Supplementary Materials to "A Unified Framework for Identification and Inference of Local Treatment Effects in Sharp Regression Kink Designs"

Zhixin Wang, Zhengyu Zhang

This supplementary material provides the mathematical proofs for the identification results and asymptotic properties presented in the main text.

Contents

S.1 PROOFS FOR IDENTIFICATION

S.1.1 Proof of Lemma 1

Proof. Under Assumption S1, we have

$$\begin{aligned}\Delta_\phi &= \lim_{\delta\to 0}\frac{\phi\big(F_{Y(b_0)|X=x_0}+\delta h_\delta\big)-\phi\big(F_{Y(b_0)|X=x_0}\big)}{\delta}\\ &= \phi'_{F_{Y|X=x_0}}\big(\Delta_{Id}\big),\end{aligned} \tag{S.1}$$

where $h_\delta := (F_{Y(b_0+\delta)|X=x_0}-F_{Y(b_0)|X=x_0})/\delta$ and $\Delta_{Id}=\lim_{\delta\to 0}h_\delta$. Note that

$$\begin{aligned}&F_{Y(b_0+\delta)|X}(y|x_0)-F_{Y(b_0)|X}(y|x_0)\\ &\overset{(1)}{=} P\left[g(b_0+\delta,x_0,\varepsilon)\leqslant y\middle|X=x_0\right]-P\left[g(b_0,x_0,\varepsilon)\leqslant y\middle|X=x_0\right]\\ &\overset{(2)}{=} P\left[Y\leqslant y-\{g(b_0+\delta,x_0,\varepsilon)-g(b_0,x_0,\varepsilon)\}\middle|X=x_0\right]-P\left[Y\leqslant y\middle|X=x_0\right]\\ &\overset{(3)}{=} P\left[y<Y\leqslant y-\{g(b_0+\delta,x_0,\varepsilon)-g(b_0,x_0,\varepsilon)\}\middle|X=x_0\right]\\ &\qquad -P\left[y-\{g(b_0+\delta,x_0,\varepsilon)-g(b_0,x_0,\varepsilon)<Y\leqslant y\}\middle|X=x_0\right]\\ &\overset{(4)}{=} P\left[y\leqslant Y\leqslant y-\delta g_1(b_0,x_0,\varepsilon)\middle|X=x_0\right]\\ &\qquad -P\left[y-\delta g_1(b_0,x_0,\varepsilon)\leqslant Y\leqslant y\middle|X=x_0\right]+o(\delta),\end{aligned}$$

where equalities (1)–(2) follow from Assumptions 1; equality (3) follows from the law of total probability; equality (4) follows from Assumptions S2 and R1(i).

Then, as $\delta\to 0^+$, it follows that:

$$\begin{aligned}&F_{Y(b_0+\delta)|X}(y|x_0)-F_{Y(b_0)|X}(y|x_0)\\ &= P\left[Y\geqslant y,\ g_1(b_0,x_0,\varepsilon)\leqslant\frac{y-Y}{\delta}\middle|X=x_0\right]\\ &\qquad -P\left[Y\leqslant y,\ g_1(b_0,x_0,\varepsilon)\geqslant\frac{y-Y}{\delta}\middle|X=x_0\right]+o(\delta)\\ &= \int_y^\infty\int_{-\infty}^{(y-a)/\delta} f_{Y,g_1|X}(a,y'|x_0)\,dy'\,da\\ &\qquad -\int_{-\infty}^{y}\int_{(y-a)/\delta}^{\infty} f_{Y,g_1|X}(a,y'|x_0)\,dy'\,da+o(\delta)\\ &\overset{(5)}{=} \delta\int_{-\infty}^{0}\int_{-\infty}^{u} f_{Y,g_1|X}(y-\delta u,y'|x_0)\,dy'\,du\\ &\qquad -\delta\int_0^\infty\int_u^\infty f_{Y,g_1|X}(y-\delta u,y'|x_0)\,dy'\,du+o(\delta)\\ &\overset{(6)}{=} \delta\int_{-\infty}^{0}\int_{y'}^{0} f_{Y,g_1|X}(y-\delta u,y'|x_0)\,du\,dy'\end{aligned}$$

$$
\begin{aligned}
&- \delta \int_0^{\infty} \int_0^{y'} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' + o(\delta) \\
&= \delta \int_{-\infty}^{\infty} \int_{y'}^{0} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' + o(\delta),
\end{aligned}
$$

where equality (5) follows from the change of variable $u = (y - a)/\delta$ (for $\delta > 0$); equality (6) follows from interchanging the order of integration.

Similarly, as $\delta \to 0^-$, we have

$$
\begin{aligned}
&F_{Y(b_0+\delta)|X}(y|x_0) - F_{Y(b_0)|X}(y|x_0) \\
&= P\left[Y \geqslant y,\ g_1(b_0, x_0, \varepsilon) \geqslant \frac{y - Y}{\delta} \middle| X = x_0\right] \\
&\quad - P\left[Y \leqslant y,\ g_1(b_0, x_0, \varepsilon) \leqslant \frac{y - Y}{\delta} \middle| X = x_0\right] + o(\delta) \\
&= \int_y^{\infty} \int_{(y-a)/\delta}^{\infty} f_{Y,g_1|X}(a, y'|x_0)\, dy'\, da \\
&\quad - \int_{-\infty}^{y} \int_{-\infty}^{(y-a)/\delta} f_{Y,g_1|X}(a, y'|x_0)\, dy'\, da + o(\delta) \\
&= -\delta \int_0^{\infty} \int_u^{\infty} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, dy'\, du \\
&\quad + \delta \int_{-\infty}^{0} \int_{-\infty}^{u} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, dy'\, du + o(\delta) \\
&= -\delta \int_0^{\infty} \int_0^{y'} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' \\
&\quad + \delta \int_{-\infty}^{0} \int_{y'}^{0} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' + o(\delta) \\
&= \delta \int_{-\infty}^{\infty} \int_{y'}^{0} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' + o(\delta).
\end{aligned}
$$

Hence, the results for the cases $\delta \to 0^+$ and $\delta \to 0^-$ are the same. Then, dividing by δ and letting $\delta \to 0$ yields that

$$
\begin{aligned}
\Delta_{Id}(y) :&= \lim_{\delta \to 0} \frac{F_{Y(b_0+\delta)|X}(y|x_0) - F_{Y(b_0)|X}(y|x_0)}{\delta} \\
&= \lim_{\delta \to 0} \int_{-\infty}^{\infty} \int_{y'}^{0} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' \\
&\overset{(7)}{=} \int_{-\infty}^{\infty} \int_{y'}^{0} \lim_{\delta \to 0} f_{Y,g_1|X}(y - \delta u, y'|x_0)\, du\, dy' \\
&\overset{(8)}{=} \int_{-\infty}^{\infty} (-y') f_{Y,g_1|X}(y, y'|x_0)\, dy' \\
&= E\big[- f_{Y|X}(y|x_0) \cdot g_1(b_0, x_0, \varepsilon) \big| Y = y, X = x_0 \big].
\end{aligned}
$$

Here, the equalities in (7)–(8) are justified by an application of the dominated convergence theorem, which is permissible under Assumption R1(ii). Substituting this final expression into Equation (S.1) yields the desired result and completes the proof. ■

S.1.2 Proof of Theorem 1

Proof. Our proof strategy is to first identify the parameter $E[g_1(b_0, x_0, \varepsilon)|Y = y, X = x_0]$. The conclusion will then follow from an application of Lemma 1. To simplify notation, we adopt the shorthand $y_\tau(x) := Q_{Y|X}(\tau|x)$. For any $x \in I^o_{x_0}$ and any $t > 0$ such that $x + t \in I^o_{x_0}$, it follows from Assumptions 1 and S4 that:

$$\begin{aligned} 0 &= P\left[Y \leqslant y_\tau(x+t)\middle|X = x+t\right] - P\left[Y \leqslant y_\tau(x)\middle|X = x\right] \\ &= P\left[h(x+t,\varepsilon) \leqslant y_\tau(x+t)\middle|X = x+t\right] - P\left[h(x,\varepsilon) \leqslant y_\tau(x)\middle|X = x\right] \\ &= A_{1,t} + A_{2,t} + A_{3,t}, \end{aligned} \tag{S.2}$$

where $h(x, e) := g(b(x), x, e)$ and

$$\begin{aligned} A_{1,t} &:= P\left[h(x+t,\varepsilon) \leqslant y_\tau(x+t)\middle|X = x+t\right] - P\left[h(x+t,\varepsilon) \leqslant y_\tau(x)\middle|X = x+t\right], \\ A_{2,t} &:= P\left[h(x+t,\varepsilon) \leqslant y_\tau(x)\middle|X = x+t\right] - P\left[h(x+t,\varepsilon) \leqslant y_\tau(x)\middle|X = x\right], \\ A_{3,t} &:= P\left[h(x+t,\varepsilon) \leqslant y_\tau(x)\middle|X = x\right] - P\left[h(x,\varepsilon) \leqslant y_\tau(x)\middle|X = x\right]. \end{aligned}$$

Dividing both sides of Equation (S.2) by t and taking the limit as $t \to 0^+$ yields:

$$0 = \lim_{t\to 0^+} A_{1,t}/t + \lim_{t\to 0^+} A_{2,t}/t + \lim_{t\to 0^+} A_{3,t}/t \tag{S.3}$$

provided each of these limits exists. We now analyze each term on the right-hand side in turn.

We begin with the term $A_{1,t}$. By the mean value theorem for integrals, which is applicable under Assumptions 1 and S4, we have:

$$\begin{aligned} A_{1,t} &= P\big[Y \leqslant y_\tau(x+t)\big|X = x+t\big] - P\big[Y \leqslant y_\tau(x)\big|X = x+t\big] \\ &= \int_{y_\tau(x)}^{y_\tau(x+t)} f_{Y|X}(y|x+t)\,dy \\ &= \big(y_\tau(x+t) - y_\tau(x)\big) \cdot f_{Y|X}(y_{\tau,t}(x)|x+t) \end{aligned}$$

for some value $y_{\tau,t}(x)$ lying between $y_\tau(x)$ and $y_\tau(x+t)$. Next, differentiating both sides of the identity $F_{Y|X}(y_\tau(x)|x) \equiv \tau$ with respect to x, which is permissible under the smoothness conditions of Assumptions S4 and R1(iii), yields that $\frac{\partial}{\partial x} y_\tau(x) \cdot f_{Y|X}(y_\tau(x)|x) = -\frac{\partial}{\partial x} F_{Y|X}(y_\tau(x)|x)$. Thus,

$$\lim_{t\to 0^+} A_{1,t}/t = \lim_{t\to 0^+} \left\{ \frac{y_\tau(x+t) - y_\tau(x)}{t} \cdot f_{Y|X}(y_{\tau,t}(x)|x+t) \right\}$$

$$= -\frac{\partial}{\partial x} F_{Y|X}(y_\tau(x)|x). \tag{S.4}$$

For the term $A_{2,t}$, we note that

$$A_{2,t} = \int \mathbb{1}\left(h(x+t, e) \leqslant y_\tau(x)\right) \cdot \left[f_{\varepsilon|X}(e|x+t) - f_{\varepsilon|X}(e|x)\right] de.$$

Assumption R1(v) implies that $f_{\varepsilon|X} = \int f_{\varepsilon|Y,X}\, dF_{Y|X}$ is dominated by the function $C|\chi_1|$ for some positive constant C. Then, under Assumptions S3 and R1(v), it follows from the dominated convergence theorem that

$$\begin{aligned}
\lim_{t \to 0^+} A_{2,t}/t &= \int \lim_{t \to 0^+} \mathbb{1}\left(h(x+t, e) \leqslant y_\tau(x)\right) \cdot \frac{f_{\varepsilon|X}(e|x+t) - f_{\varepsilon|X}(e|x)}{t}\, de \\
&= \int_{\mathcal{V}(y_\tau(x), x)} \frac{\partial}{\partial x} f_{\varepsilon|x}(e|x)\, de, \qquad \text{(S.5)}
\end{aligned}$$

where $\mathcal{V}(y, x) := \{e \in \mathbb{R}^{d_\varepsilon} : h(x, e) \leqslant y\}$.

