frac{\sigma_0^2(\mathsf{x},\mathbf{A})}{g_0(\mathsf{x},\mathbf{A})}]>0$. $\lim_{\eta\downarrow 0}\mathcal{C}_{\mathsf{x}}(1,\eta)/\sqrt{\eta}=0$ follows from the same computation as in the monotone density estimation. The remaining conditions follow from verifying (E5) above.

(B6) (B7) (B8). Verifying these conditions is the same as in the classical monotone regression case.

SA.2. Primitive sufficient conditions for Assumption MRC (iii). Here we provide primitive sufficient conditions for Assumption MRC (iii) by focusing on specific estimators $\widehat{\mu}_n$ and $\widehat{g}_n$. As discussed by Westling, Gilbert and Carone (2020), cross-fitting avoids restrictions on uniform entropy, allowing for a large class of flexible preliminary estimators. Here we use sample splitting to simplify exposition, but the proposed procedure can be straightforwardly modified for cross-fitting.

Suppose there is a separate random sample $\tilde{\mathbf{Z}}_1,\ldots,\tilde{\mathbf{Z}}_n$ drawn from the distribution of $\mathbf{Z}$, which is independent of $\mathbf{Z}_1,\ldots,\mathbf{Z}_n$. Preliminary estimators $\widehat{\mu}_n$ and $\widehat{g}_n$ are constructed from $\tilde{\mathbf{Z}}_1,\ldots,\tilde{\mathbf{Z}}_n$. For concreteness, we consider a partitioning-based least squares estimator $\widehat{\mu}_n$ (Cattaneo, Farrell and Feng, 2020) and local polynomial kernel-based estimators $\widehat{f}_{X|\mathbf{A},n}(x|\mathbf{a})$ and $\widehat{f}_n(x)$ of $f_{X|\mathbf{A}}(x|\mathbf{a})$ and $f_0(x)$ (Cattaneo, Jansson and Ma, 2020; Cattaneo et al., 2024), from which we construct $\widehat{g}_n(x,\mathbf{a})=\widehat{f}_{X|\mathbf{A},n}(x|\mathbf{a})/\widehat{f}_n(x)$.

Let $d=\dim(\mathbf{A})$. For simplicity, suppose the support of $(X,\mathbf{A}')'$ equals $[0,1]^{1+d}$. Let $\mathbf{p}(x,\mathbf{a})$ be a k_n-dimensional vector of bounded basis functions of order m on $\mathcal{S}$ which are locally supported e.g., splines (see Cattaneo, Farrell and Feng, 2020, for details and examples of basis functions). We consider the estimator

$$\widehat{\mu}_n(x,\mathbf{a})=\mathbf{p}(x,\mathbf{a})'\bigg(\sum_{i=1}^n\mathbf{p}(\tilde{X}_i,\tilde{\mathbf{A}}_i)\mathbf{p}(\tilde{X}_i,\tilde{\mathbf{A}}_i)'\bigg)^{-1}\sum_{i=1}^n\mathbf{p}(\tilde{X}_i,\tilde{\mathbf{A}}_i)\tilde{Y}_i.$$

For the estimator of $f_{X|\mathbf{A}}(x|\mathbf{a})$, letting $\widehat{F}_{X|\mathbf{A},n}(\cdot|\mathbf{a})$ be an estimator of $\mathbb{P}[X\leq\cdot|\mathbf{A}=\mathbf{a}]$ specified below, $\widehat{f}_{X|\mathbf{A},n}(x|\mathbf{a})$ is obtained by local polynomial regression:

$$\widehat{f}_{X|\mathbf{A},n}(x|\mathbf{a})=\mathbf{e}_2'\widehat{\boldsymbol{\beta}}(x|\mathbf{a}),\quad \widehat{\boldsymbol{\beta}}(x|\mathbf{a})=\operatorname*{argmin}_{\mathbf{u}\in\mathbb{R}^{\mathfrak{p}_1+1}}\sum_{i=1}^n\Big(\widehat{F}_{X|\mathbf{A},n}(\tilde{X}_i|\mathbf{a})-\mathbf{q}_1(\tilde{X}_i-x)'\mathbf{u}\Big)^2K_h(\tilde{X}_i-x)$$

where $\mathfrak{p}_1\geq 1$ is the order of the polynomial basis $\mathbf{q}_1(x)=(1,x/1!,x^2/2!,\ldots,x^{\mathfrak{p}_1}/\mathfrak{p}_1!)'$, $\mathbf{e}_l$ is the conformable unit vector whose lth element is unity, and $K_h(x)=K(x/h)/h$ for some kernel function K and some positive bandwidth h. The estimator $\widehat{F}_{X|\mathbf{A},n}(x|\mathbf{a})$ is constructed via local polynomial regression of order $\mathfrak{p}_2=\mathfrak{p}_1-1$:

$$\widehat{F}_{X|\mathbf{A},n}(x|\mathbf{a})=\mathbf{e}_1'\widehat{\boldsymbol{\gamma}}(x|\mathbf{a}),\quad \widehat{\boldsymbol{\gamma}}(x|\mathbf{a})=\operatorname*{argmin}_{\mathbf{v}\in\mathbb{R}^{k_{\mathfrak{p}_2}}}\sum_{i=1}^n\big(\mathbb{1}\{\tilde{X}_i\leq x\}-\mathbf{q}_2(\tilde{\mathbf{A}}_i-\mathbf{a})'\mathbf{v}\big)^2L_h(\tilde{\mathbf{A}}_i-\mathbf{a})$$

where, using standard multi-index notation, $\mathbf{q}_2(\mathbf{a})$ denotes the $k_{\mathfrak{p}_2}$-dimensional vector collecting the polynomials $\mathbf{a}^{\mathbf{m}}/\mathbf{m}!$ for $0\leq|\mathbf{m}|\leq\mathfrak{p}_2$ with $\mathbf{a}^{\mathbf{m}}=a_1^{m_1}a_2^{m_2}\ldots a_d^{m_d}$, $|\mathbf{m}|=\sum_{j=1}^d m_j$, and $k_{\mathfrak{p}_2}=\frac{(d+\mathfrak{p}_2)!}{d!\mathfrak{p}_2!}+1$, and $L_h(\mathbf{a})=L(\mathbf{a}/h)/h^d$ for $L(\mathbf{a})=\prod_{j=1}^d K(a_j)$ i.e., product kernel. The estimator $\widehat{f}_n(x)$ is constructed in a similar manner. First, the empirical cdf $\widehat{F}_n$ of $\{\tilde{X}_i\}$ is constructed and then $\widehat{f}_n(x)$ is formed via local polynomial regression:

$$\widehat{f}_n(x)=\mathbf{e}_2'\widehat{\boldsymbol{\delta}}_n(x),\quad \widehat{\boldsymbol{\delta}}_n(x)=\operatorname*{argmin}_{\mathbf{w}\in\mathbb{R}^{\mathfrak{p}_1+1}}\sum_{i=1}^n\Big(\widehat{F}_n(\tilde{X}_i)-\mathbf{q}_1(\tilde{X}_i-x)'\mathbf{w}\Big)^2K_b(\tilde{X}_i-x)$$

where $b>0$ is some bandwidth.

Now we state sufficient conditions for Assumption MRC (ii) based on the partitioning-based series estimator $\widehat{\mu}_n$ and the kernel-based estimator $\widehat{g}_n$.

Primitive Conditions MRC
(i) $\tilde{\mathbf{Z}}_1,\ldots,\tilde{\mathbf{Z}}_n$ are independent of $\mathbf{Z}_1,\ldots,\mathbf{Z}_n$.
(ii) The support of $(X,\mathbf{A}')'$ is $[0,1]^{1+d}$, and the distribution of $(X,\mathbf{A}')'$ is absolutely continuous. The Lebesgue density of $(X,\mathbf{A}')'$ and the conditional variance of Y given $(X,\mathbf{A}')'$ are bounded away from zero and continuous on $[0,1]^{1+d}$. μ_0 is $(m+1)$-times continuously differentiable on $[0,1]^{1+d}$.
(iii) The vector of basis functions p satisfies Assumptions 2, 3, and 4 of Cattaneo, Farrell and Feng (2020).
(iv) $f_{X|\mathbf{A}}(x|\mathbf{a})$ and $f_0(x)$ are $\mathfrak{p}_1$-times continuously differentiable in x, and $f_{X|\mathbf{A}}(x|\mathbf{a})$ is $\mathfrak{p}_1$-times continuously differentiable in $\mathbf{a}$.
(v) K is a symmetric, Lipschitz continuous probability density function supported on $[-1,1]$.