For the term $A_{3,t}$, as $t \to 0^+$, it follows that

$$\begin{aligned}
A_{3,t} &\overset{(1)}{=} P\left[Y \leqslant y_\tau(x) - \{h(x+t, \varepsilon) - h(x, \varepsilon)\} \middle| X = x\right] \\
&\quad - P\left[Y \leqslant y_\tau(x) \middle| X = x\right] \\
&\overset{(2)}{=} P\left[y_\tau(x) < Y \leqslant y_\tau(x) - \{h(x+t, \varepsilon) - h(x, \varepsilon)\} \middle| X = x\right] \\
&\quad - P\left[y_\tau(x) - \{h(x+t, \varepsilon) - h(x, \varepsilon)\} < Y \leqslant y_\tau(x) \middle| X = x\right] \\
&\overset{(3)}{=} P\left[y_\tau(x) \leqslant Y \leqslant y_\tau(x) - t h_x(x, \varepsilon) \middle| X = x\right] \\
&\quad - P\left[y_\tau(x) - t h_x(x, \varepsilon) \leqslant Y \leqslant y_\tau(x) \middle| X = x\right] + o(t) \qquad (h_x := \tfrac{\partial}{\partial x} h) \\
&= P\left[Y \geqslant y_\tau(x),\, h_x(x, \varepsilon) \leqslant \frac{y_\tau(x) - Y}{t} \middle| X = x\right] \\
&\quad - P\left[Y \leqslant y_\tau(x),\, h_x(x, \varepsilon) \geqslant \frac{y_\tau(x) - Y}{t} \middle| X = x\right] + o(t) \\
&= \int_{y_\tau(x)}^{\infty} \int_{-\infty}^{(y_\tau(x) - y)/t} f_{Y, h_x|X}(y, y'|x)\, dy'\, dy \\
&\quad - \int_{-\infty}^{y_\tau(x)} \int_{(y_\tau(x) - y)/t}^{\infty} f_{Y, h_x|X}(y, y'|x)\, dy'\, dy + o(t) \\
&\overset{(4)}{=} t \int_{-\infty}^{0} \int_{-\infty}^{u} f_{Y, h_x|X}(y_\tau(x) - tu, y'|x)\, dy'\, du \\
&\quad - t \int_{0}^{\infty} \int_{u}^{\infty} f_{Y, h_x|X}(y_\tau(x) - tu, y'|x)\, dy'\, du + o(t) \\
&\overset{(5)}{=} t \int_{-\infty}^{0} \int_{y'}^{0} f_{Y, h_x|X}(y_\tau(x) - \delta u, y'|x)\, du\, dy'
\end{aligned}$$

$$
\begin{aligned}
&- t \int_0^{\infty} \int_0^{y'} f_{Y,h_x|X}(y_\tau(x) - \delta u, y'|x)\, du\, dy' + o(t) \\
&= t \int_{-\infty}^{\infty} \int_{y'}^{0} f_{Y,h_x|X}(y_\tau(x) - \delta u, y'|x)\, du\, dy' + o(t),
\end{aligned}
$$

where equality (1) follows from Assumption 1; equality (2) follows from the law of total probability; equality (3) follows from Assumption R1(iii); equality (4) follows from the change of variable $u = (y_\tau(x) - y)/t$; equality (5) follows from interchanging the order of integration. Then, by the dominated convergence theorem, which is permissible under Assumption R1(iv), it follows that:

$$
\begin{aligned}
\lim_{t \to 0^+} A_{3,t}/t &= \int_{-\infty}^{\infty} \int_{y'}^{0} \lim_{t \to 0^+} f_{Y,h_x|X}(y_\tau(x) - \delta u, y'|x)\, du\, dy' \\
&= \int_{-\infty}^{\infty} (-y') f_{Y,h_x|X}(y_\tau(x), y'|x)\, dy' \\
&= E\left[h_x(x,\varepsilon) \middle| Y = y_\tau(x), X = x\right] \cdot \big(- f_{Y|X}(y_\tau(x)|x)\big). \qquad \text{(S.6)}
\end{aligned}
$$

Hence, substituting Equations (S.4)–(S.6) into Equation (S.3) yields:

$$
\frac{\partial}{\partial x} F_{Y|X}(y_\tau(x)|x) = E\left[h_x(x,\varepsilon) \middle| Y = y_\tau(x), X = x\right] \cdot \big(-f_{Y|X}(y_\tau(x)|x)\big) + \int_{\mathcal{V}(y_\tau(x),x)} \frac{\partial}{\partial x} f_{\varepsilon|x}(e|x)\, de.
$$

Under Assumptions S3–S4, R1(iii), and (v), it follows as $x \to x_0^+$ that:

$$
\begin{aligned}
\frac{\partial}{\partial x} F_{Y|X}(y_\tau|x_0^+) \;&=\; \lim_{x \to x_0^+} \int h_x(x,e) f_{\varepsilon|Y,X}(e|y_\tau, x)\, de \cdot \big(- f_{Y|X}(y_\tau(x)|x)\big) \\
&\quad + \lim_{x \to x_0^+} \int_{\mathcal{V}(y_\tau(x),x)} \frac{\partial}{\partial x} f_{\varepsilon|x}(e|x)\, de \\
&\overset{(6)}{=} \int \lim_{x \to x_0^+} \left\{ \frac{\partial}{\partial x} g(b(x), x, e) \right\} f_{\varepsilon|Y,X}(e|y_\tau(x), x)\, de \cdot \big(- f_{Y|X}(y_\tau|x_0)\big) \\
&\quad + \int \lim_{x \to x_0^+} \mathbb{1}\left(h(x,e) \leqslant y_\tau(x)\right) \frac{\partial}{\partial x} f_{\varepsilon|x}(e|x)\, de \\
&= b'(x_0^+) \int g_1(b_0, x_0, e) f_{\varepsilon|Y,X}(e|y_\tau, x_0)\, de \cdot \big(- f_{Y|X}(y_\tau|x_0)\big) \qquad (b_0 := b(x_0)) \\
&\quad + \int g_2(b_0, x_0, e) f_{\varepsilon|Y,X}(e|y_\tau, x_0)\, de \cdot \big(- f_{Y|X}(y_\tau|x_0)\big) \qquad (y_\tau := y_\tau(x_0)) \\
&\quad + \int \mathbb{1}\left(g(b_0, x_0, e) \leqslant y_\tau\right) \frac{\partial}{\partial x} f_{\varepsilon|x}(e|x_0)\, de,
\end{aligned}
$$

where equality (6) follows from the dominated convergence theorem under Assumption R1(v). Similarly, it follows as $x \to x_0^-$ that

$$
\frac{\partial}{\partial x} F_{Y|X}(y_\tau|x_0^-) = b'(x_0^-) \int g_1(b_0, x_0, e) f_{\varepsilon|Y,X}(e|y_\tau, x_0)\, de \cdot \big(- f_{Y|X}(y_\tau|x_0)\big)
$$

$$+\int g_2(b_0,x_0,e) f_{\varepsilon|Y,X}(e|y_\tau,x_0)\,de\cdot\big(-f_{Y|X}(y_\tau|x_0)\big)$$
$$+\int \mathbb{1}\left(g(b_0,x_0,e)\leqslant y_\tau\right)\frac{\partial}{\partial x} f_{\varepsilon|x}(e|x_0)\,de.$$

Taking the difference then yields:

$$\begin{aligned}
&\frac{\partial}{\partial x}F_{Y|X}(y_\tau|x_0^+)-\frac{\partial}{\partial x}F_{Y|X}(y_\tau|x_0^-)\\
&=\big[b'(x_0^+)-b'(x_0^-)\big]\cdot E\big[-f_{Y|X}(y_\tau|x_0)\cdot g_1(b_0,x_0,e)|Y=y_\tau,X=x_0\big].
\end{aligned}$$

Under Assumptions 2 and S4, dividing both sides of this equation by $b'(x_0^+)-b'(x_0^-)$ and evaluating at $\tau=F_{Y|X}(y|x_0)$ yields

$$\frac{\frac{\partial}{\partial x}F_{Y|X}(y|x_0^+)-\frac{\partial}{\partial x}F_{Y|X}(y|x_0^-)}{b'(x_0^+)-b'(x_0^-)}=E\big[-f_{Y|X}(y|x_0)\cdot g_1(b_0,x_0,e)\big|Y=y,X=x_0\big]$$

for all $y\in\mathcal{Y}_{x_0}$, where the left-hand side is DRKD(y). Finally, substituting this result into Lemma 1 completes the proof. ■

S.1.3 Derivation of Example 1

Proof. The mean functional $\mu(F)=\int w\,dF(w)$ is linear, so its Hadamard derivative at F is the functional itself. Applying Theorem 1 with $\phi=\mu$ yields:

$$\begin{aligned}
\Delta_\mu &= \mu'_{F_{Y|X=x_0}}(\mathrm{DRKD})\\
&=\int y\,d\mathrm{DRKD}(y)\\
&=\frac{1}{b'(x_0^+)-b'(x_0^-)}\left[\int y\,d\left(\frac{\partial}{\partial x}F_{Y|X}(y|x_0^+)\right)-\int y\,d\left(\frac{\partial}{\partial x}F_{Y|X}(y|x_0^-)\right)\right].
\end{aligned}$$

Note that

$$\begin{aligned}
\int y\,d\left(\frac{\partial}{\partial x}F_{Y|X}(y|x_0^+)\right) &\overset{(1)}{=} y\frac{\partial}{\partial x}F_{Y|X}(y|x_0^+)-\int\frac{\partial}{\partial x}F_{Y|X}(y|x_0^+)\,dy\\
&\overset{(2)}{=}\lim_{x\to x_0^+} y\frac{\partial}{\partial x}F_{Y|X}(y|x)-\lim_{x\to x_0^+}\frac{\partial}{\partial x}\int F_{Y|X}(y|x)\,dy\\
&\overset{(3)}{=}\lim_{x\to x_0^+} y\frac{\partial}{\partial x}F_{Y|X}(y|x)-\lim_{x\to x_0^+}\frac{\partial}{\partial x}\left\{yF_{Y|X}(y|x)-\int y\,dF_{Y|X}(y|x)\right\}\\
&=\lim_{x\to x_0^+}\frac{\partial}{\partial x}E\left[Y|X=x\right],
\end{aligned}$$

where equalities (1) and (3) follow from integration by parts; equality (2) follows from the dominated convergence theorem. By similar reasoning, it follows that

$$\int y\, d\left(\frac{\partial}{\partial x}F_{Y|X}(y|x_0^-)\right) = \lim_{x\to x_0^-}\frac{\partial}{\partial x}E\left[Y|X=x\right].$$

This completes the proof. ■

S.1.4 Derivation of Example 2

Proof. The result for the distributional effect, $\Delta_{Id}(y) = \mathrm{DRKD}(y)$, is straightforward. For the quantile effect Δ_Q, if $\mathcal{Y}_{x_0}$ is regularized to a compact set, then Lemma 21.4 of Van der Vaart (2000) implies that the quantile functional Q_τ is Hadamard differentiable at $F_{Y|X=x_0}$, with its derivative given by:

$$[Q_\tau]'_{F_{Y|X=x_0}}(h) = -\frac{h(y_\tau)}{f_{Y|X}(y_\tau|x_0)}$$

for all $\tau \in (0,1)$, where $y_\tau := Q_{Y|X}(\tau|x_0)$. Under Assumption S4, differentiating both sides of the identity $F_{Y|X}(y_\tau(x)|x) \equiv \tau$ with respect to x yields: $\frac{\partial}{\partial x}F_{Y|X}(y_\tau(x)|x) = -\frac{\partial}{\partial x}Q_{Y|X}(\tau|x)f_{Y|X}(y_\tau(x)|x)$. Then, it follows that

$$\begin{aligned}
&\lim_{x\to x_0^+}\frac{\partial}{\partial x}F_{Y|X}(y_\tau(x)|x) - \lim_{x\to x_0^-}\frac{\partial}{\partial x}F_{Y|X}(y_\tau(x)|x)\\
&= -\left[\lim_{x\to x_0^+}\frac{\partial}{\partial x}Q_{Y|X}(\tau|x) - \lim_{x\to x_0^-}\frac{\partial}{\partial x}Q_{Y|X}(\tau|x)\right] f_{Y|X}(y_\tau|x_0).
\end{aligned}$$

Dividing both sides by $b'(x_0^+)-b'(x_0^-)$ implies that $\mathrm{DRKD}(y_\tau) = -\mathrm{QRKD}(\tau)f_{Y|X}(y_\tau|x_0)$. Therefore,

$$\Delta_Q(\tau) = [Q_\tau]'_{F_{Y|X=x_0}}(\mathrm{DRKD}) = -\frac{\mathrm{DRKD}(y_\tau)}{f_{Y|X}(y_\tau|x_0)} = \mathrm{QRKD}(\tau),$$

This completes the proof. ■

S.1.5 Derivation of Example 3

Proof. For the Lorenz effect Δ_L, the Hadamard differentiability of the Lorenz functional $L_\tau(F)$ depends on that of the underlying quantile functional Q_τ. This latter condition is satisfied, for example, under the regularity conditions of Proposition 2 in Bhattacharya (2007) . The Hadamard derivative is then given by:

$$[L_\tau]'_F(h) == \frac{\int_0^\tau [Q_p]'_F(h)\,dp}{\mu(F)} - \frac{L_\tau(F)}{\mu(F)}\cdot\mu(h).$$

Then, it follows that

$$\Delta_L(\tau) = [L_\tau]'_F\,(\text{DRKD}) = \frac{\int_0^\tau \text{QRKD}(p)\,dp - L_{Y|X}(\tau|x_0)\cdot \text{MRKD}}{\mu_0}.$$

This completes the proof. ■

S.1.6 Proof of Proposition 1

Proof. For the part (a), it follows from the proof of Lemma 1 in Sasaki (2015) that:

$$\begin{aligned}
f_{Y|X}(y_\tau(x)|x) &\overset{(1)}{=} \frac{\partial}{\partial y} P\left[h(x,\varepsilon) \leqslant y \middle| X = x\right]\Big|_{y=y_\tau(x)} \\
&\overset{(2)}{=} \frac{\partial}{\partial y}\int_{\mathcal{V}(y,x)} f_{\varepsilon|X}(e|x)\,de\Big|_{y=y_\tau(x)} \\
&\overset{(3)}{=} \int_{\partial\mathcal{V}(y_\tau(x),x)} \frac{1}{\|\nabla_e h(x,e)\|}\frac{f_{\varepsilon|X}(e|x)\cdot M\pi^{(M-1)/2}}{2^{M-1}\Gamma\left(\frac{M+1}{2}\right)}\,dH^{M-1}(e),
\end{aligned}$$

where $\mathcal{V}(y,x) := \{e \in \mathbb{R}^{d_\varepsilon} : h(x,e) \leqslant y\}$; Γ denotes the Gamma function. Equality (1) follows from Assumptions 1 and 2. Equalities (2) and (3) follow from Assumption S4'(i)–(ii). Then, it follows that

$$\begin{aligned}
\lim_{x\to x_0} f_{Y|X}(y_\tau(x)|x) &= \lim_{x\to x_0}\int_{\partial\mathcal{V}(y_\tau(x),x)} \frac{1}{\|\nabla_e h(x,e)\|}\frac{f_{\varepsilon|X}(e|x)\cdot M\pi^{(M-1)/2}}{2^{M-1}\Gamma\left(\frac{M+1}{2}\right)}\,dH^{M-1}(e) \\
&= \frac{M\pi^{(M-1)/2}}{2^{M-1}\Gamma\left(\frac{M+1}{2}\right)}\lim_{x\to x_0}\tilde{f}(y_\tau(x),x),
\end{aligned}$$

where $y_\tau := y_\tau(x_0)$. Hence, under Assumptions S4'(iii) and S5, the continuity of $\tilde{f}$ implies that the function $x \mapsto f_{Y|X}(y_\tau(x)|x)$ is continuous at $x = x_0$ and strictly positive there. For the part (b), we note that under Assumption S6,

$$\begin{aligned}
F_{Y|X}(y|x) &= \int P\big[h(x,\varepsilon_1,a) \leqslant y \big| A = a, X = x\big]\,dF_{A|X}(a|x) \\
&= \int P\big[\varepsilon_1 \leqslant h^{-1}(x,y,a) \big| A = a, X = x\big]\,dF_{A|X}(a|x) \\
&= \int F_{\varepsilon_1|A,X}(h^{-1}(x,y,a)|a,x) f_{A|X}(a|x)\,da.
\end{aligned}$$

Then, differentiating both sides with respect to y and evaluating at $y = y_\tau(x)$ yields:

$$\begin{aligned}
f_{Y|X}(y_\tau(x)|x) &= \frac{\partial}{\partial y}F_{Y|X}(y|x)\Big|_{y=y_\tau(x)} \\
&\overset{(4)}{=} \int f_{\varepsilon_1|A,X}(h^{-1}(x,y,a)|a,x)\cdot\frac{\partial h^{-1}(x,y,a)}{\partial y}\cdot f_{A|X}(a|x)\,da\Big|_{y=y_\tau(x)}
\end{aligned}$$

$$\stackrel{(5)}{=} \int f_{\varepsilon_1|A,X}(h^{-1}(x, y_\tau(x), a)|a, x) \cdot \frac{f_{A|X}(a|x)}{\frac{\partial}{\partial e_1} h(x, h^{-1}(x, y_\tau(x), a), a)} \, da$$

where equality (4) follows from Assumption S6(i); equality (5) follows from the inverse function theorem under Assumption S6(ii). Finally, by the dominated convergence theorem, it follows that

$$\begin{aligned}
\lim_{x \to x_0} f_{Y|X}(y_\tau(x)|x) &= \int \lim_{x \to x_0} \left\{ f_{\varepsilon_1|A,X}(h^{-1}(x, y_\tau(x), a)|a, x) \cdot \frac{f_{A|X}(a|x)}{\frac{\partial}{\partial e_1} h(x, h^{-1}(x, y_\tau(x), a), a)} \right\} da \\
&= \int f_{\varepsilon_1|A,X}(h^{-1}(x_0, y_\tau, a)|a, x_0) \cdot \frac{f_{A|X}(a|x_0)}{\frac{\partial}{\partial e_1} h(x, h^{-1}(x_0, y_\tau, a), a)} \, da.
\end{aligned}$$

Therefore, $x \mapsto f_{Y|X}(y_\tau(x)|x)$ is continuous at $x = x_0$ and strictly positive under Assumption S6. This completes the proof. ∎

S.1.7 Proof of Proposition 2

Proof. By a similar argument as in the proof of Theorem 1, it follows that:

$$\begin{aligned}
\frac{\partial}{\partial x} Q_{Y|X}(\tau|x) &= b'(x) E\left[g_1(b(x), x, \varepsilon)|Y = y_\tau(x), X = x\right] \\
&\quad + E\left[g_2(b(x), x, \varepsilon)|Y = y_\tau(x), X = x\right] \\
&\quad - \frac{\int_{\mathcal{V}(y_\tau(x),x)} \frac{\partial}{\partial x} f_{\varepsilon|x}(e|x) \, de}{f_{Y|X}(y_\tau(x)|x)} \\
&= b'(x) \int g_1(b(x), x, e) \frac{f_{Y,X|\varepsilon}(y_\tau(x), x|e)}{f_{Y,X}(y_\tau(x), x)} \, dF_\varepsilon(e) \\
&\quad + \int g_2(b(x), x, e) \frac{f_{Y,X|\varepsilon}(y_\tau(x), x|e)}{f_{Y,X}(y_\tau(x), x)} \, dF_\varepsilon(e) \\
&\quad - \frac{f_{Y,X}(y_\tau(x), x)}{f_X(x)} \int_{\mathcal{V}(y_\tau(x),x)} \frac{\partial}{\partial x} \frac{f_{X|\varepsilon}(x|e)}{f_X(x)} \, de.
\end{aligned}$$

First, note that under Assumptions 1, 2, and S3', Lemmas 1–2 of Card et al. (2015) imply that the functions $x \mapsto f_X(x)$, $x \mapsto f_{\varepsilon|X}(e|x)$, and $x \mapsto \frac{\partial}{\partial x} f_{\varepsilon|X}(e|x)$ are continuous on I_{x_0}. Furthermore, Assumption S7(ii) states that the function $(y, x) \mapsto f_{Y,X|\varepsilon}(y, x|e)$ is continuous. Combined with the continuity of $x \mapsto y_\tau(x)$ (implied by previous assumptions), this ensures the continuity of the composite function $x \mapsto f_{Y,X|\varepsilon}(y_\tau(x), x|e)$ on I_{x_0}. Therefore, for any $\eta > 0$ and any fixed e, by the definition of continuity, there exists a $\delta > 0$ such that for all $x_1, x_2 \in I_{x_0}$ with $|x_1 - x_2| < \delta$, $\left|f_{Y,X|\varepsilon}(y_\tau(x_1), x_1|e) - f_{Y,X|\varepsilon}(y_\tau(x_2), x_2|e)\right| < \eta / \int_{\mathcal{A}_\varepsilon(\tau)} dF_\varepsilon(e)$. Hence, for every $\eta > 0$, there exists $\delta > 0$ with $|x_1 - x_2| < \delta$, the following inequality holds:

$$|f_{Y,X}(y_\tau(x_1), x_1) - f_{Y,X}(y_\tau(x_2), x_2)| = \left| \int_{\mathcal{A}_\varepsilon(\tau)} \left[f_{Y,X|\varepsilon}(y_\tau(x_1), x_1|e) - f_{Y,X|\varepsilon}(y_\tau(x_2), x_2|e)\right] dF_\varepsilon(e) \right|$$

$$
\begin{aligned}
&\leqslant \int_{\mathcal{A}_\varepsilon(\tau)} \left| f_{Y,X|\varepsilon}(y_\tau(x_1), x_1|e) - f_{Y,X|\varepsilon}(y_\tau(x_2), x_2|e) \right| dF_\varepsilon(e) \\
&< \eta,
\end{aligned}
$$

where $\mathcal{A}_\varepsilon(\tau) := \{e \in \mathbb{R}^{d_\varepsilon} : f_{Y,X|\varepsilon}(y_\tau(x), x|e) > 0, \forall x \in I_{x_0}\}$. This implies $f_{Y,X}(y_\tau(x), x)$ is continuous in $x \in I_{x_0}$ for each $\tau \in (0,1)$. Furthermore, under Assumption S7(i), $f_{Y,X}(y_\tau(x), x) \geqslant \int_{\mathcal{A}_\varepsilon(\tau)} f_{Y,X|\varepsilon}(y_\tau(x), x|e)\, dF_\varepsilon(e) > 0$. Therefore, under Assumptions 1, 2, 3' S2, and R1(v), it follows from the dominated convergence theorem that

$$
\begin{aligned}
&\lim_{x \to x_0^+} \frac{\partial}{\partial x} Q_{Y|X}(\tau|x) - \lim_{x \to x_0^-} \frac{\partial}{\partial x} Q_{Y|X}(\tau|x) \\
&= \left[b'(x_0^+) - b'(x_0^-) \right] \int g_1(b(x_0), x_0, e) \frac{f_{Y,X|\varepsilon}(y_\tau, x_0|e)}{f_{Y,X}(y_\tau, x_0)}\, dF_\varepsilon(e).
\end{aligned}
$$

Dividing both sides by $b'(x_0^+) - b'(x_0^-)$ completes the proof. ■

S.2 PROOFS FOR ASYMPTOTIC THEORY

Notation: Let $(\Omega, \mathcal{S}, P)$ denote the underlying probability space; for every measurable $f : \Omega \to \mathbb{R}$, let $\|f\|_{P,p} := (\int |f|^p \, dP)^{1/p}$ and the $\mathcal{L}^2(P)$-semimetric $e_{P,2}(f,g) := \|f - g\|_{P,2}$. Let $(\mathbb{T}, \mathrm{d})$ be a semimetric space, $\ell^\infty(\mathbb{T})$ denote the space of all bounded functions $f : \mathbb{T} \to \mathbb{R}$ equipped with the uniform norm $\|f\|_{\mathbb{T}} := \sup_{t \in \mathbb{T}} |f(t)|$, where $|\cdot|$ is the standard Euclidean norm. Let $\mathbb{G}_n f := \frac{1}{\sqrt{n}} \sum_{i=1}^n (f(X_i) - E[f(X_i)])$ denote the empirical process indexed by $\mathcal{F} \ni f$, and $\|\mathbb{G}_n\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |\mathbb{G}_n f|$. $a_n \lesssim b_n$ means $a_n \leqslant C b_n$ for a constant $C > 0$ independent of n. Unless otherwise stated, we write $C > 0$ for a universal constant whose values may vary from place to place.

Definition S.2.1.

(i) (covering number) A ε-cover of $\mathbb{T}$ with respect to a semi-metric $\mathrm{d}(\cdot,\cdot)$ is defined as

$$
\{\{\theta_1, \ldots, \theta_N\} : \forall \theta \in \mathbb{T} \;\; \exists\, \nu \in \{1, 2 \ldots N\} \; s.t. \; \mathrm{d}(\theta, \theta_\nu) \leqslant \varepsilon\}.
$$

The ε*–covering number* of $\mathbb{T}$ is defined as

$$
N(\mathbb{T}, \mathrm{d}, \varepsilon) := \inf\{N \in \mathbb{N} : \exists\, \varepsilon\text{-cover } \{\theta_1, \ldots, \theta_N\} \text{ of } \mathbb{T}\}
$$

(ii) (Uniform entropy integral) The uniform entropy integral for a class $\mathcal{F}$ of measurable function with envelope F is defined as

$$
J(\delta, \mathcal{F}, F) := \int_0^\delta \sup_P \sqrt{1 + \log N(\mathcal{F}, e_{P,2}, \varepsilon \|F\|_{P,2})}\, d\varepsilon \quad \forall \delta > 0,
$$

where " sup" is taken over all probability measures P on $(\Omega, \mathcal{S})$ with $\|F\|_{P,2} > 0$.

(iii) (VC-type class) Let $\mathcal{F}$ be a class of measurable $f : \Omega \to \mathbb{R}$ with envelop F. Then $\mathcal{F}$ is a VC-type class with envelope F if there exist constants $A, v \geqslant 1$ such that

$$\sup_P N(\mathcal{F}, e_{P,2}, \varepsilon \|F\|_{P,2}) \leqslant \left(\frac{A}{\varepsilon}\right)^v, \quad \forall \varepsilon \in (0,1],$$

where "sup" is taken over all probability measures P on $(\Omega, \mathcal{S})$.