As verified by Cattaneo, Farrell and Feng (2020), (iii) holds for widely used local basis functions such as splines and wavelets.

LEMMA SA-7. *Suppose $\tilde{\mathbf{Z}}_1,\ldots,\tilde{\mathbf{Z}}_n$ is a random sample drawn from the distribution of $\mathbf{Z}$ and Primitive Conditions MRC hold. In addition, with $\tau_n=n^{-1}\log n$, $k_n=O(\tau_n^{-\frac{d+1}{2m+d+1}})$, $h=O(\tau_n^{\frac{d+1}{2\mathfrak{p}_1+d+1}})$, $b=O(\tau_n^{\frac{1}{2\mathfrak{p}_1+1}})$, $\frac{m}{2m+d+1}+\frac{\mathfrak{p}_1}{2\mathfrak{p}_1+d+1}\geq\frac{1}{2}$, and $\min\{\frac{2m}{2m+d+1},\frac{2\mathfrak{p}_1}{2\mathfrak{p}_1+d+1}\}>\frac{1}{2\mathfrak{q}+1}$. Then, $\widehat{\mu}_n$ and $\widehat{g}_n$ described above and $\widehat{\Gamma}_n$ based on the $\widehat{\mu}_n,\widehat{g}_n$ satisfy Assumption MRC (iii) with $\delta=\min\{\mathsf{x},1-\mathsf{x}\}/4$. In particular,*

$$\sqrt{n}\sup_{|v|\leq V}\left|\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+v)+\bar{\Gamma}_n(\mathsf{x})\right|\leq VO_{\mathbb{P}}(1)+o_{\mathbb{P}}(a_n^{-1})$$

uniformly over $V\in(0,2\delta]$.

PROOF. By Theorem 4.3 of Cattaneo, Farrell and Feng (2020),

$$\sup_{(x,\mathbf{a})\in[0,1]^{1+d}}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|=O_{\mathbb{P}}\left(\sqrt{\frac{k_n\log n}{n}}+k_n^{-\frac{m}{d+1}}\right)$$

and by Theorem 1 of Cattaneo et al. (2024),

$$\sup_{(x,\mathbf{a})\in[0,1]^{1+d}}|\widehat{f}_{X|\mathbf{A},n}(x|\mathbf{a})-f_{X|\mathbf{A}}(x|\mathbf{a})|=O_{\mathbb{P}}\left(\sqrt{\frac{\log n}{nh^{1+d}}}+h^{\mathfrak{p}_1}\right).$$

Also, one can show

$$\sup_{x\in[0,1]}|\widehat{f}_n(x)-f_0(x)|=O_{\mathbb{P}}\left(\sqrt{\frac{\log n}{nb}}+b^{\mathfrak{p}_1}\right).$$

Then, with the specified rate of k_n,h,b,

$$\sup_{(x,\mathbf{a})\in[0,1]^{1+d}}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|=O_{\mathbb{P}}\left(\tau_n^{\frac{m}{2m+d+1}}\right),$$

$$\sup_{(x,\mathbf{a})\in[0,1]^{1+d}}|\widehat{g}_n(x,\mathbf{a})-g_0(x,\mathbf{a})|=O_{\mathbb{P}}\left(\tau_n^{\frac{\mathfrak{p}_1}{2\mathfrak{p}_1+d+1}}\right).$$

Since $\min\{\frac{2m}{2m+d+1},\frac{2\mathfrak{p}_1}{2\mathfrak{p}_1+d+1}\}>\frac{1}{2\mathfrak{q}+1}$, it follows $a_n\frac{1}{n}\sum_{i=1}^n|\widehat{\mu}_n(X_i,\mathbf{A}_i)-\mu_0(X_i,\mathbf{A}_i)|^2=o_{\mathbb{P}}(1)$, $a_n\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n|\widehat{\mu}_n(X_i,\mathbf{A}_j)-\mu_0(X_i,\mathbf{A}_j)|^2=o_{\mathbb{P}}(1)$, and $a_n\frac{1}{n}\sum_{i=1}^n\varepsilon_i^2|\widehat{g}_n(X_i,\mathbf{A}_i)-g_0(X_i,\mathbf{A}_i)|^2=o_{\mathbb{P}}(1)$. Also, by $\frac{m}{2m+d+1}+\frac{\mathfrak{p}_1}{2\mathfrak{p}_1+d+1}\geq\frac{1}{2}$,

(SA.21)
$$\sup_{(x,\mathbf{a})\in[0,1]^{1+d}}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|\sup_{(x,\mathbf{a})\in[0,1]^{1+d}}|\widehat{g}_n(x,\mathbf{a})-g_0(x,\mathbf{a})|=O_{\mathbb{P}}\Big(n^{-1/2}\Big).$$

Decompose $\widehat{\gamma}_n$ into $\widehat{\gamma}_{1,n}$ and $\widehat{\gamma}_{2,n}$ where

$$\widehat{\gamma}_{1,n}(x;\mathbf{Z})=\mathbb{1}\{X\leq x\}\frac{Y-\widehat{\mu}_n(X,\mathbf{A})}{\widehat{g}_n(X,\mathbf{A})},\quad \widehat{\gamma}_{2,n}(x;\mathbf{Z})=\mathbb{1}\{X\leq x\}\frac{1}{n}\sum_{j=1}^n\widehat{\mu}_n(X,\mathbf{A}_j)$$

and let $\widehat{\Gamma}_{k,n}(x)=\frac{1}{n}\sum_{i=1}^n\widehat{\gamma}_{1,n}(x;\mathbf{Z}_i)$ for $k=1,2$. Define $\gamma_{k,0},\bar{\Gamma}_{k,n}$ $k=1,2$ in the same manner.

Letting $\tilde{\mathfrak{Z}}_n$ be the σ-field generated by $\tilde{\mathbf{Z}}_1,\ldots,\tilde{\mathbf{Z}}_n$,

$$\begin{aligned}
&\mathbb{V}\Big[\widehat{\Gamma}_{1,n}(\mathsf{x}+v)-\widehat{\Gamma}_{1,n}(\mathsf{x})-\bar{\Gamma}_{1,n}(\mathsf{x}+v)+\bar{\Gamma}_{1,n}(\mathsf{x})\Big|\tilde{\mathfrak{Z}}_n\Big]\\
&\leq n^{-1}\mathbb{E}\left[\big(\mathbb{1}\{X\leq\mathsf{x}+v\}-\mathbb{1}\{X\leq\mathsf{x}\}\big)^2\left(\frac{Y-\widehat{\mu}_n(X,\mathbf{A})}{\widehat{g}_n(X,\mathbf{A})}-\frac{Y-\mu_0(X,\mathbf{A})}{g_0(X,\mathbf{A})}\right)^2\middle|\tilde{\mathfrak{Z}}_n\right]\\
&\leq n^{-1}C\mathbb{E}\big[\big|\mathbb{1}\{X\leq\mathsf{x}+v\}-\mathbb{1}\{X\leq\mathsf{x}\}\big|\varepsilon^2\big]\sup_{|x-\mathsf{x}|\leq|v|}\sup_{\mathbf{a}\in[0,1]^d}|\widehat{g}_n(x,\mathbf{a})-g_0(x,\mathbf{a})|^2\\
&\quad+n^{-1}C\mathbb{E}\big[\big|\mathbb{1}\{X\leq\mathsf{x}+v\}-\mathbb{1}\{X\leq\mathsf{x}\}\big|\big]\sup_{|x-\mathsf{x}|\leq|v|}\sup_{\mathbf{a}\in[0,1]^d}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|^2\\
&=o_{\mathbb{P}}\big((na_n)^{-1}\big)
\end{aligned}$$

where the inequalities hold with probability one.