Let $\mathbb{T}$ denote an arbitrary index set, and $\odot$ denote the pointwise product.

Definition S.2.2. Let $\mathscr{F}_{n\omega} := \{(f_{n1}(\omega,t), \ldots, f_{nm_n}(\omega,t)) \in \mathbb{R}^{m_n} : t \in \mathbb{T}\}$. A triangular array of real-valued processes $\{f_{ni}(\omega,t) : t \in \mathbb{T}, 1 \leqslant i \leqslant n\}$ is *manageable* with respect to nonnegative envelopes $\boldsymbol{F}_n(\omega) := (F_{n1}(\omega), \ldots, F_{nm_n}(\omega)) \in \mathbb{R}^{m_n}$ if there exists a deterministic function $\lambda : \mathbb{R} \to \mathbb{R}$ (the capacity bound) such that

(i) $\int_0^1 \log \sqrt{\lambda(\epsilon)}\, d\epsilon < \infty$;

(ii) There exists $N \subset \Omega$ such that $P^*(N) = 0$ and for each $\omega \notin N$, [1]

$$D_{m_n}\left(\epsilon \|\alpha \odot \boldsymbol{F}_n(\omega)\|, \alpha \odot \mathscr{F}_{n\omega}\right) \leqslant \lambda(\epsilon)$$

for every $0 < \epsilon \leqslant 1$, $\alpha \in \mathbb{R}^{m_n}$ with $\alpha_j > 0$, $j = 1, \ldots, m_n$ and $n \geqslant 1$, where $\lambda(\cdot)$ does not depend on ω and n.

We present Pollard's functional central limit theorem (Pollard's FCLT) as follows. See also Theorem 1 of Kosorok (2003) or Theorem 11.16 of Kosorok (2008).

Lemma S.2.1 (Pollard, 1990, Theorem 10.6)**.** *Let $B_n(\omega,t) := \sum_{i=1}^{m_n} (f_{ni}(\omega,t) - E[f_{ni}(\cdot,t)])$. Suppose that the stochastic processes from the triangular array $\{f_{ni}(\omega,t) : t \in \mathbb{T}\}$ are independent within rows and satisfy*

(A) $\{f_{ni}\}$ are manageable with envelopes $\{F_{ni}\}$ which are also independent within rows;

(B) $\lim_{n\to\infty} E[B_n(\cdot,t_1)B_n(\cdot,t_2)]$ exists for every t_1, $t_2 \in \mathbb{T}$;

(C) $\limsup_{n\to\infty} \sum_{i=1}^{m_n} E[F_{ni}^2] < \infty$;

(D) $\lim_{n\to\infty} \sum_{i=1}^{m_n} E[F_{ni}^2]\mathbb{1}(F_{ni} > \epsilon) = 0$ for each $\epsilon > 0$;

(E) Let $\rho(t_1,t_2) := \lim_{n\to\infty} \rho_n(t_1,t_2)$, where

$$\rho_n(t_1,t_2) := \left(\sum_{i=1}^{m_n} E\big[f_{ni}(\cdot,t_1) - f_{ni}(\cdot,t_2)\big]^2\right)^{1/2}$$

[1] The packing number $D_m(\epsilon, \mathbb{T})$ is the largest integer k such that there exist k points in $\mathbb{T}$ with the smallest *Euclidean distance* between any two distinct points being greater than $\epsilon > 0$. See Definition 3.3 in Pollard (1990) for more details.

exists for every $t_1, t_2 \in \mathbb{T}$, and for all deterministic sequences $\{t_{1n}\}$ and $\{t_{2n}\}$ in $\mathbb{T}$, if $\rho(t_{1n}, t_{2n}) \to 0$, then $\rho_n(t_{1n}, t_{2n}) \to 0$ as $n \to \infty$.

Then, provided we have the sufficient measurability, we have

(i) $\mathbb{T}$ is totally bounded under the ρ pseudometric;

(ii) $B_n(t)$ converges weakly to a tight zero-mean Gaussian process $G \in \ell^\infty(\mathbb{T})$ concentrated on $UC_\rho(\mathbb{T})$, with covariance function $\lim_{n\to\infty} E[B_n(t_1)B_n(t_2)]$, where $UC_\rho(\mathbb{T}) := \{z \in \mathbb{T} : z \text{ is uniformly } \rho\text{-continuous}\}$.

Lemma S.2.2 (Wooldridge, 2010, Corollary 3.1)**.** *Let $\{\mathbf{Z}_n : n = 1, 2, \ldots\}$ be a sequence of random $p \times p$ matrices, and let $\mathbf{A}$ be a nonrandom, invertible $p \times p$ matrix. If $\mathbf{Z}_n \xrightarrow{P} \mathbf{A}$, then*

(i) $\mathbf{Z}_n^{-1}$ exists with probability approaching one;

(ii) $\mathbf{Z}_n^{-1} \xrightarrow{P} \mathbf{A}^{-1}$.

Lemma S.2.3 (Chiang et al., 2019, Lemma 2)**.** *Let $(\Omega^w \times \Omega^\xi, \mathcal{S}^w \times \mathcal{S}^\xi, P := P^{w\times\xi})$ be the product probability space of $(\Omega^w, \mathcal{S}^w, P^w)$ and $(\Omega^\xi, \mathcal{S}^\xi, P^\xi)$. For a metric space $(\mathbb{T}, d)$, consider the random sequences X_n, $Y_n : \Omega^w \times \Omega^\xi \to \mathbb{T}$, if $X_n \underset{\xi}{\overset{P}{\rightsquigarrow}} X$ for some random element $X : \Omega^w \times \Omega^\xi \to \mathbb{T}$ and $d(X_n, Y_n) \xrightarrow{P} 0$ as $n \to \infty$, then $Y_n \underset{\xi}{\overset{P}{\rightsquigarrow}} X$.*

Lemma S.2.4 (Gin and Nickl, 2015, Proposition 3.6.12)**.** *Let f be a function of bounded p-variation, $p \geqslant 1$. Then the collection $\mathscr{F} := \{x \mapsto f(tx - s) : t > 0, s \in \mathbb{R}\}$ of translations and dilations of f is of VC type.*

Lemma S.2.5 (Chiang et al., 2019, Lemma 6)**.** *Let $\mathscr{F}$ and $\mathscr{G}$ be of VC type with envelopes F and G, respectively. Then the collections of element-wise sums $\mathscr{F} + \mathscr{G}$ and element-wise products $\mathscr{F} \cdot \mathscr{G}$ are of VC type with envelope $F + G$ and $F \cdot G$, respectively.*

Lemma S.2.6 (Chernozhukov et al., 2014, Theorem 5.2)**.** *Suppose that $F \in \mathcal{L}^2(P)$. Let $\delta := \sigma/\|F\|_{P,2}$ and $M := \max_{1\leqslant i\leqslant n} F(X_i)$, where $\sigma^2 > 0$ is a constant such that $\sup_{f\in\mathcal{F}} E_P[f^2] \leqslant \sigma^2 \leqslant \|F\|_{P,2}^2$. Then,*

$$E\left[\|\mathbb{G}_n\|_{\mathcal{F}}\right] \lesssim J(\delta, \mathcal{F}, F)\,\|F\|_{P,2} + \frac{\|M\|_2 J^2(\delta, \mathcal{F}, F)}{\delta^2\sqrt{n}}.$$

S.2.1 Proof of Lemma 4

Proof. The proof is analogous to those for Lemma 1 and Theorem 1 in Chiang et al. (2019).

(i) Let $\alpha(\theta) = \left(m(\theta, x_0), \frac{m^{(1)}(\theta, x_0^+)}{1!}, \frac{m^{(1)}(\theta, x_0^-)}{1!}, \ldots, \frac{m^{(p)}(\theta, x_0^+)}{p!}, \frac{m^{(p)}(\theta, x_0^-)}{p!}\right)^\top$. The local pth-order polynomial constrained regression estimator of $\alpha(\theta)$ is given by

$$\hat{\alpha}(\theta) = \arg\min_{\alpha\in\mathbb{R}^{2p+1}} \sum_{i=1}^n \left(\varphi(Y_i, \theta) - \bar{r}_p(X_i - x_0)^\top \alpha\right)^2 K\left(\frac{X_i - x_0}{h_{n,\theta}}\right).$$

The first-order condition is

$$\begin{aligned}&\bar{H}_{n,\theta}^{-1}\sum_{i=1}^{n}\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)^{\top}K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\bar{H}_{n,\theta}^{-1}\hat{\alpha}(\theta)\\&\quad=\bar{H}_{n,\theta}^{-1}\sum_{i=1}^{n}\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\varphi(Y_i,\theta)K\left(\frac{X_i-x_0}{h_{n,\theta}}\right),\end{aligned}\tag{S.7}$$

where $\bar{H}_{n,\theta}:=\mathrm{diag}(1,1/h_{n,\theta},1/h_{n,\theta},\dots,1/h_{n,\theta}^p,1/h_{n,\theta}^p)$ is a diagonal matrix. Under Assumption 3(ii)(a), the mean value expansion of $\varphi(y,\theta)$ at $x=x_0$ yields that

$$\begin{aligned}\varphi(y,\theta)&=m(x_0,\theta)+\frac{x-x_0}{h_{n,\theta}}\frac{m^{(1)}(\theta,x_0^+)h_{n,\theta}}{1!}\delta_{x-x_0}^{+}+\frac{x-x_0}{h_{n,\theta}}\frac{m^{(1)}(\theta,x_0^-)h_{n,\theta}}{1!}\delta_{x-x_0}^{-}+\dots\\&\quad+\left(\frac{x-x_0}{h_{n,\theta}}\right)^p\frac{m^{(p)}(\theta,x_0^+)h_{n,\theta}^p}{p!}\delta_{x-x_0}^{+}+\left(\frac{x-x_0}{h_{n,\theta}}\right)^p\frac{m^{(p)}(\theta,x_0^-)h_{n,\theta}^p}{p!}\delta_{x-x_0}^{-}\\&\quad+\left(\frac{x-x_0}{h_{n,\theta}}\right)^{p+1}\frac{m^{(p+1)}(\theta,\bar{x}^+)h_{n,\theta}^{p+1}}{(p+1)!}\delta_{x-x_0}^{+}+\left(\frac{x-x_0}{h_{n,\theta}}\right)^{p+1}\frac{m^{(p+1)}(\theta,\bar{x}^-)h_{n,\theta}^{p+1}}{(p+1)!}\delta_{x-x_0}^{-}\\&\quad+\varepsilon^m(y,x,\theta)\\&=\bar{r}_p\left(\frac{x-x_0}{h_{n,\theta}}\right)^{\top}\bar{H}_{n,\theta}^{-1}\alpha(\theta)+\frac{h_{n,\theta}^{p+1}}{(p+1)!}\left(\frac{x-x_0}{h_{n,\theta}}\right)^{p+1}\left(m^{(p+1)}(\theta,\bar{x}^+)\delta_{x-x_0}^{+}+m^{(p+1)}(\theta,\bar{x}^-)\delta_{x-x_0}^{-}\right)\\&\quad+\varepsilon^m(y,x,\theta)\end{aligned}$$

for some $\bar{x}^+\in(x_0,x)$ and $\bar{x}^-\in(x,x_0)$, where $\delta_u^+:=\mathbb{1}(u\geqslant 0)$, $\delta_u^-:=\mathbb{1}(u<0)$, and $\bar{r}_p(u):=(1,u\delta_u^+,u\delta_u^-,\dots,u^p\delta_u^+,u^p\delta_u^-)^{\top}$. Then, substituting the above expansion into the right-hand side of Equation (S.7) yields:

$$\begin{aligned}&\bar{\Gamma}_{n,p,\theta}\left(\bar{H}_{n,\theta}^{-1}\hat{\alpha}(\theta)-\bar{H}_{n,\theta}^{-1}\alpha(\theta)\right)\\&=\frac{h_{n,\theta}^{p+1}}{nh_{n,\theta}}\sum_{i=1}^{n}\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\frac{\left(\frac{X_i-x_0}{h_{n,\theta}}\right)^{p+1}}{(p+1)!}K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\left(m^{(p+1)}(\theta,\bar{X}_i^+)\delta_i^++m^{(p+1)}(\theta,\bar{X}_i^-)\delta_i^-\right)\\&\quad+\frac{1}{nh_{n,\theta}}\sum_{i=1}^{n}\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\varepsilon(Y_i,X_i,\theta),\end{aligned}\tag{S.8}$$

where $\delta_i^{\pm}:=\delta_{X_i-x_0}^{\pm}$, and

$$\bar{\Gamma}_{n,p,\theta}:=\frac{1}{nh_{n,\theta}}\sum_{i=1}^{n}\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)^{\top}K\left(\frac{X_i-x_0}{h_{n,\theta}}\right).$$

Step 1. We show that $\bar{\Gamma}_{n,p,\theta}=\bar{\Gamma}_pf_X(x_0)+o_P(1)$ uniformly in $\theta\in\bar{\Theta}$, where $\bar{\Gamma}_p:=\int_{\mathbb{R}}\bar{r}_p(u)\bar{r}_p(u)^{\top}K(u)\,du$.