Note $\widehat{\Gamma}_{2,n}(x)=\frac{1}{n^2}\sum_{1\leq i\neq j\leq n}\mathbb{1}\{X_i\leq x\}\widehat{\mu}_n(X_i,\mathbf{A}_j)+O_{\mathbb{P}}(n^{-1})$, and

$$\begin{aligned}
&\mathbb{V}\left[\frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n}\big(\mathbb{1}\{X_i\leq\mathsf{x}+v\}-\mathbb{1}\{X_i\leq\mathsf{x}\}\big)\big(\widehat{\mu}_n(X_i,\mathbf{A}_j)-\mu_0(X_i,\mathbf{A}_j)\big)\middle|\tilde{\mathfrak{Z}}_n\right]\\
&\leq n^{-1}C|v|\sup_{|x-\mathsf{x}|\leq|v|}\sup_{\mathbf{a}\in[0,1]^d}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|^2\\
&\quad+Cn^{-2}|v|\sup_{|x-\mathsf{x}|\leq|v|}\sup_{\mathbf{a}\in[0,1]^d}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|^2=o_{\mathbb{P}}\big((na_n)^{-1}\big)
\end{aligned}$$

where we use Hoeffding decomposition and the inequality holds with probability one.

To complete the proof, it suffices to show that there is a sequence of random variables $A_n'=O_{\mathbb{P}}(1)$ such that for $V\in(0,a_n\delta]$,

$$\sqrt{n}\sup_{|v|\leq V}\Big|\mathbb{E}\big[\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+v)+\bar{\Gamma}_n(\mathsf{x})\big|\tilde{\mathfrak{Z}}_n\big]\Big|\leq A_n'V.$$

With $f_{\mathbf{A}}(\mathbf{a})$ denoting the Lebesgue density of $\mathbf{A}$,

$$\begin{aligned}
&\mathbb{E}\big[\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})\big|\tilde{\mathfrak{Z}}_n\big]\\
&=\int\int\big(\mathbb{1}\{u\leq\mathsf{x}+v\}-\mathbb{1}\{u\leq\mathsf{x}\}\big)\frac{g_0(u,\mathbf{a})}{\widehat{g}_n(u,\mathbf{a})}\big[\mu_0(u,\mathbf{a})-\widehat{\mu}_n(u,\mathbf{a})\big]f_0(u)duf_{\mathbf{A}}(\mathbf{a})d\mathbf{a}
\end{aligned}$$

$$+\int\left(\mathbb{1}\{u\le \mathsf{x}+v\}-\mathbb{1}\{u\le \mathsf{x}\}\right)\int\widehat{\mu}_n(u,\mathbf{a})f_{\mathbf{A}}(\mathbf{a})d\mathbf{a}du$$

$$\mathbb{E}\big[\bar{\Gamma}_n(\mathsf{x}+v)-\bar{\Gamma}_n(\mathsf{x})\big|\tilde{\mathfrak{Z}}_n\big]$$
$$=\int\left(\mathbb{1}\{u\le \mathsf{x}+v\}-\mathbb{1}\{u\le \mathsf{x}\}\right)\int\mu_0(u,\mathbf{a})f_{\mathbf{A}}(\mathbf{a})d\mathbf{a}f_0(u)du.$$

Then, for $v\ge 0$,

$$\mathbb{E}\big[\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+v)+\bar{\Gamma}_n(\mathsf{x})\big|\tilde{\mathfrak{Z}}_n\big]$$
$$=\int_{\mathsf{x}\le u\le \mathsf{x}+v}\big[\widehat{\mu}_n(u,\mathbf{a})-\mu_0(u,\mathbf{a})\big]\Big[1-\frac{g_0(u,\mathbf{a})}{\widehat{g}_n(u,\mathbf{a})}\Big]f_0(u)f_{\mathbf{A}}(\mathbf{a})d(u,\mathbf{a}).$$

A similar expression holds for $v<0$. Then, for some fixed $C>0$,

$$\Big|\mathbb{E}\big[\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+v)+\bar{\Gamma}_n(\mathsf{x})\big|\tilde{\mathfrak{Z}}_n\big]\Big|$$
$$\le C|v|\sup_{|x-\mathsf{x}|\le|v|}\sup_{\mathbf{a}\in[0,1]^d}|\widehat{\mu}_n(x,\mathbf{a})-\mu_0(x,\mathbf{a})|\sup_{|x-\mathsf{x}|\le|v|}\sup_{\mathbf{a}\in[0,1]^d}|\widehat{g}_n(x,\mathbf{a})-g_0(x,\mathbf{a})|$$

and the desired result follows from (SA.21). □

SA.3. Additional example: monotone density function with conditionally independent right-censoring. We introduce covariates $\mathbf{A}$ and consider the case of censoring at random: $X\perp\!\!\!\perp C|\mathbf{A}$. See van der Laan and Robins (2003); Zeng (2004) and references therein for existing analysis of this problem. We have

$$\gamma_0(x;\mathbf{Z})=F_0(x|\mathbf{A})+S_0(x|\mathbf{A})\left[\frac{\Delta\mathbb{1}\{\check{X}\le x\}}{S_0(\check{X}|\mathbf{A})G_0(\check{X}|\mathbf{A})}-\int_0^{\check{X}\wedge x}\frac{\Lambda_0(du|\mathbf{A})}{S_0(u|\mathbf{A})G_0(u|\mathbf{A})}\right]$$

where $F_0(x|A)=1-S_0(x|A)$, $S_0(x|\mathbf{A})=\mathbb{P}[X>x|\mathbf{A}]$, $G_0(c|\mathbf{A})=\mathbb{P}[C>c|\mathbf{A}]$, and $\Lambda_0(x|\mathbf{A})=\int_0^x\frac{f_0(u|\mathbf{A})}{S_0(u|\mathbf{A})}du$ with f_0 being the Lebesgue density of X. Denote by $\widehat{S}_n(\cdot|\cdot)$, $\widehat{G}_n(\cdot|\cdot)$, $\widehat{\Lambda}_n(\cdot|\cdot)$ preliminary estimates of S_0,G_0,Λ_0, respectively.

Assumption SA.3. Let $\mathfrak{S}_n$, $\mathfrak{G}_n,\mathfrak{L}_n$ be sequences of function classes that contain $S_0(\cdot|\cdot)$, $G_0(\cdot|\cdot)$, $\Lambda_0(\cdot|\cdot)$, respectively.

(i) x is in the interior of $I=[0,u_0]$, $X\perp\!\!\!\perp C|A$, and $\theta_0=f_0$ satisfies Assumption (A2).
(ii) There exist $c,c_1,c_2>0,\rho_\gamma\in(0,2)$such that for $n\ge 1$, for any $S\in\mathfrak{S}_n$, $G\in\mathfrak{G}_n$, and $\Lambda\in\mathfrak{L}_n$, the following hold: $\log N_U(\varepsilon,\{S(x|\cdot):x\in I\})\le c\varepsilon^{-\rho_\gamma}$ for $\varepsilon\in(0,1)$, where N_U is as defined in Section 4.2 of the main paper, and $c_1\le S(x|\mathbf{A})\le c_2$, $c_1\le G(x|\mathbf{A})\le c_2$ for $x\in I$, and $\Lambda(u_0|\mathbf{A})\le c_2$ with probability one.
(iii) There exist $\delta>0,\beta_\gamma\in[1/2,2)$ such that for $V\in(0,2\delta]$, $\sqrt{na_n}\sup_{|v|\le V}|\widehat{\Gamma}_n(\mathsf{x}+v)-\widehat{\Gamma}_n(\mathsf{x})-\bar{\Gamma}_n(\mathsf{x}+v)+\bar{\Gamma}_n(\mathsf{x})|\le o_{\mathbb{P}}(1)+V^{\beta_\gamma}o_{\mathbb{P}}(a_n^{\beta_\gamma})$ where $o_{\mathbb{P}}$ terms do not depend on V.
(iv) With probability approaching one, $\widehat{S}_n\in\mathfrak{S}_n$, $\widehat{G}_n\in\mathfrak{G}_n$, $\widehat{\Lambda}_n\in\mathfrak{L}_n$. For $(\widehat{h}_n,h_0)\in\{(\widehat{S}_n,S_0),(\widehat{G}_n,G_0),(\widehat{\Lambda}_n,\Lambda_0)\}$,

$$a_n\frac{1}{n}\sum_{i=1}^n\sup_{x\in I}|\widehat{h}_n(x|\mathbf{A}_i)-h_0(x|\mathbf{A}_i)|^2=o_{\mathbb{P}}(1).$$

(v) The conditional distribution of X given $\mathbf{A}$ has bounded Lebesgue density $f_{X|\mathbf{A}}$, $\mathbb{E}[\frac{f_{X|\mathbf{A}}(\mathsf{x}|\mathbf{A})}{G_0(\mathsf{x}|\mathbf{A})}]>0$, and there are real-valued functions B,ω such that $\mathbb{E}[B(\mathbf{A})]<\infty$,

$\lim_{\eta\downarrow 0}\omega(\eta)=0$, and for $|x-\mathsf{x}|$ sufficiently small, $|\frac{f_{X|\mathbf{A}}(x|\mathbf{A})}{S_0(x|\mathbf{A})G_0(x|\mathbf{A})}-\frac{f_{X|\mathbf{A}}(\mathsf{x}|\mathbf{A})}{S_0(\mathsf{x}|\mathbf{A})G_0(\mathsf{x}|\mathbf{A})}|\leq \omega(|x-\mathsf{x}|)B(\mathbf{A})$.