For every $\eta > 0$, it follows that

$$
\begin{aligned}
& P\left[\sup_{\theta\in\bar{\Theta}}\left|\bar{\Gamma}_{n,p,\theta} - \bar{\Gamma}_p f_X(x_0)\right| > \eta\right] \\
& \overset{(1)}{\leqslant} \frac{1}{\eta} E\left[\sup_{\theta\in\bar{\Theta}}\left|\bar{\Gamma}_{n,p,\theta} - \bar{\Gamma}_p f_X(x_0)\right|\right] \\
& \overset{(2)}{\leqslant} \frac{1}{\eta}\left(E\left[\sup_{\theta\in\bar{\Theta}}\left|\bar{\Gamma}_{n,p,\theta} - E\left[\bar{\Gamma}_{n,p,\theta}\right]\right|\right] + E\left[\sup_{\theta\in\bar{\Theta}}\left|E\left[\bar{\Gamma}_{n,p,\theta}\right] - \bar{\Gamma}_p f_X(x_0)\right|\right]\right),
\end{aligned}
$$

where inequality (1) follows from Markov's inequality; inequality (2) follows from Minkowski's inequality. For the deterministic term, we note that

$$
\begin{aligned}
E\left[\bar{\Gamma}_{n,p,\theta}\right] & \overset{(3)}{=} \frac{1}{h_{n,\theta}}\int \bar{r}_p\left(\frac{x-x_0}{h_{n,\theta}}\right)\bar{r}_p\left(\frac{x-x_0}{h_{n,\theta}}\right)^{\top} K\left(\frac{x-x_0}{h_{n,\theta}}\right) f_X(x)\,dx \\
& \overset{(4)}{=} \int \bar{r}_p(u)\bar{r}_p(u)^{\top}K(u)\,du f_X(x_0) + O(h_n),
\end{aligned}
$$

where equality (3) follows from Assumption 3(i)(a); equality (4) follows from the change of variable $u = (x-x_0)/h_{n,\theta}$ and Assumption 3(i)(b). Then, it follows that

$$
E\left[\sup_{\theta\in\bar{\Theta}}\left|E\left[\bar{\Gamma}_{n,p,\theta}\right] - \bar{\Gamma}_p f_X(x_0)\right|\right] = O(h_n).
$$

For the stochastic term, we note that

$$
\bar{r}_p(u)\bar{r}_p(u)^{\top} = \begin{bmatrix}
1 & u\delta_u^+ & u\delta_u^- & \dots & u^p\delta_u^+ & u^p\delta_u^- \\
u\delta_u^+ & u^2\delta_u^+ & 0 & \dots & u^{p+1}\delta_u^+ & 0 \\
u\delta_u^- & 0 & u^2\delta_u^- & \dots & 0 & u^{p+1}\delta_u^- \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
u^p\delta_u^+ & u^{p+1}\delta_u^+ & 0 & \dots & u^{2p}\delta_u^+ & 0 \\
u^p\delta_u^- & 0 & u^{p+1}\delta_u^- & \dots & 0 & u^{2p}\delta_u^-
\end{bmatrix}.
$$

Hence, it suffices to bound the non-zero component of the stochastic term:

$$
E\left[\sup_{\theta\in\bar{\Theta}}\left|\frac{1}{nh_{n,\theta}}\sum_{i=1}^{n}\left(f_n^{\pm}(X_i,\theta) - E\left[f_n^{\pm}(X_i,\theta)\right]\right)\right|\right], \tag{S.9}
$$

where $f_n^{\pm}(x,\theta) := \delta_{x-x_0}^{\pm} K\left(\frac{x-x_0}{h_{n,\theta}}\right)\left(\frac{x-x_0}{h_{n,\theta}}\right)^s$ for any positive integer $s \in \{1,\dots,2p\}$. We only prove the case "$x \geqslant x_0$"; similar arguments apply to the case "$x < x_0$". Consider the following function classes:

$$
\begin{aligned}
& \mathscr{F}_n^+ := \{x \mapsto f_n(x,\theta) : \theta \in \bar{\Theta}\}, \\
\text{and} \quad & \mathscr{F}^+ := \left\{x \mapsto \mathbb{1}(x \geqslant x_0)K(a(x-x_0))(a(x-x_0))^s\mathbb{1}(a(x-x_0) \in [-1,1]) : a > 1/h_0\right\}.
\end{aligned}
$$

By Lemma S.2.5 and under Assumption 3(iii)(a), the function class $\mathscr{F}^+$ is of VC type with envelope $F(x) = \sup_{a\in(1/h_0,\infty)} |K(a(x - x_0))|$. This VC property ensures that the uniform entropy integral is finite, i.e., $J(1, \mathscr{F}^+, F) < \infty$. Then, we obtain that

$$
\begin{aligned}
\text{(S.9)} &\overset{(5)}{\lesssim} \frac{1}{\sqrt{n}h_n} E\left[\sup_{f\in\mathscr{F}_n^+} |\mathbb{G}_n f|\right] \\
&\overset{(6)}{\leqslant} \frac{1}{\sqrt{n}h_n} E\left[\sup_{f\in\mathscr{F}^+} |\mathbb{G}_n f|\right] \\
&\overset{(7)}{\lesssim} \frac{1}{\sqrt{n}h_n} \left(J(1, \mathscr{F}^+, F)\|F\|_{P,2} + \frac{\|\max_{1\leqslant i\leqslant n} F(X_i)\|_2\, J^2(1, \mathscr{F}^+, F)}{\sqrt{n}} \right) \\
&= O\left(\frac{1}{\sqrt{n}h_n}\right)
\end{aligned}
$$

as $n \to \infty$, where inequality (5) follows from Assumption 3(iv); inequality (6) follows from $\mathscr{F}_n^+ \subset \mathscr{F}^+$; inequality (7) follows from Lemma S.2.6 with $\delta = \sigma/\|F\|_{P,2} \leqslant 1$. Therefore,

$$
P\left[\sup_{\theta\in\bar{\Theta}} \left|\bar{\Gamma}_{n,p,\theta} - \bar{\Gamma}_p f_X(x_0)\right| > \eta\right] \leqslant \frac{1}{\eta}\left(O\left(\frac{1}{\sqrt{n}h_n}\right) + O(h_n) \right) \to 0
$$

as $n \to \infty$. This proves $\bar{\Gamma}_{n,p,\theta} \xrightarrow{P} \bar{\Gamma}_p f_X(x_0)$ uniformly in $\theta \in \bar{\Theta}$. |||

By Lemma S.2.2, $\bar{\Gamma}_{n,p,\theta}^{-1}$ exists with probability approaching one, and $\bar{\Gamma}_{n,p,\theta}^{-1} = \bar{\Gamma}_p^{-1}/f_X(x_0) + o_P(1)$. Note that

$$
\begin{aligned}
&\bar{H}_{n,\theta}^{-1}\hat{\alpha}(\theta) - \bar{H}_{n,\theta}^{-1}\alpha(\theta) \\
&= \Bigg(\hat{m}(\theta, x_0) - m(\theta, x_0), \frac{(\hat{m}^{(1)}(\theta, x_0^+) - m^{(1)}(\theta, x_0^+))h_{n,\theta}}{1!}, \frac{(\hat{m}^{(1)}(\theta, x_0^-) - m^{(1)}(\theta, x_0^-))h_{n,\theta}}{p!}, \\
&\qquad \ldots, \frac{(\hat{m}^{(p)}(\theta, x_0^+) - m^{(p)}(\theta, x_0^+))h_{n,\theta}^p}{1!}, \frac{(\hat{m}^{(p)}(\theta, x_0^-) - m^{(p)}(\theta, x_0^-))h_{n,\theta}^p}{p!} \Bigg)^\top.
\end{aligned}
$$

Then, multiplying both sides of Equation (S.8) by $\sqrt{nh_{n,\theta}}\iota_2^\top \bar{\Gamma}_{n,p,\theta}^{-1}$ yields

$$
\sqrt{nh_{n,\theta}^3}\left(\hat{m}^{(1)}(\theta, x_0^+) - m^{(1)}(\theta, x_0^+)\right) = \frac{\iota_2^\top \bar{\Gamma}_{n,p,\theta}^{-1}}{(p+1)!}\bar{A}_{n,\theta} + \iota_2^\top \bar{\Gamma}_{n,p,\theta}^{-1}\bar{B}_{n,\theta}, \tag{S.10}
$$

where

$$
\begin{aligned}
\bar{A}_{n,\theta} &:= \frac{\sqrt{h_{n,\theta}^{2p+1}}}{n} \sum_{i=1}^n \bar{r}_p\left(\frac{X_i - x_0}{h_{n,\theta}}\right)\left(\frac{X_i - x_0}{h_{n,\theta}}\right)^{p+1} K\left(\frac{X_i - x_0}{h_{n,\theta}}\right) \\
&\qquad \cdot \left(m^{(p+1)}(\theta, \bar{X}_i^+)\delta_i^+ + m^{(p+1)}(\theta, \bar{X}_i^-)\delta_i^- \right), \\
\bar{B}_{n,\theta} &:= \frac{1}{\sqrt{nh_{n,\theta}}} \sum_{i=1}^n \bar{r}_p\left(\frac{X_i - x_0}{h_{n,\theta}}\right) K\left(\frac{X_i - x_0}{h_{n,\theta}}\right) \varepsilon(Y_i, X_i, \theta).
\end{aligned}
$$

By similar reasoning, $\sqrt{nh_{n,\theta}^3}\left(\hat{m}^{(1)}(\theta,x_0^-)-m^{(1)}(\theta,x_0^-)\right)=\frac{\iota_3^\top\bar{\Gamma}_{n,p,\theta}^{-1}}{(p+1)!}\bar{A}_{n,\theta}+\iota_3^\top\bar{\Gamma}_{n,p,\theta}^{-1}\bar{B}_{n,\theta}$.

To derive the uniform Bahadur representation, we now analyze the consistency of $\bar{A}_{n,\theta}$ and the boundedness of $\bar{B}_{n,\theta}$.

Step 2. We show that $\bar{A}_{n,\theta}=\bar{A}_{n,\theta}^*+o_P\left(\sqrt{nh_{n,\theta}^{2p+3}}\right)$ uniformly in $\theta\in\bar{\Theta}$, where

$$\bar{A}_{n,\theta}^*:=\sqrt{nh_{n,\theta}^{2p+3}}\left(m^{(p+1)}(\theta,x_0^+)\bar{\vartheta}_{p,p+1}^+ + m^{(p+1)}(\theta,x_0^-)\bar{\vartheta}_{p,p+1}^-\right)f_X(x_0).$$

For every $\eta>0$,

$$\begin{aligned}P\left[\sup_{\theta\in\bar{\Theta}}\left|\bar{A}_{n,\theta}-\bar{A}_{n,\theta}^*\right|\right]&\overset{(8)}{\leqslant}\frac{1}{\eta}E\left[\sup_{\theta\in\bar{\Theta}}\left|\bar{A}_{n,\theta}-\bar{A}_{n,\theta}^*\right|\right]\\&\overset{(9)}{\leqslant}\frac{1}{\eta}\left(E\left[\sup_{\theta\in\bar{\Theta}}\left|\bar{A}_{n,\theta}-E[\bar{A}_{n,\theta}]\right|\right]+E\left[\sup_{\theta\in\bar{\Theta}}\left|E[\bar{A}_{n,\theta}]-\bar{A}_{n,\theta}^*\right|\right]\right),\end{aligned}$$

where inequality (8) follows from Markov's inequality; inequality (9) follows from Minkowski's inequality. For the deterministic term, note that

$$\begin{aligned}E\left[\bar{A}_{n,\theta}\right]&=\sqrt{nh_{n,\theta}^{2p+3}}\left(E\left[\frac{1}{nh_{n,\theta}}\sum_{i=1}^n\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\left(\frac{X_i-x_0}{h_{n,\theta}}\right)^{p+1}K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)m^{(p+1)}(\theta,\bar{X}_i^+)\delta_i^+\right]\right.\\&\quad\left.+E\left[\frac{1}{nh_{n,\theta}}\sum_{i=1}^n\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\left(\frac{X_i-x_0}{h_{n,\theta}}\right)^{p+1}K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)m^{(p+1)}(\theta,\bar{X}_i^-)\delta_i^-\right]\right)\\&\overset{(10)}{=}\sqrt{nh_{n,\theta}^{2p+3}}\left(\int_{\mathbb{R}_+}\bar{r}_p(u)u^{p+1}K(u)\,du\cdot m^{(p+1)}(\theta,x_0^+)f_X(x_0)\right.\\&\quad\left.+\int_{\mathbb{R}_-}\bar{r}_p(u)u^{p+1}K(u)\,du\cdot m^{(p+1)}(\theta,x_0^-)f_X(x_0)+O(h_n)\right)\\&=\underbrace{\sqrt{nh_{n,\theta}^{2p+3}}\left(m^{(p+1)}(\theta,x_0^+)\bar{\vartheta}_{p,p+1}^+ + m^{(p+1)}(\theta,x_0^-)\bar{\vartheta}_{p,p+1}^-\right)f_X(x_0)}_{=:\bar{A}_{n,\theta}^*}+O\left(\sqrt{nh_n^{2p+5}}\right),\end{aligned}$$

where equality (10) follows from Assumptions 3(ii) and (iii)(a). Hence, $E\left[\sup_{\theta\in\bar{\Theta}}\left|E[\bar{A}_{n,\theta}]-\bar{A}_{n,\theta}^*\right|\right]=O\left(\sqrt{nh_n^{2p+5}}\right)$ uniformly in $\theta\in\bar{\Theta}$.

For the stochastic term, recalling that $\bar{r}_p(u)=(1,u\delta_u^+,u\delta_u^-,\dots,u^p\delta_u^+,u^p\delta_u^-)^\top$, it suffices to bound the relevant component of $\bar{B}_{n,\theta}$:

$$E\left[\sup_{\theta\in\bar{\Theta}}\left|\sqrt{\frac{h_{n,\theta}^{2p+1}}{n}}\sum_{i=1}^n\left(g_n(X_i,\theta)-E[g_n(X_i,\theta)]\right)\right|\right]\tag{S.11}$$

for every positive integer $s \in \{0, \dots, p\}$, where

$$g_n(x,\theta) = \left(\frac{x-x_0}{h_{n,\theta}}\right)^{s+p+1} K\left(\frac{x-x_0}{h_{n,\theta}}\right)\left(m^{(p+1)}(\theta, x_0^+)\delta_{x-x_0}^+ + m^{(p+1)}(\theta, x_0^-)\delta_{x-x_0}^-\right).$$

Consider the function classes

$$\begin{aligned}
\mathscr{G}_n &:= \{x \mapsto g_n(x,\theta) : \theta \in \bar{\Theta}\}, \\
\text{and } \mathscr{G} &:= \big\{x \mapsto (a(x-x_0))^{s+p+1} K(a(x-x_0)) \mathbb{1}(a(x-x_0) \in [-1,1]) \\
&\qquad \cdot (m^{(p+1)}(\theta, x_0^+)\delta_{x-x_0}^+ + m^{(p+1)}(\theta, x_0^-)\delta_{x-x_0}^-) : a \geqslant 1/h_0, \theta \in \bar{\Theta}\big\}.
\end{aligned}$$

Since $x \mapsto (a(x-x_0))^{s+p+1}\mathbb{1}(a(x-x_0) \in [-1,1])$ is of bounded $(s+p+1)$-variation, by Lemma S.2.4, $\{x \mapsto (a(x-x_0))^{s+p+1}\mathbb{1}(a(x-x_0) \in [-1,1]) : a \geqslant 1/h_0\}$ is of VC type. Furthermore, under Assumption 3(ii)(a), (ii)(b), and (iii)(a), $m^{(p+1)}(\theta, \cdot)$ is Lipschitz on $I_{x_0}\backslash\{x_0\}$ and $\{x \mapsto K((x-x_0)/h) : h > 0\}$ is of VC type. Then, Lemma S.2.5 implies that the composite class $\mathscr{G}$ is of VC type with envelope $G(x) := \sup_{a \in (1/h_0, \infty)} |K(a(x-x_0))| \cdot \int_{\mathcal{Y}\times\mathcal{X}} \chi(y,x)\, dF_{Y,X}(y,x)$, and its uniform entropy integral is finite: $J(1, \mathscr{G}, G) < \infty$. Hence,

$$\begin{aligned}
\text{(S.11)} &\overset{(11)}{\lesssim} \sqrt{h_n^{2p+1}} E\left[\sup_{g \in \mathscr{G}_n} |\mathbb{G}_n g|\right] \\
&\overset{(12)}{\leqslant} \sqrt{h_n^{2p+1}} E\left[\sup_{g \in \mathscr{G}} |\mathbb{G}_n g|\right] \\
&\overset{(13)}{\lesssim} \sqrt{h_n^{2p+1}} \left(J(1, \mathscr{G}, G)\|G\|_{P,2} + \frac{\|\max_{1 \leqslant i \leqslant n} G(X_i)\|_2 J^2(1, \mathscr{G}, G)}{\sqrt{n}}\right) \\
&= O\left(\sqrt{h_n^{2p+1}}\right) + O\left(\sqrt{\frac{h_n^{2p+1}}{n}}\right),
\end{aligned}$$

where inequality (11) follows from Assumption 3(iv); inequality (12) follows from $\mathscr{G}_n \subset \mathscr{G}$ for all $n \in \mathbb{N}_+$; inequality (13) follows from Lemma S.2.6. Therefore,

$$\begin{aligned}
P\left[\sup_{\theta \in \bar{\Theta}} |\bar{A}_{n,\theta} - \bar{A}_{n,\theta}^*|\right] &\leqslant \frac{1}{\eta}\left(O\left(\sqrt{h_n^{2p+1}}\right) + O\left(\sqrt{\frac{h_n^{2p+1}}{n}}\right) + O\left(\sqrt{n h_n^{2p+5}}\right)\right) \\
&\to 0
\end{aligned}$$

as $n \to \infty$ under Assumption 3(iv). This proves $\bar{A}_{n,\theta} = \bar{A}_{n,\theta}^* + o_P(1)$ uniformly in $\theta \in \bar{\Theta}$. |||

Step 3. We show that $\bar{B}_{n,\theta}$ is bounded in probability.

To this end, we consider each component $\iota_s^\top \bar{B}_{n,\theta}$ for $s \in \{1, \dots, 2p+1\}$. For every $\eta > 0$, there exists $M_\eta > 0$ and $N \in \mathbb{N}_+$ such that for all $n \geqslant N$,

$$
\begin{aligned}
& P\left[\sup_{\theta\in\bar\Theta} \left|\iota_s^\top \bar B_{n,\theta}\right| \geqslant M_\eta\right] \\
\overset{(14)}{\leqslant}\; & \frac{1}{M_\eta^2} E\left[\sup_{\theta\in\bar\Theta}\left|\frac{\iota_s^\top}{\sqrt{nh_{n,\theta}}}\sum_{i=1}^n \bar r_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right) K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\varepsilon^m(Y_i,X_i,\theta)\right|^2\right] \\
\overset{(15)}{\leqslant}\; & \frac{C}{M_\eta^2} E\left[\frac{\iota_s^\top}{nh_{n,\bar\theta}}\sum_{i=1}^n \bar r_p\left(\frac{X_i-x_0}{h_{n,\bar\theta}}\right)\bar r_p\left(\frac{X_i-x_0}{h_{n,\bar\theta}}\right)^\top K^2\left(\frac{X_i-x_0}{h_{n,\bar\theta}}\right)(\varepsilon^m(Y_i,X_i,\bar\theta))^2\iota_s\right] \\
\overset{(16)}{=}\; & \frac{C}{M_\eta^2}\iota_s^\top\Bigg(\int_{x_0}^\infty \frac{1}{h_{n,\bar\theta}}\bar r_p\left(\frac{x-x_0}{h_{n,\bar\theta}}\right)\bar r_p\left(\frac{x-x_0}{h_{n,\bar\theta}}\right)^\top K^2\left(\frac{x-x_0}{h_{n,\bar\theta}}\right)\sigma^m(\bar\theta,\bar\theta|x)f_X(x)\,dx \\
& + \int_{-\infty}^{x_0}\frac{1}{h_{n,\bar\theta}}\bar r_p\left(\frac{x-x_0}{h_{n,\bar\theta}}\right)\bar r_p\left(\frac{x-x_0}{h_{n,\bar\theta}}\right)^\top K^2\left(\frac{x-x_0}{h_{n,\bar\theta}}\right)\sigma^m(\bar\theta,\bar\theta|x)f_X(x)\,dx\Bigg)\iota_s \\
\overset{(17)}{=}\; & \frac{C}{M_\eta^2}\iota_s^\top\Bigg(\int_{\mathbb{R}_+}\bar r_p(u)\bar r_p(u)^\top K^2(u)\,du\cdot\sigma^m(\bar\theta,\bar\theta|x_0^+) \\
& + \int_{\mathbb{R}_-}\bar r_p(u)\bar r_p(u)^\top K^2(u)\,du\cdot\sigma^m(\bar\theta,\bar\theta|x_0^-)\Bigg)\iota_s f_X(x_0) + O(h_n) \\
\leqslant\; & \frac{C\left[\iota_s^\top\left(\bar\Psi_p^+\sigma^m(\bar\theta,\bar\theta|x_0^+) + \bar\Psi_p^-\sigma^m(\bar\theta,\bar\theta|x_0^-)\right)\iota_s f_X(x_0) + h_n\right]}{M_\eta^2}
\end{aligned}
$$

where inequality (14) follows from 2nd-order Markov's inequality; inequality (15) holds for some interior point $\bar\theta \in \bar\Theta$ and a large constant $C > 0$; equality (16) follows from the law of iterated expectations; equality (17) follows from the change of variable $u = (x-x_0)/h_{n,\bar\theta}$ and Assumption 3(ii)(c). Therefore, for every $\eta > 0$, the right-hand side can be made arbitrarily small by choosing a sufficiently large $M_\eta > 0$. This bounds $\bar B_{n,\theta}$ in probability uniformly in $\theta \in \bar\Theta$. |||

Combining the results from Steps 1–3 with Equation (S.10) yields:

$$
\begin{aligned}
& \sqrt{nh_{n,\theta}^3}\left(\hat m^{(1)}(\theta,x_0^+) - m^{(1)}(\theta,x_0^+)\right) \\
={} & \left\{\frac{\iota_2^\top\bar\Gamma_p^{-1}/f_X(x_0)}{(p+1)!} + O_P(h_n)\right\} \\
& \cdot\left\{\sqrt{nh_{n,\theta}^{2p+3}}\left(m^{(p+1)}(\theta,x_0^+)\bar\vartheta_{p,p+1}^+ + m^{(p+1)}(\theta,x_0^-)\bar\vartheta_{p,p+1}^-\right)f_X(x_0) + o_P\left(\sqrt{nh_n^{2p+3}}\right)\right\} \\
& + \left\{\frac{\iota_2^\top\bar\Gamma_p^{-1}}{f_X(x_0)} + O_P(h_n)\right\}\frac{1}{\sqrt{nh_{n,\theta}}}\sum_{i=1}^n\bar r_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\varepsilon(Y_i,X_i,\theta) \\
={} & \sqrt{nh_{n,\theta}^{2p+3}}\frac{\iota_2^\top\bar\Gamma_p^{-1}}{(p+1)!}\left(m^{(p+1)}(\theta,x_0^+)\bar\vartheta_{p,p+1}^+ + m^{(p+1)}(\theta,x_0^-)\bar\vartheta_{p,p+1}^-\right) \\
& + \frac{\iota_2^\top\bar\Gamma_p^{-1}}{f_X(x_0)\sqrt{nh_{n,\theta}}}\sum_{i=1}^n\bar r_p\left(\frac{X_i-x_0}{h_{n,\theta}}\right)K\left(\frac{X_i-x_0}{h_{n,\theta}}\right)\varepsilon(Y_i,X_i,\theta) + o_P\left(\sqrt{nh_n^{2p+3}}\right) + O_P(h_n).
\end{aligned}
$$

By similar reasoning for the case $x < x_0$, we obtain:

$$\begin{aligned}
&\sqrt{nh_{n,\theta}^3}\left(\hat{m}^{(1)}(\theta,x_0^-) - m^{(1)}(\theta,x_0^-)\right) \\
&= \sqrt{nh_{n,\theta}^{2p+3}}\frac{\iota_3^\top \bar{\Gamma}_p^{-1}}{(p+1)!}\left(m^{(p+1)}(\theta,x_0^+)\bar{\vartheta}_{p,p+1}^+ + m^{(p+1)}(\theta,x_0^-)\bar{\vartheta}_{p,p+1}^-\right) \\
&\quad + \frac{\iota_3^\top \bar{\Gamma}_p^{-1}}{f_X(x_0)\sqrt{nh_{n,\theta}}}\sum_{i=1}^n \bar{r}_p\left(\frac{X_i - x_0}{h_{n,\theta}}\right)K\left(\frac{X_i - x_0}{h_{n,\theta}}\right)\varepsilon(Y_i,X_i,\theta) + o_P\left(\sqrt{nh_n^{2p+3}}\right) + O_P(h_n).
\end{aligned}$$

Rearranging terms then yields the desired uniform Bahadur representation, which completes the proof of Part (i).