The condition (iii) is high-level, and there are a few different approaches to verify them. See Westling and Carone (2020) for details.

COROLLARY SA-1. *Under Assumption SA.3, Assumptions A and B hold with*

$$\widehat{\Gamma}_n(x)=\frac{1}{n}\sum_{i=1}^n\widehat{\gamma}_n(x;\mathbf{Z}_i),\quad \widehat{\Gamma}_n^*(x)=\frac{1}{n}\sum_{i=1}^n W_{i,n}\widehat{\gamma}_n(x;\mathbf{Z}_i)$$

$$\widehat{\gamma}_n(x;\mathbf{Z})=\widehat{F}_n(x|\mathbf{A})+\widehat{S}_n(x|\mathbf{A})\Big[\frac{\Delta\mathbb{1}(\check{X}\leq x)}{\widehat{S}_n(\check{X}|\mathbf{A})\widehat{G}_n(\check{X}|\mathbf{A})}-\int_0^{\check{X}\wedge x}\frac{\widehat{\Lambda}_n(du|\mathbf{A})}{\widehat{S}_n(u|\mathbf{A})\widehat{G}_n(u|\mathbf{A})}\Big],$$

where $\widehat{F}_n=1-\widehat{S}_n$*,*

$$\widehat{\Phi}_n(x)=\widehat{\Phi}_n^*(x)=x,\quad \widehat{u}_n=\widehat{u}_n^*=u_0,$$

$$\mathcal{C}_{\mathsf{x}}(s,t)=\mathbb{E}\Big[\frac{f_{X|\mathbf{A}}(\mathsf{x}|\mathbf{A})}{G_0(\mathsf{x}|\mathbf{A})}\Big](|s|\wedge|t|)\mathbb{1}\{\mathrm{sign}(s)=\mathrm{sign}(t)\},\quad \mathcal{D}_{\mathfrak{q}}(\mathsf{x})=\frac{\partial^{\mathfrak{q}}f_0(\mathsf{x})}{(\mathfrak{q}+1)!}.$$

SA.3.1. *Proof of Corollary SA-1.* In this example, $\widehat{\Phi}_n(x)=\widehat{\Phi}_n^*(x)=x=\Phi_0(x)$. Assumptions A and (E1)-(E2) follow from the hypothesis.

(E3). In this example, $\check{\Gamma}_n=\widehat{\Gamma}_n$. For $x\in I$,

$$\begin{aligned}
\widehat{\gamma}_n(x;\mathbf{Z})-\gamma_0(x;\mathbf{Z})&=\widehat{F}_n(x|\mathbf{A})-F_0(x|\mathbf{A})\\
&+\big[\widehat{S}_n(x|\mathbf{A})-S_0(x|\mathbf{A})\big]\Big[\frac{\Delta\mathbb{1}(\check{X}\leq x)}{\widehat{S}_n(\check{X}|\mathbf{A})\widehat{G}_n(\check{X}|\mathbf{A})}-\int_0^{\check{X}\wedge x}\frac{\widehat{\Lambda}_n(du|\mathbf{A})}{\widehat{S}_n(u|\mathbf{A})\widehat{G}_n(u|\mathbf{A})}\Big]\\
&+S_0(x|\mathbf{A})\Delta\mathbb{1}\{\check{X}\leq x\}\Big[\big(\widehat{S}_n(\check{X}|\mathbf{A})\widehat{G}_n(\check{X}|\mathbf{A})\big)^{-1}-\big(S_0(\check{X}|\mathbf{A})G_0(\check{X}|\mathbf{A})\big)^{-1}\Big]\\
&-S_0(x|\mathbf{A})\Big[\int_0^{\check{X}\wedge x}\frac{\widehat{\Lambda}_n(du|\mathbf{A})}{\widehat{S}_n(u|\mathbf{A})\widehat{G}_n(u|\mathbf{A})}-\int_0^{\check{X}\wedge x}\frac{\Lambda_0(du|\mathbf{A})}{S_0(u|\mathbf{A})G_0(u|\mathbf{A})}\Big]
\end{aligned}$$

and

$$\begin{aligned}
&\left|\int_0^{\check{X}\wedge x}\frac{\widehat{\Lambda}_n(du|\mathbf{A})}{\widehat{S}_n(u|\mathbf{A})\widehat{G}_n(u|\mathbf{A})}-\int_0^{\check{X}\wedge x}\frac{\Lambda_0(du|\mathbf{A})}{S_0(u|\mathbf{A})G_0(u|\mathbf{A})}\right|\\
&\leq\sup_{x'\in I}\Big|\big(\widehat{S}_n(x'|\mathbf{A})\widehat{G}_n(x'|\mathbf{A})\big)^{-1}-\big(S_0(x'|\mathbf{A})G_0(x'|\mathbf{A})\big)^{-1}\Big|\widehat{\Lambda}_n(u_0|\mathbf{A})\\
&\quad+\left|\frac{\widehat{\Lambda}_n(\check{X}\wedge x|\mathbf{A})-\Lambda_0(\check{X}\wedge x|\mathbf{A})}{S_0(\check{X}\wedge x|\mathbf{A})G_0(\check{X}\wedge x|\mathbf{A})}\right|+\left|\frac{\widehat{\Lambda}_n(0|\mathbf{A})-\Lambda_0(0|\mathbf{A})}{S_0(0|\mathbf{A})G_0(0|\mathbf{A})}\right|\\
&\quad+\left|\int_0^{\check{X}\wedge x}[\widehat{\Lambda}_n(u|\mathbf{A})-\Lambda_0(u|\mathbf{A})]J_0(du|\mathbf{A})\right|
\end{aligned}$$

using integration by parts, where $J_0(u|\mathbf{a}) = [S_0(u|\mathbf{a})G_0(u|\mathbf{a})]^{-1}$. Thus, there is a fixed $C > 0$ such that

$$|\widehat{\gamma}_n(x;\mathbf{Z}) - \gamma_0(x;\mathbf{Z})| \le C\Big[\sup_{x\in I}|\widehat{S}_n(x|\mathbf{A}) - S_0(x|\mathbf{A})| + \sup_{x\in I}|\widehat{G}_n(x|\mathbf{A}) - G_0(x|\mathbf{A})| + \sup_{x\in I}|\widehat{\Lambda}_n(x|\mathbf{A}) - \Lambda_0(x|\mathbf{A})|\Big],$$

From the hypothesis,

$$\frac{1}{n}\sum_{i=1}^{n}\sup_{x\in I}|\widehat{\gamma}_n(x;\mathbf{Z}_i) - \gamma_0(x;\mathbf{Z}_i)|^2 = o_{\mathbb{P}}(1),\ a_n\frac{1}{n}\sum_{i=1}^{n}\sup_{x\in I^\delta_{\mathsf{x}}}\left|\widehat{\gamma}_n(x;\mathbf{Z}_i) - \gamma_0(x;\mathbf{Z}_i)\right|^2 = o_{\mathbb{P}}(1)$$

follow.