(ii) We first define

$$G_{n,1}(\theta,k) := \sum_{i=1}^n f_{ni}(\theta,k) := \sum_{i=1}^n f_n(Y_i,X_i,\theta,k) \tag{S.12}$$

for every $(\theta,k) \in \bar{\Theta} \times \{2,3\}$, where

$$\begin{aligned}
f_n(y,x,\theta,k) &= \frac{\iota_k^\top \bar{\Gamma}_p^{-1}\bar{r}_p\left(\frac{x-x_0}{h_{n,\theta}}\right)K\left(\frac{x-x_0}{h_{n,\theta}}\right)\varepsilon^m(y,x,\theta)}{f_X(x_0)\sqrt{nh_{n,\theta}}} \\
&= \frac{\left\{c_{k,0} + c_{k,1}\left(\frac{x-x_0}{h_{n,\theta}}\right) + \ldots + c_{k,2p+1}\left(\frac{x-x_0}{h_{n,\theta}}\right)^p\right\}K\left(\frac{x-x_0}{h_{n,\theta}}\right)\varepsilon^m(y,x,\theta)}{f_X(x_0)\sqrt{nh_{n,\theta}}},
\end{aligned}$$

with constants $\{c_{k,0},\ldots,c_{k,2p+1}\}$ determined by the term $\iota_k^\top \bar{\Gamma}_p^{-1}\bar{r}_p\left(\frac{x-x_0}{h_{n,\theta}}\right)$. The weak convergence result follows from an application of Pollard's functional central limit theorem (Lemma S.2.1), provided Conditions (A)–(E) of that lemma are satisfied.

For Condition (A), we note that: (1) $x \mapsto \frac{\left(c_{k,0}+c_{k,1}\left(\frac{x-x_0}{h_{n,\theta}}\right)+\ldots+c_{k,2p+1}\left(\frac{x-x_0}{h_{n,\theta}}\right)^p\right)\mathbb{1}\left(\left|\frac{x-x_0}{h_{n,\theta}}\right|\leqslant 1\right)}{f_X(x_0)\sqrt{nh_{n,\theta}}}$ is of bounded variation, and thus $\left\{x \mapsto \frac{\left(c_{k,0}+c_{k,1}\left(\frac{x-x_0}{h_{n,\theta}}\right)+\ldots+c_{k,2p+1}\left(\frac{x-x_0}{h_{n,\theta}}\right)^p\right)\mathbb{1}\left(\left|\frac{x-x_0}{h_{n,\theta}}\right|\leqslant 1\right)}{f_X(x_0)\sqrt{nh_{n,\theta}}} : \theta \in \bar{\Theta}\right\}$ is of VC type for each $k \in \{2,3\}$ by Lemma S.2.4; (2) under Assumption 3(ii)(b), $\{(y,x) \mapsto \varepsilon^m(y,x,\theta) : \theta \in \bar{\Theta}\}$ is of VC type with envelope $2\chi(y,x)$; (3) under Assumption 3(iii)(a), $\{x \mapsto K((x-x_0)/h) : h > 0\}$ is of VC type with envelope $\|K\|_\infty$. Then, using Lemma S.2.5, the product class $\{(y,x) \mapsto f_n(y,x,\theta,k) : \theta \in \bar{\Theta}\}$ is of VC type for each $k \in \{2,3\}$, with envelope

$$F_n(y,x,\theta,k) := \frac{C_k}{\sqrt{nh_{n,\theta}}}\chi(y,x)\mathbb{1}\left(\left|\frac{x-x_0}{h_{n,\theta}}\right| \leqslant 1\right)$$

for some constant $C_k > 0$. We let $\mathscr{F}_{n,k} := \{(f_{n1}(\theta,k),\ldots,f_{nn}(\theta,k)) \in \mathbb{R}^n : \theta \in \bar{\Theta}\}$, $F_{ni,k} := F_n(Y_i,X_i,\theta,k)$, and $\boldsymbol{F}_{n,k} := (F_{n1,k},\ldots,F_{nn,k}) \in \mathbb{R}^n$. Since $\mathscr{F}_{n,k}$ is a VC class and the Pollard's packing number satisfies[2]

$$\sup_{\omega\in\Omega,n\geqslant 1,\alpha\in\mathbb{R}^n} D\left(\epsilon\|\alpha\odot \boldsymbol{F}_{n,k}(\omega)\|,\mathscr{F}_{n,k}\right)\leqslant \sup_P N\left(\mathscr{F}_{n,k},e_{P,2},\epsilon/2\|\alpha\odot \boldsymbol{F}_{n,k}(\omega)\|\right),$$

then there exists a function λ_k independent of ω and n satisfying the conditions in Definition S.2.2 for $k\in\{2,3\}$. This implies the array $\{f_{ni}(\theta,k):\theta\in\bar{\Theta}\}$ is manageable. Condition (A) is thus verified.

For Condition (B), since $\{f_{ni}(\theta,k):\theta\in\bar{\Theta}\}$ is row independent by Assumption 3(i)(a) and $E[f_{ni}(\theta,k)]=0$ for each $i=1,\dots,n$, it follows that:

$$\begin{aligned}
&\lim_{n\to\infty} E\left[G_{n,1}(\theta_1,k_1)G_{n,1}(\theta_2,k_2)\right]\\
&=\lim_{n\to\infty} E\left[\sum_{i=1}^n f_{ni}(\theta_1,k_1)f_{ni}(\theta_2,k_2)\right]\\
&=\frac{\iota_{k_1}^\top\bar{\Gamma}_p^{-1}}{f_X^2(x_0)}E\left[\frac{\sum_{i=1}^n\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta_1}}\right)\bar{r}_p\left(\frac{X_i-x_0}{h_{n,\theta_2}}\right)K\left(\frac{X_i-x_0}{h_{n,\theta_1}}\right)K\left(\frac{X_i-x_0}{h_{n,\theta_2}}\right)\varepsilon^m(Y_i,X_i,\theta_1)\varepsilon^m(Y_i,X_i,\theta_2)}{n\sqrt{h_{n,\theta_1}h_{n,\theta_2}}}\right]\bar{\Gamma}_p^{-1}\iota_{k_2}\\
&\overset{(18)}{=}\frac{\iota_{k_1}^\top\bar{\Gamma}_p^{-1}}{f^2(x_0)\sqrt{\varsigma(\theta_1)\varsigma(\theta_2)}}E\left[\frac{\sum_{i=1}^n\bar{r}_p\left(\frac{X_i-x_0}{\varsigma(\theta_1)h_n}\right)\bar{r}_p\left(\frac{X_i-x_0}{\varsigma(\theta_2)h_n}\right)K\left(\frac{X_i-x_0}{\varsigma(\theta_1)h_n}\right)K\left(\frac{X_i-x_0}{\varsigma(\theta_2)h_n}\right)\sigma^m(\theta_1,\theta_2|X_i)\delta_i^+}{nh_n}\right.\\
&\qquad\left.+\frac{\sum_{i=1}^n\bar{r}_p\left(\frac{X_i-x_0}{\varsigma(\theta_1)h_n}\right)\bar{r}_p\left(\frac{X_i-x_0}{\varsigma(\theta_2)h_n}\right)K\left(\frac{X_i-x_0}{\varsigma(\theta_1)h_n}\right)K\left(\frac{X_i-x_0}{\varsigma(\theta_2)h_n}\right)\sigma^m(\theta_1,\theta_2|X_i)\delta_i^-}{nh_n}\right]\bar{\Gamma}_p^{-1}\iota_{k_2}\\
&\overset{(19)}{=}\frac{\iota_{k_1}^\top\bar{\Gamma}_p^{-1}}{f_X(x_0)}\left(\frac{\int_{\mathbb{R}_+}\bar{r}_p\left(\frac{u}{\varsigma(\theta_1)}\right)\bar{r}_p\left(\frac{u}{\varsigma(\theta_2)}\right)^\top K\left(\frac{u}{\varsigma(\theta_1)}\right)K\left(\frac{u}{\varsigma(\theta_2)}\right)du}{\sqrt{\varsigma(\theta_1)\varsigma(\theta_2)}}\cdot\sigma^m(\theta_1,\theta_2|x_0^+)\right.\\
&\qquad\left.+\frac{\int_{\mathbb{R}_-}\bar{r}_p\left(\frac{u}{\varsigma(\theta_1)}\right)\bar{r}_p\left(\frac{u}{\varsigma(\theta_2)}\right)^\top K\left(\frac{u}{\varsigma(\theta_1)}\right)K\left(\frac{u}{\varsigma(\theta_2)}\right)du}{\sqrt{\varsigma(\theta_1)\varsigma(\theta_2)}}\cdot\sigma^m(\theta_1,\theta_2|x_0^-)\right)\bar{\Gamma}_p^{-1}\iota_{k_2}+O_P(h_n),
\end{aligned}$$

where equality (18) follows from Assumption 3(iv); equality (19) follows from the change of variable $u=(x-x_0)/h_n$. Then, under Assumption 3(iii)(c),

$$\lim_{n\to\infty} E\left[G_{n,1}(\theta_1,k_1)G_{n,1}(\theta_2,k_2)\right]<\infty,$$

This verifies the condition (B).

For Condition (C), we note that

$$\lim_{n\to\infty} E\left[F_{ni}^2(\theta,k)\right]\ =\lim_{n\to\infty} E\left[\frac{C_k^2}{nh_{n,\theta}}\sum_{i=1}^n\chi^2(Y_i,X_i)\mathbb{1}\left(\left|\frac{X_i-x_0}{h_{n,\theta}}\right|\leqslant 1\right)\right]$$

[2]See an analogous inequality in the proof of Theorem 1 in Andrews (1994).

$$
\begin{aligned}
&\overset{(20)}{=} \lim_{n\to\infty} C_k^2 \iint_{-1}^{1} \chi^2(y, x_0 + uh_n\varsigma(\theta)) f_{Y,X}(y, x_0 + uh_n\varsigma(\theta))\, du\, dy \\
&\overset{(21)}{=} C_k^2 \iint_{-1}^{1} \chi^2(y, x_0) f_{Y,X}(y, x_0)\, du\, dy \\
&< \infty,
\end{aligned}
$$

where equality (20) follows from the change of variable $u = (x - x_0)/h_{n,\theta}$; equality (21) follows from the dominated convergence theorem under Assumption 3(ii.b). This verifies the condition (C).

For Condition (D), by a similar argument to that used for Condition (C), it follows that:

$$
\begin{aligned}
&\lim_{n\to\infty} E\left[F_{ni}^2(\theta, k)\mathbb{1}\left(F_{ni}(\theta, k) > \epsilon\right)\right] \\
&= \lim_{n\to\infty} C_k^2 \iint_{-1}^{1} \chi^2(y, x_0 + uh_n\varsigma(\theta))\mathbb{1}\left(\frac{C_k\chi(y, x_0 + uh_n\varsigma(\theta))}{\sqrt{nh_n\varsigma(\theta)}} > \epsilon\right) f_{Y,X}(y, x_0 + uh_n\varsigma(\theta))\, du\, dy \\
&= C_k^2 \iint_{-1}^{1} \lim_{n\to\infty} \chi^2(y, x_0 + uh_n\varsigma(\theta))\mathbb{1}\left(\frac{C_k\chi(y, x_0 + uh_n\varsigma(\theta))}{\sqrt{nh_n\varsigma(\theta)}} > \epsilon\right) f_{Y,X}(y, x_0 + uh_n\varsigma(\theta))\, du\, dy \\
&= 0
\end{aligned}
$$

for each $\epsilon > 0$. Condition (D) is thus verified.