For uniform covering numbers, it suffices to show that each of $\{S(x|\cdot) : x \in I\}$, $\{\mathbb{1}\{\cdot \le x\} : x \in I\}$, and $\{\int_0^{\cdot\wedge x}\frac{\Lambda(du|\cdot)}{S(u|\cdot)G(u|\cdot)} : x \in I\}$ has an appropriate bound on the uniform covering number by Lemma 5.1 of van der Vaart and van der Laan (2006) (see examples after the lemma). For $\{\int_0^{\cdot\wedge x}\frac{\Lambda(du|\cdot)}{S(u|\cdot)G(u|\cdot)} : x \in I\}$ with $(S, G, \Lambda) \in \mathfrak{S}_n \times \mathfrak{G}_n \times \mathfrak{L}_n$, the mapping $x \mapsto \int_0^{x}\frac{\Lambda(du|\cdot)}{S(u|\cdot)G(u|\cdot)}$ is monotone (by the non-decreasing property of Λ and $S, G \ge c_1 > 0$) and Lemma 2.6.18 of van der Vaart and Wellner (1996) implies the desired result.

There is a fixed $C > 0$ such that for $x \in I^\eta_{\mathsf{x}}$,

$$|\gamma_0(x;\mathbf{Z}) - \gamma_0(\mathsf{x};\mathbf{Z})| \le C\Big[|S_0(x|\mathbf{A}) - S_0(\mathsf{x}|\mathbf{A})| + C\Delta|\mathbb{1}\{\check{X} \le x\} - \mathbb{1}\{\check{X} \le \mathsf{x}\}|\Big]$$

and using $1 - S_0(x|\cdot) = \int_0^x f_{X|\mathbf{A}}(u|\cdot)du$ with $f_{X|\mathbf{A}}$ being bounded, we can take

$$\bar{D}^\eta_\gamma(\mathbf{Z}) = C\Delta\mathbb{1}\{\mathsf{x} - \eta \le \check{X} \le \mathsf{x} + \eta\} + C\eta,$$

which satisfies the desired bound condition.

(E5). We have

$$\psi_{\mathsf{x}}(v;\mathbf{Z}) = S_0(\mathsf{x}|\mathbf{A})\frac{(\mathbb{1}\{\check{X} \le \mathsf{x} + v\} - \mathbb{1}\{\check{X} \le \mathsf{x}\})\Delta}{S_0(\check{X}|\mathbf{A})G_0(\check{X}|\mathbf{A})} + O(|v|)$$

and the first display follows as in the independent censoring case. For the covariance kernel,

$$\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})] = \mathbb{E}\left[S_0(\mathsf{x}_n|\mathbf{A})^2\int_{\mathsf{x}_n}^{\mathsf{x}_n+\eta_n(s\wedge t)}\frac{f_{X|\mathbf{A}}(u|\mathbf{A})}{S_0(u|\mathbf{A})^2G_0(u|\mathbf{A})}du\mathbb{1}\{s,t>0\}\right] + \mathbb{E}\left[S_0(\mathsf{x}_n|\mathbf{A})^2\int_{\mathsf{x}_n+\eta_n(s\vee t)}^{\mathsf{x}_n}\frac{f_{X|\mathbf{A}}(u|\mathbf{A})}{S_0(u|\mathbf{A})^2G_0(u|\mathbf{A})}du\mathbb{1}\{s,t<0\}\right] + O(\eta_n^2)$$

and $\eta_n^{-1}\mathbb{E}[\psi_{\mathsf{x}_n}(s\eta_n;\mathbf{Z})\psi_{\mathsf{x}_n}(t\eta_n;\mathbf{Z})]$ converges to $\mathbb{E}[\frac{f_{X|\mathbf{A}}(\mathsf{x}|\mathbf{A})}{G_0(\mathsf{x}|\mathbf{A})}](|s|\wedge|t|)\mathbb{1}\{\mathrm{sign}(s) = \mathrm{sign}(t)\}$.

(B5). $\mathcal{C}_{\mathsf{x}}(1,1) > 0$ follows from $\mathbb{E}[f_{X|\mathbf{A}}(\mathsf{x}|A)/G_0(\mathsf{x}|A)] > 0$. $\lim_{\eta\downarrow 0}\mathcal{C}_{\mathsf{x}}(1,\eta)/\sqrt{\eta} = 0$ follows from the same computation as in the no censoring case. The remaining conditions follow from verifying (E5).

(B6), (B7), and (B8) hold since in this example, $\widehat{u}_n = \widehat{u}^*_n = u_0$ and $\widehat{\Phi}_n, \widehat{\Phi}^*_n$ are the identity map.

SA.4. Additional example: monotone hazard function. Let X be a non-negative random variable, f_0 be its Lebesgue density, and $S_0(x) = \mathbb{P}[X > x]$ be its survival function. We consider the parameter of estimating the hazard function of X, $\theta_0(\mathsf{x}) = f_0(\mathsf{x})/S_0(\mathsf{x})$, with possible right-censoring as in the monotone density function example. Observations $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ come from a random sample of $\mathbf{Z} = (\check{X}, \Delta)'$ where $\check{X} = \min\{X, C\}$ and $\Delta = \mathbb{1}\{X \le C\}$, C being a random censoring time. As pointed out by Westling and Carone (2020), with strictly increasing Φ_0, the function Γ_0 takes the form $\Gamma_0(x) = \int_0^x \frac{f_0(u)}{S_0(u)} \Phi_0(du)$, and by taking $\Phi_0(x) = \int_0^x S_0(u)du$, $\Gamma_0(x) = F_0(x) = \mathbb{P}[X \le x]$. Since Γ_0 is identical to the monotone density case with the choice $\Phi_0 = \int_0^x S_0(u)du$, we can leverage the analysis for the monotone density. The interval I equals $[0, u_0^{\mathtt{MD}}]$ where $u_0^{\mathtt{MD}}$ is u_0 in the monotone density example. The u_0 for the monotone hazard function estimation is $u_0 = \Phi_0(u_0^{\mathtt{MD}})$.

Consider the case of completely random censoring i.e., $X \perp\!\!\!\perp C$. As in the setup for Corollary 2, let $\widehat{S}_n(x)$ be the Kaplan-Meier estimator for $S_0(x) = 1 - F_0(x) = \mathbb{P}[X > x]$, $\widehat{F}_n = 1 - \widehat{S}_n$, and $\widehat{G}_n$ be the Kaplan-Meier estimator for $G_0(x) = \mathbb{P}[C > x]$. Also,

$$\gamma_0(x; \mathbf{Z}) = F_0(x) + S_0(x) \left[\frac{\Delta \mathbb{1}\{\check{X} \le x\}}{S_0(\check{X}) G_0(\check{X})} - \int_0^{\check{X} \wedge x} \frac{\Lambda_0(du)}{S_0(u) G_0(u)} \right]$$

and $\phi_0(x; \mathbf{Z}) = x - \int_0^x \gamma_0(u; \mathbf{Z}) du$.

COROLLARY SA-2. *Suppose that the hypothesis of Corollary 2 and Assumption BW hold. Then, Assumptions A and B hold with*

$$\widehat{\Gamma}_n(x) = 1 - \widehat{S}_n(x), \quad \widehat{\Gamma}_n^*(x) = \frac{1}{n} \sum_{i=1}^n W_{i,n} \widehat{\gamma}_n(x; \mathbf{Z}_i),$$

$$\widehat{\gamma}_n(x; \mathbf{Z}) = \widehat{F}_n(x) + \widehat{S}_n(x) \left[\frac{\Delta \mathbb{1}\{\check{X} \le x\}}{\widehat{S}_n(\check{X}) \widehat{G}_n(\check{X})} - \int_0^{\check{X} \wedge x} \frac{\widehat{\Lambda}_n(du)}{\widehat{S}_n(u) \widehat{G}_n(u)} \right],$$

$$\widehat{\Phi}_n(x) = \int_0^x \widehat{F}_n(u) du, \quad \widehat{\Phi}_n^*(x) = \int_0^x [1 - \widehat{\Gamma}_n^*(u)] du = \frac{1}{n} \sum_{i=1}^n W_{i,n} \widehat{\phi}_n(x; \mathbf{Z}_i),$$

$$\widehat{\phi}_n(x; \mathbf{Z}) = x - \int_0^x \widehat{\gamma}_n(u; \mathbf{Z}) du, \quad \widehat{u}_n = \widehat{u}_n^* = \widehat{\Phi}_n(u_0^{\mathtt{MD}}),$$

$$\mathcal{C}_{\mathsf{x}}(s,t) = \frac{f_0(\mathsf{x})}{G_0(\mathsf{x})} (|s| \wedge |t|) \mathbb{1}\{\operatorname{sign}(s) = \operatorname{sign}(t)\}, \quad \mathcal{D}_{\mathfrak{q}}(\mathsf{x}) = \frac{S_0(\mathsf{x}) \partial^{\mathfrak{q}} f_0(\mathsf{x})}{(\mathfrak{q}+1)!}.$$

SA.4.1. *Proof of Corollary SA-2.* We use the same $\widehat{\gamma}_n$ function and assumptions as in the monotone density setting. Also, the covariance kernels are the same as in the monotone density case. Thus, (E3) and part of (B5) follow from the same argument. We focus on (E4)-(E5) and (B6)-(B8).

(E4). Since

$$\widehat{\phi}_n(x; \mathbf{Z}) - \phi_0(x; \mathbf{Z}) = -\int_0^x [\widehat{\gamma}_n(u; \mathbf{Z}) - \gamma_0(u; \mathbf{Z})] du,$$

$$\frac{1}{n} \sum_{i=1}^n \sup_{x \in I} |\widehat{\phi}_n(x; \mathbf{Z}_i) - \phi_0(x; \mathbf{Z}_i)|^2 = o_{\mathbb{P}}(1), \; a_n \frac{1}{n} \sum_{i=1}^n \sup_{x \in I_{\mathsf{x}}^{\delta}} |\widehat{\phi}_n(x; \mathbf{Z}_i) - \phi_0(x; \mathbf{Z}_i)|^2 = o_{\mathbb{P}}(1)$$

follow from $a_n \frac{1}{n}\sum_{i=1}^n \sup_{x\in I} |\widehat{\gamma}_n(x;\mathbf{Z}_i) - \gamma_0(x;\mathbf{Z}_i)|^2$, which was verified in Section SA.1.12. To check $\sup_{x\in I}|\widehat{\Phi}_n(x) - \Phi_0(x)| = o_{\mathbb{P}}(1)$, $\sup_{x\in I}|\frac{1}{n}\sum_{i=1}^n \phi_0(x;\mathbf{Z}_i) - \Phi_0(x)| = o_{\mathbb{P}}(1)$ follows from Glivenko-Cantelli, and

$$\sup_{x\in I}\left|\frac{1}{n}\sum_{i=1}^n \left[\widehat{\phi}_n(x;\mathbf{Z}_i) - \phi_0(x;\mathbf{Z}_i)\right]\right| \leq \frac{1}{n}\sum_{i=1}^n \sup_{x\in I}|\widehat{\gamma}_n(x;\mathbf{Z}_i) - \gamma_0(x;\mathbf{Z}_i)| = o_{\mathbb{P}}(1),$$

where the last equality follows from $\frac{1}{n}\sum_{i=1}^n \sup_{x\in I}|\widehat{\gamma}_n(x;\mathbf{Z}_i) - \gamma_0(x;\mathbf{Z}_i)|^2 = o_{\mathbb{P}}(1)$. Now $\sup_{x\in I}|\widehat{\Phi}_n(x) - \Phi_0(x)| = o_{\mathbb{P}}(1)$ follows by the triangle inequality.

For $|v| \leq V$

$$\begin{aligned}\left|\widehat{\Phi}_n(\mathsf{x}+v) - \widehat{\Phi}_n(\mathsf{x}) - \bar{\Phi}_n(\mathsf{x}+v) + \bar{\Phi}_n(\mathsf{x})\right| &= \left|\int_{\mathsf{x}-v}^{\mathsf{x}+v} [\widehat{\Gamma}_n(u) - \bar{\Gamma}_n(u)]du\right| \\ &\leq 2|V|\Big(\sup_{|v|\leq|V|}|\widehat{\Gamma}_n(\mathsf{x}+v) - \widehat{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+v) + \bar{\Gamma}_n(\mathsf{x})| + |\widehat{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x})|\Big)\end{aligned}$$

Using the argument in Section SA.1.12, we can bound $\sup_{|v|\leq|V|}|\widehat{\Gamma}_n(\mathsf{x}+v) - \widehat{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x}+v) + \bar{\Gamma}_n(\mathsf{x})|$. Then, $\sqrt{n}|\widehat{\Gamma}_n(\mathsf{x}) - \bar{\Gamma}_n(\mathsf{x})| = O_{\mathbb{P}}(1)$ implies

$$\sqrt{na_n}\sup_{|v|\leq V}\left|\widehat{\Phi}_n(\mathsf{x}+v) - \widehat{\Phi}_n(\mathsf{x}) - \bar{\Phi}_n(\mathsf{x}+v) + \bar{\Phi}_n(\mathsf{x})\right| \leq o_{\mathbb{P}}(1) + VO_{\mathbb{P}}\big(\sqrt{a_n}\big)$$

uniformly over $V \in (0, 2\delta]$. Theorem 1 of Lo and Singh (1986) implies $\sqrt{na_n}\sup_{x\in I}|\check{\Phi}_n(x) - \check{\Phi}_n(\mathsf{x}) - \bar{\Phi}_n(x) + \bar{\Phi}_n(\mathsf{x})| = o_{\mathbb{P}}(1)$.

The conditions on the uniform covering number hold because γ_0 and $\widehat{\gamma}_n$ are bounded (for $\widehat{\gamma}_n$, with probability approaching one) and thus $|\phi_0(x_1;\mathbf{Z}) - \phi_0(x_2;\mathbf{Z})| \leq C|x_1 - x_2|$ and $|\widehat{\phi}_n(x_1;\mathbf{Z}) - \widehat{\phi}_n(x_2;\mathbf{Z})| \leq C|x_1 - x_2|$ with probability approaching one. By this Lipschitz property, the condition on $\bar{D}_\phi^\eta(\mathbf{Z})$ also holds.

(E5). Let $\psi_{\mathsf{x}}^{\mathtt{MD}}(v;\mathbf{Z}) = \gamma_0(\mathsf{x}+v;\mathbf{Z}) - \gamma_0(\mathsf{x};\mathbf{Z}) - \theta_0(\mathsf{x})v$ be the ψ_{x} function for the monotone density. Then, for x sufficiently close to x and $|v|$ small enough,

$$\begin{aligned}\psi_x(v;\mathbf{z}) &= \gamma_0(x+v;\mathbf{z}) - \gamma_0(x;\mathbf{z}) - \theta_0(x)[\phi_0(x+v;\mathbf{z}) - \phi_0(x;\mathbf{z})] \\ &= \psi_x^{\mathtt{MD}}(v;\mathbf{Z}) + \theta_0(x)\int_x^{x+v}\gamma_0(u;\mathbf{Z})du = \psi_x^{\mathtt{MD}}(v;\mathbf{Z}) + O(|v|).\end{aligned}$$

Then, the same argument as in the monotone density case implies the desired result.

(B6) follows from consistency of $\widehat{\Phi}_n$ and $\widehat{\Phi}_n^*$.

(B7). $\widehat{\Phi}_n(x) = \int_0^x \widehat{F}_n(u)du, \widehat{\Phi}_n^*(x) = \int_0^x 1 - \widehat{\Gamma}_n^*(u)du$ are non-negative since $\widehat{F}_n \geq 0$ and $1 - \widehat{\Gamma}_n^* \geq 0$ with probability approaching one. This property also implies the non-decreasing property as $\widehat{\Phi}_n, \widehat{\Phi}_n^*$ are integrals. The continuity property also follows from the integral representation. By definition, $\widehat{\Phi}_n(0) = 0 = \widehat{\Phi}_n^*(0)$ and $\widehat{\Phi}_n(u_0^{\mathtt{MD}}) = \widehat{u}_n = \widehat{u}_n^* = \widehat{\Phi}_n^*(u_0^{\mathtt{MD}})$ with $I = [0, u_0^{\mathtt{MD}}]$. The closedness of the range follows from continuity and I being a compact interval.

(B8) follows from continuity of $\widehat{\Phi}_n$ and $\widehat{\Phi}_n^*$.

SA.5. Additional example: distribution function estimation with current status data. We consider the problem of estimating the cdf of X at x, $\theta_0(\mathsf{x}) = F_0(\mathsf{x})$. Observations $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ come from a random sample of $\mathbf{Z} = (\Delta, C, \mathbf{A}')'$ where $\Delta = \mathbb{1}\{X \leq C\}$, C is a random censoring time, and $\mathbf{A}$ is a vector of covariates. In this example, we do not observe $\check{X} = X \wedge C$. Instead, we observe the censoring time and whether the observation was censored. This setup is often referred to as current status data. Let $H_0(x) = \mathbb{P}[C \leq x]$ be the cdf of C. We can use $\Gamma_0(x) = \int_0^x F_0(u) H_0(du)$ and $\Phi_0(x) = H_0(x)$. The interval I is the support of X and $u_0 = 1$. We also assume H_0 admits a Lebesgue density h_0. The structure of the estimation problem turns out to be identical to the one for the monotone regression example, and we can leverage the common structure.

SA.5.1. *Independent right-censoring.* First we consider the case of completely at random censoring $X \perp\!\!\!\perp C$. See Groeneboom and Wellner (1992) for existing analysis. In this exaple, we do not use covariates $\mathbf{A}$. We set $\gamma_0(x; \mathbf{Z}) = \Delta \mathbb{1}\{C \leq x\}$ and $\phi_0(x; \mathbf{Z}) = \mathbb{1}\{C \leq x\}$. Note that if the notation is mapped by $(\Delta, C) \leftrightarrow (Y, X)$, then these functions are identical to those of the classical monotone regression problem (Corollary 3). Thus, the following result is identical to Corollary 3, up to notation and some changes due to boundedness of Δ.

COROLLARY SA-3. *Let $\varepsilon = \Delta - \mathbb{E}[\Delta|C]$ and x be an interior point of I. Suppose that Assumption BW holds, $\theta_0 = F_0$ satisfies Assumption (A2), the cdf $\Phi_0 = H_0$ satisfies Assumption (A3), and $\sigma_0^2(x) = \mathbb{E}[\varepsilon^2 | C = x]$ is continuous and positive at x. Then Assumptions A and B hold with*

$$\widehat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \Delta \mathbb{1}\{C \leq x\}, \quad \widehat{\Gamma}_n^*(x) = \frac{1}{n} \sum_{i=1}^n W_{i,n} \Delta \mathbb{1}\{C \leq x\},$$

$$\widehat{\Phi}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{C \leq x\}, \quad \widehat{\Phi}_n^*(x) = \frac{1}{n} \sum_{i=1}^n W_{i,n} \mathbb{1}\{C \leq x\}, \quad \widehat{u}_n = \widehat{u}_n^* = 1,$$

$$\mathcal{C}_{\mathsf{x}}(s,t) = h_0(\mathsf{x}) \sigma_0^2(\mathsf{x}) (|s| \wedge |t|) \mathbb{1}\{\mathrm{sign}(s) = \mathrm{sign}(t)\}, \qquad \mathcal{D}_{\mathfrak{q}}(\mathsf{x}) = \frac{h_0(\mathsf{x}) \partial^{\mathfrak{q}} F_0(\mathsf{x})}{(\mathfrak{q}+1)!}.$$

SA.5.2. *Conditionally independent right-censoring.* We consider the case where right-censoring is conditionally independent i.e., $X \perp\!\!\!\perp C | \mathbf{A}$. van der Vaart and van der Laan (2006) analyzed this example as well as settings with time-varying covariates. We are focusing on time-invariant covariates. Define $F_0(C, \mathbf{A}) = \mathbb{E}[\Delta | C, \mathbf{A}]$ and $g_0(C, \mathbf{A}) = \frac{h_{C|\mathbf{A}}(C|\mathbf{A})}{h_0(C)}$ where $h_{C|\mathbf{A}}$ is the conditional density of C given $\mathbf{A}$ and h_0 is the marginal density of C. Let $\widehat{F}_n(c, \mathbf{a})$ and $\widehat{g}_n(c, \mathbf{a})$ be preliminary estimators for $F_0(c, \mathbf{a})$ and $g_0(c, \mathbf{a})$, respectively.

Identical to the censoring completely at random case, with appropriate changes in the notation (i.e., $(\Delta, C) \leftrightarrow (Y, X)$), the setup is equivalent to that of the monotone regression with covariates.

Assumption SA.5.2. Let $\varepsilon = \Delta - \mathbb{E}[\Delta | C, \mathbf{A}]$, $\sigma_0^2(C, \mathbf{A}) = \mathbb{E}[\varepsilon^2 | C, \mathbf{A}]$, and $\eta > 0$ be some fixed number.

(i) x is in the interior of I, $\theta_0 = F_0$ satisfies (A2), and $\Phi_0 = H_0$ satisfies (A3).

(ii) The conditional distribution of C given $\mathbf{A}$ has a bounded Lebesgue density $h_{C|A}$, and there is $c > 0$ such that $g_0(C, \mathbf{A}) \geq c$ with probability one.

(iii) There exist $\delta > 0$ and random variables $A_n = o_{\mathbb{P}}(a_n^{-1/2})$, $B_n = O_{\mathbb{P}}(1)$ such that $\sqrt{n} \sup_{|v| \leq V} |\widehat{\Gamma}_n(\mathsf{x} + v) - \widehat{\Gamma}_n(\mathsf{x}) - \Gamma_0(\mathsf{x} + v) + \Gamma_0(\mathsf{x})| \leq A_n + B_n V$ for $V \in (0, 2\delta]$. For each $n \geq 1$, $(\mathbf{Z}_1, \ldots, \mathbf{Z}_n, \widehat{F}_n, \widehat{g}_n) \perp\!\!\!\perp (W_{1,n}, \ldots, W_{n,n})$. Also, $a_n \frac{1}{n} \sum_{i=1}^n |\widehat{F}_n(C_i, \mathbf{A}_i) - F_0(C_i, \mathbf{A}_i)|^2 = o_{\mathbb{P}}(1)$, $a_n \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |\widehat{F}_n(C_i, \mathbf{A}_j) - F_0(C_i, \mathbf{A}_j)|^2 = o_{\mathbb{P}}(1)$, $a_n \frac{1}{n} \sum_{i=1}^n |\widehat{g}_n(C_i, \mathbf{A}_i) - g_0(C_i, \mathbf{A}_i)|^2 = o_{\mathbb{P}}(1)$.

(iv) There exists a real-valued function $\bar{F}$ such that $|F_0(c_1,\mathbf{A}) - F_0(c_2,\mathbf{A})| \le |c_1 - c_2|\bar{F}(\mathbf{A})$ for $|c_1 - c_2| \le \eta$ and $\mathbb{E}[\bar{F}(\mathbf{A})^2] < \infty$.
(v) $\mathbb{E}[\frac{\sigma_0^2(\mathsf{x},\mathbf{A})}{g_0(\mathsf{x},\mathbf{A})}] > 0$, and there are real-valued functions B, ω such that $\mathbb{E}[B(\mathbf{A})] < \infty$, $\lim_{\eta\downarrow 0}\omega(\eta) = 0$, and for $|x - \mathsf{x}| \le \eta$, $|\frac{\sigma_0^2(x,\mathbf{A})h_{C|A}(x|\mathbf{A})}{g_0(x,\mathbf{A})^2} - \frac{\sigma_0^2(\mathsf{x},\mathbf{A})h_{C|A}(\mathsf{x}|\mathbf{A})}{g_0(\mathsf{x},\mathbf{A})^2}| \le \omega(|x - \mathsf{x}|)B(\mathbf{A})$.

COROLLARY SA-4. *Under Assumption SA.5.2and Assumption BW, Assumptions A and B hold with*

$$\widehat{\Gamma}_n(x) = \frac{1}{n}\sum_{i=1}^n \widehat{\gamma}_n(x;\mathbf{Z}_i), \quad \widehat{\Gamma}_n^*(x) = \frac{1}{n}\sum_{i=1}^n W_{i,n}\widehat{\gamma}_n(x;\mathbf{Z}_i),$$

$$\widehat{\gamma}_n(x;\mathbf{Z}) = \mathbb{1}\{C \le x\}\left[\frac{\Delta - \widehat{F}_n(C,\mathbf{A})}{\widehat{g}_n(C,\mathbf{A})} + \frac{1}{n}\sum_{j=1}^n \widehat{F}_n(C,\mathbf{A}_j)\right],$$

$$\widehat{\Phi}_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{C_i \le x\}, \quad \widehat{\Phi}_n^*(x) = \frac{1}{n}\sum_{i=1}^n W_{i,n}\mathbb{1}\{C_i \le x\}, \quad \widehat{u}_n = \widehat{u}_n^* = 1,$$

$$\mathcal{C}_{\mathsf{x}}(s,t) = h_0(\mathsf{x})\mathbb{E}\left[\frac{\sigma_0^2(\mathsf{x},\mathbf{A})}{g_0(\mathsf{x},\mathbf{A})}\right](|s|\wedge|t|)\mathbb{1}\{\mathrm{sign}(s) = \mathrm{sign}(t)\}, \qquad \mathcal{D}_{\mathfrak{q}}(\mathsf{x}) = \frac{h_0(\mathsf{x})\partial^{\mathfrak{q}}F_0(\mathsf{x})}{(\mathfrak{q}+1)!}.$$

SA.5.3. *Proof of Corollaries SA-3 and SA-4.* As noted above, by mapping the notation $(\Delta, C) \leftrightarrow (Y, X)$, the arguments in Sections SA.1.13 and SA.1.14 directly apply to the current status estimators.

SA.6. Rule-of-thumb step size selection. Here we develop a rule-of-thumb procedure to choose a step size for the bias-reduced numerical derivative estimator in the context of isotonic regression without covariates. Specifically, we consider the numerical derivative estimator

$$\widetilde{\mathcal{D}}_{j,n}^{\mathtt{BR}}(\mathsf{x}) = \epsilon_n^{-(j+1)}\sum_{k=1}^{\underline{\mathfrak{s}}+1}\lambda_j^{\mathtt{BR}}(k)[\widehat{\Upsilon}_n(\mathsf{x} + c_k\epsilon_n) - \widehat{\Upsilon}_n(\mathsf{x})]$$

with $\underline{\mathfrak{s}} = 3$, $c_1 = 1, c_2 = -1, c_3 = 2, c_4 = -2$. Then,

$$\lambda_1^{\mathtt{BR}}(1) = \frac{2}{3} = \lambda_1^{\mathtt{BR}}(2), \quad \lambda_1^{\mathtt{BR}}(3) = -\frac{1}{24} = \lambda_1^{\mathtt{BR}}(4),$$

$$\lambda_3^{\mathtt{BR}}(1) = -\frac{1}{6} = \lambda_3^{\mathtt{BR}}(2), \quad \lambda_3^{\mathtt{BR}}(3) = \frac{1}{24} = \lambda_3^{\mathtt{BR}}(4).$$

We use the (asymptotic) MSE-optimal step size discussed in the main paper. See also SA.1.7. Yet, with the choice of c_k's, part of the bias constant $\sum_{k=1}^{\underline{\mathfrak{s}}+1}\lambda_j^{\mathtt{BR}}(k)c_k^{\underline{\mathfrak{s}}+2}$ equals zero, and we need to turn to the next leading term of the bias, which is

$$\epsilon_n^{\underline{\mathfrak{s}}+2-j}\frac{\partial^{\underline{\mathfrak{s}}+3}\Upsilon_0(\mathsf{x})}{(\underline{\mathfrak{s}}+3)!}\sum_{k=1}^{\underline{\mathfrak{s}}+1}\lambda_j^{\mathtt{BR}}(k)c_k^{\underline{\mathfrak{s}}+3}.$$

Then, letting $\mathsf{B}_j^{\mathtt{BR}}(\mathsf{x}) = \frac{\partial^6\Upsilon_0(\mathsf{x})}{6!}\sum_{k=1}^4\lambda_j^{\mathtt{BR}}(k)c_k^6$, the MSE-optmal step size is

$$\epsilon_{j,n}^{\mathtt{BR}} = \left(\frac{(2j+1)\mathsf{V}_j^{\mathtt{BR}}(\mathsf{x})}{2(5-j)\mathsf{B}_j^{\mathtt{BR}}(\mathsf{x})^2}\right)^{1/11} n^{-1/11}.$$

The bias and variance constants depend on unknown features of the data generating process. Specifically, $\mathsf{B}_j^{\mathtt{BR}}(\mathsf{x})$ depends on the regression function θ_0, the Lebesgue density of X, and their derivatives at $X = \mathsf{x}$ while $\mathsf{V}_j^{\mathtt{BR}}(\mathsf{x})$ is determined by the density of X and the conditional variance of the regression error $\varepsilon = Y - \theta_0(X)$ at $X = \mathsf{x}$. To operationalize the construction of the step size, we posit a simple parametric model:

$$\mathbb{E}[Y|X] = \gamma_0 + \sum_{k=1}^{5} \gamma_k (X - \mathsf{x}_0)^k, \quad X \sim \mathrm{Normal}(\mu, \sigma^2)$$

where $\{\gamma_0, \gamma_1, \gamma_2, \gamma_4, \gamma_4, \gamma_5, \mu, \sigma\}$ are parameters to be estimated. Once we estimate the parameters of this reference model, we can construct a rule-of-thumb step size $\epsilon_{j,n}^{\mathtt{ROT}}$ by replacing $\mathsf{B}_j^{\mathtt{BR}}(\mathsf{x})$ and $\mathsf{V}_j^{\mathtt{BR}}(\mathsf{x})$ with their estimates. Note that although the bias and variance constant estimators may not be consistent for the true $\mathsf{B}_j^{\mathtt{BR}}(\mathsf{x})$ and $\mathsf{V}_j^{\mathtt{BR}}(\mathsf{x})$, the rate of $\epsilon_{j,n}^{\mathtt{ROT}}$ is MSE-optimal, and the numerical derivative estimator converges to $\mathcal{D}_j(\mathsf{x})$ sufficiently fast to satisfy Equation (11) in the main paper.

REFERENCES

BURKE, M. D., CSÖRGŐ, S. and HORVÁTH, L. (1988). A Correction to and Improvement of 'Strong Approximations of Some Biometric Estimates under Random Censorship'. *Probability Theory and Related Fields* **79** 51–57.

CATTANEO, M. D., FARRELL, M. and FENG, Y. (2020). Large Sample Properties of Partitioning-Based Series Estimators. *Annals of Statistics* **48** 1718-1741.

CATTANEO, M. D., JANSSON, M. and MA, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association* **115** 1449-1455.

CATTANEO, M. D., CHANDAK, R., JANSSON, M. and MA, X. (2024). Boundary Adaptive Local Polynomial Conditional Density Estimators. *Bernoulli*.

COX, G. (2022). A Generalized Argmax Theorem with Applications. arXiv:2209.08793.

GROENEBOOM, P. and WELLNER, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Birkhäuser Basel.

JAIN, N. C. and MARCUS, M. B. (1978). Continuity of Subgaussian Processes. In *Probability in Banach Spaces. Advances in Probability and Related Topics, Volume 4* 81–196. Dekker.

KIM, J. and POLLARD, D. (1990). Cube Root Asymptotics. *Annals of Statistics* **18** 191–219.

KOSOROK, M. R. (2008). Bootstrapping the Grenander Estimator. In *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen* 282–292. Institute of Mathematical Statistics.

LO, S.-H. and SINGH, K. (1986). The Product-Limit Estimator and the Bootstrap: Some Asymptotic Representations. *Probability Theory and Related Fields* **71** 455–465.

POLLARD, D. (1989). Asymptotics via Empirical Processes. *Statistical Science* **4** 341–354.

SEN, B., BANERJEE, M. and WOODROOFE, M. (2010). Inconsistency of Bootstrap: The Grenander Estimator. *Annals of Statistics* **38** 1953–1977.

VAN DER LAAN, M. J. and ROBINS, J. M. (2003). *Unified Methods for Censored longitudinal Data and Causality*. Springer Science & Business Media.

VAN DER VAART, A. W. and VAN DER LAAN, M. J. (2006). Estimating a Survival Distribution with Current Status Data and High-Dimensional Covariates. *International Journal of Biostatistics* **2**.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer.

WESTLING, T. and CARONE, M. (2020). A Unified Study of Nonparametric Inference for Monotone Functions. *Annals of Statistics* **48** 1001–1024.

WESTLING, T., GILBERT, P. and CARONE, M. (2020). Causal isotonic regression. *Journal of the Royal Statistical Society, Series B* **82** 719–747.

ZENG, D. (2004). Estimating Marginal Survival Function by Adjusting for Dependent Censoring using Many Covariates. *Annals of Statistics* **32** 1533–1555.