For Condition (E), by the calculation used for Condition (B), it follows that:

$$
\begin{aligned}
&\sum_{i=1}^{n} E\left[f_{ni}(\theta_1, k_1) - f_{ni}(\theta_2, k_2)\right]^2 \\
&= E\left[\sum_{i=1}^{n} f_{ni}^2(\theta_1, k_1)\right] + E\left[\sum_{i=1}^{n} f_{ni}^2(\theta_2, k_2)\right] - 2E\left[\sum_{i=1}^{n} f_{ni}(\theta_1, k_1) f_{ni}(\theta_2, k_2)\right] \\
&= \frac{\bar{\varGamma}_p^{-1}\iota_{k_1}\left(\bar{\varPsi}_p^{+}(\theta_1, \theta_1)\sigma^m(\theta_1, \theta_1|x_0^{+}) + \bar{\varPsi}_p^{-}(\theta_1, \theta_1)\sigma^m(\theta_1, \theta_1|x_0^{-})\right)\bar{\varGamma}_p^{-1}\iota_{k_1}}{f_X(x_0)} \\
&\quad + \frac{\bar{\varGamma}_p^{-1}\iota_{k_2}\left(\bar{\varPsi}_p^{+}(\theta_2, \theta_2)\sigma^m(\theta_2, \theta_2|x_0^{+}) + \bar{\varPsi}_p^{-}(\theta_2, \theta_2)\sigma^m(\theta_2, \theta_2|x_0^{-})\right)\bar{\varGamma}_p^{-1}\iota_{k_2}}{f_X(x_0)} \\
&\quad - 2\frac{\bar{\varGamma}_p^{-1}\iota_{k_1}\left(\bar{\varPsi}_p^{+}(\theta_1, \theta_2)\sigma^m(\theta_1, \theta_2|x_0^{+}) + \bar{\varPsi}_p^{-}(\theta_1, \theta_2)\sigma^m(\theta_1, \theta_2|x_0^{-})\right)\bar{\varGamma}_p^{-1}\iota_{k_2}}{f_X(x_0)} + O(h_n)
\end{aligned}
$$

for every $\theta_1, \theta_2 \in \bar{\varTheta}$, and $k \in \{2, 3\}$, where

$$
\bar{\varPsi}_p^{\pm}(\theta_1, \theta_2) := \frac{1}{\sqrt{\varsigma(\theta_1)\varsigma(\theta_2)}} \int_{\mathbb{R}_{\pm}} \bar{r}_p\left(\frac{u}{\varsigma(\theta_1)}\right)\bar{r}_p\left(\frac{u}{\varsigma(\theta_2)}\right)^{\top} K\left(\frac{u}{\varsigma(\theta_1)}\right) K\left(\frac{u}{\varsigma(\theta_2)}\right) du.
$$

This implies that both

$$
\rho_n((\theta_1, k_1), (\theta_2, k_2)) := \left(\sum_{i=1}^{n} E\left[f_{ni}(\theta_1, k_1) - f_{ni}(\theta_2, k_2)\right]^2\right)^{1/2}
$$

and $\rho((\theta_1,k_1),(\theta_2,k_2)) := \lim_{n\to\infty}\rho_n((\theta_1,k_1),(\theta_2,k_2))$ exist. Furthermore, we note that all terms depending on n in $\rho_n(\cdot,\cdot)$ and $\rho(\cdot,\cdot)$ are included in $O(h_n)$. Thus, for every sequences $\{(\theta_{1n},k_{1n})\}$ and $\{(\theta_{2n},k_{2n})\}$ in $\bar{\Theta}\times\{2,3\}$, if $\rho((\theta_{1n},k_{1n}),(\theta_{2n},k_{2n}))\to 0$, then $\rho_n((\theta_{1n},k_{1n}),(\theta_{2n},k_{2n}))\to 0$ as $n\to\infty$. This verifies Condition (E).

Therefore, Lemma S.2.1 implies $G_{n,1}(\cdot,\cdot)\rightsquigarrow\mathbb{Z}^m(\cdot,\cdot)$ as $n\to\infty$, where $\mathbb{Z}^m:\Omega\to\ell^\infty(\bar{\Theta}\times\{2,3\})$ is a tight zero-mean Gaussian process with covariance function given by:

$$
\begin{aligned}
E\left[\mathbb{Z}^m(\theta_1,k_1)\mathbb{Z}^m(\theta_2,k_2)\right] &:= \lim_{n\to\infty} E\left[G_{n,1}(\theta_1,k_1)G_{n,1}(\theta_2,k_2)\right]\\
&= \frac{\iota_{k_1}^\top \bar{\Gamma}_p^{-1}\left(\bar{\Psi}_p^{+}(\theta_1,\theta_2)\sigma^m(\theta_1,\theta_2|x_0^+)+\bar{\Psi}_p^{-}(\theta_1,\theta_2)\sigma^m(\theta_1,\theta_2|x_0^-)\right)\bar{\Gamma}_p^{-1}\iota_{k_2}}{f_X(x_0)}.
\end{aligned}
$$

This completes the proof. ∎

S.2.2 Proof of Lemma 6

Proof. We note that

$$
\begin{aligned}
&\left(\hat{\varepsilon}^m(y,x,\theta)-\varepsilon^m(y,x,\theta)\right)\mathbb{1}\left(|x-x_0|\leqslant h_{n,\theta}\right)\\
&=\left(m(x,\theta)-\bar{r}_p(x-x_0)^\top\hat{\alpha}(\theta)\right)\mathbb{1}\left(|x-x_0|\leqslant h_{n,\theta}\right)\\
&=\left(\bar{r}_p\left(\frac{x-x_0}{h_{n,\theta}}\right)^\top\left(\bar{H}_{n,\theta}^{-1}\alpha(\theta)-\bar{H}_{n,\theta}^{-1}\hat{\alpha}(\theta)\right)+\frac{(x-x_0)^{p+1}}{(p+1)!}\left(m^{(p+1)}(\theta,\bar{x}^+)\delta_{x-x_0}^+ + m^{(p+1)}(\theta,\bar{x}^-)\delta_{x-x_0}^-\right)\right)\\
&\quad\cdot\mathbb{1}\left(|x-x_0|\leqslant h_{n,\theta}\right)\\
&\overset{(1)}{\lesssim}\left(O\left(h_n^{p+1}\right)+O(1)\right)O(h_n)\\
&=O(h_n)
\end{aligned}
$$

uniformly in $\theta\in\bar{\Theta}$, where inequality (1) follows from Equation (S.8). Therefore, we have $\hat{\varepsilon}^m(y,x,\theta)\mathbb{1}\left(|x-x_0|\leqslant h_{n,\theta}\right)=\varepsilon^m(y,x,\theta)\mathbb{1}\left(|x-x_0|\leqslant h_{n,\theta}\right)+O_P(h_n)$ uniformly in $\theta\in\bar{\Theta}$. This completes the proof. ∎

S.2.3 Proof of Theorem 2

Proof. **(i)** The estimator of RKD_m is given by

$$
\widehat{\mathrm{RKD}}_m(\cdot)=\frac{\hat{m}^{(1)}(\cdot,x_0^+)-\hat{m}^{(1)}(\cdot,x_0^-)}{b'(x_0^+)-b'(x_0^-)}.
$$

Then, it follows from Lemma 4 that

$$
\sqrt{nh_{n,\theta}}\left(\widehat{\mathrm{RKD}}_m(\theta)-\mathrm{RKD}_m(\theta)\right)
$$

$$= \frac{\sqrt{nh_{n,\theta}^3}\left(\hat{m}^{(1)}(\theta,x_0^+) - m^{(1)}(\theta,x_0^+)\right)}{b'(x_0^+) - b'(x_0^-)} + \frac{\sqrt{nh_{n,\theta}^3}\left(\hat{m}^{(1)}(\theta,x_0^-) - m^{(1)}(\theta,x_0^-)\right)}{b'(x_0^+) - b'(x_0^-)}$$
$$\rightsquigarrow \frac{\mathbb{Z}^m(\theta,2) - \mathbb{Z}^m(\theta,3)}{b'(x_0^+) - b'(x_0^-)} := \mathbb{G}^m(\theta)$$

uniformly in $\theta \in \bar{\Theta}$. The covariance function is

$$E\left[\mathbb{G}^m(\theta_1)\mathbb{G}^m(\theta_2)\right] = \frac{1}{(b'(x_0^+) - b'(x_0^-))^2} E\left[\left(\mathbb{Z}^m(\theta_1,2) - \mathbb{Z}^m(\theta_1,3)\right)\left(\mathbb{Z}^m(\theta_2,2) - \mathbb{Z}^m(\theta_2,3)\right)\right]$$
$$= \frac{(\iota_2 - \iota_3)^\top \bar{\Gamma}_p^{-1}\left(\bar{\Psi}_p^+(\theta_1,\theta_2)\sigma^m(\theta_1,\theta_2|x_0^+) + \bar{\Psi}_p^-(\theta_1,\theta_2)\sigma^m(\theta_1,\theta_2|x_0^-)\right)\bar{\Gamma}_p^{-1}(\iota_2 - \iota_3)}{(b'(x_0^+) - b'(x_0^-))^2 f_X(x_0)}$$

for all $\theta_1, \theta_2 \in \bar{\Theta}$.

(ii) Under Assumption 5(i),

$$\widehat{\mathrm{RKD}}_{\psi|m}(\cdot) = \psi\left(\widehat{\mathrm{RKD}}_m, \widehat{\mathrm{QRKD}}\right)(\cdot) + o_P(1).$$

Then, under Assumptions 3(iv), 4(v), and 5(ii)–(iii), an application of Lemmas 4–5 and the functional Delta method (e.g., Theorem 20.8 in Van der Vaart, 2000) yields:

$$\sqrt{nh_n^3}\left(\widehat{\mathrm{RKD}}_{\psi|m}(\theta') - \mathrm{RKD}_{\psi|m}(\theta')\right)$$
$$= \sqrt{nh_n^3}\left\{\psi\left(\widehat{\mathrm{RKD}}_m(\cdot), \widehat{\mathrm{QRKD}}(\cdot)\right)(\theta') - \psi\left(\mathrm{RKD}_m(\cdot), \mathrm{QRKD}(\cdot)\right)(\theta')\right\} + o_P(1)$$
$$= \psi'_{(\mathrm{RKD}_m,\mathrm{QRKD})}\left[\sqrt{nh_n^3}\left(\widehat{\mathrm{RKD}}_m(\cdot) - \mathrm{RKD}_m(\cdot)\right), \sqrt{nh_n^3}\left(\widehat{\mathrm{QRKD}}(\cdot) - \mathrm{QRKD}(\cdot)\right)\right](\theta') + o_P(1)$$
$$\rightsquigarrow \psi'_{(\mathrm{RKD}_m,\mathrm{QRKD})}\left(\mathbb{G}^m(\cdot)/\sqrt{\varsigma^3(\cdot)}, \mathbb{G}^Q(\cdot)/\sqrt{c^3(\cdot)}\right)(\theta')$$

uniformly in $\theta' \in \bar{\Theta}'$. Finally, the bilinearity of $\psi'_{(\mathrm{RKD}_m,\mathrm{QRKD})}$ implies $\mathbb{G}^{\psi|m}$ is mean zero (Alt, 2016, Theorem 5.11). This completes the proof. ■

S.2.4 Proof of Theorem 3

Proof. We define the EMPs of $\sqrt{nh_{n,\theta}}\left(\widehat{\mathrm{RKD}}_m(\theta) - \mathrm{RKD}_m(\theta)\right)$ as

$$\widehat{\mathbb{G}}_\xi^m(\theta) := \frac{\sum_{i=1}^n f_{\xi,ni}(\theta)\hat{\varepsilon}^m(Y_i,X_i,\theta)}{\hat{f}_X(x_0)}, \quad \mathbb{G}_\xi^m(\theta) := \frac{\sum_{i=1}^n f_{\xi,ni}(\theta)\varepsilon^m(Y_i,X_i,\theta)}{f_X(x_0)}$$

where

$$f_{\xi,ni}(\theta) := \xi_i \frac{(\iota_2 - \iota_3)^\top \bar{\Gamma}_p^{-1}\bar{r}_p\left(\frac{X_i - x_0}{h_{n,\theta}}\right)K\left(\frac{X_i - x_0}{h_{n,\theta}}\right)}{(b'(x_0^+) - b'(x_0^-))\sqrt{nh_{n,\theta}}}.$$

We first establish the uniform stochastic equivalence between the processes $\widehat{\mathbb{G}}_\xi^m(\cdot)$ and $\mathbb{G}_\xi^m(\cdot)$ on $\bar{\Theta}$. To show this, consider the difference:

$$\widehat{\mathbb{G}}_\xi^m(\theta) - \mathbb{G}_\xi^m(\theta) = \frac{\sum_{i=1}^n f_{\xi,ni}(\theta)\left(\hat{\varepsilon}^m(Y_i,X_i,\theta) - \varepsilon^m(Y_i,X_i,\theta)\right)}{\hat{f}_X(x_0)} + \frac{f_X(x_0) - \hat{f}_X(x_0)}{\hat{f}_X(x_0)}\mathbb{G}_\xi^m(\theta).$$

Then, for every $\eta > 0$, by the Markov's and Minkowski's inequalities, we obtain:

$$\begin{aligned}
P\left[\sup_{\theta\in\bar{\Theta}}\left|\widehat{\mathbb{G}}_\xi^m(\theta) - \mathbb{G}_\xi^m(\theta)\right|\right] &\leqslant \frac{1}{\eta}E\left[\sup_{\theta\in\bar{\Theta}}\left|\widehat{\mathbb{G}}_\xi^m(\theta) - \mathbb{G}_\xi^m(\theta)\right|\right] \\
&\leqslant \frac{1}{\eta}\Bigg(\underbrace{E\left[\sup_{\theta\in\bar{\Theta}}\left|\frac{\sum_{i=1}^n f_{\xi,ni}(\theta)\left(\hat{\varepsilon}^m(Y_i,X_i,\theta) - \varepsilon^m(Y_i,X_i,\theta)\right)}{\hat{f}_X(x_0)}\right|\right]}_{\text{(I)}} + \\
&\qquad + \underbrace{E\left[\sup_{\theta\in\bar{\Theta}}\left|\frac{f_X(x_0) - \hat{f}_X(x_0)}{\hat{f}_X(x_0)}\mathbb{G}_\xi^m(\theta)\right|\right]}_{\text{(II)}}\Bigg),
\end{aligned}$$

where the probability space is $(\Omega^w \times \Omega^\xi, \mathcal{S}^w \times \mathcal{S}^\xi, P := P_{w,\xi})$. The expectation operator is $E := E_{w,\xi} = E_w E_{\xi|w}$, where $E_{\xi|w}$ denotes the conditional expectation given the remaining data $\{(Y_i, X_i) : 1 \leqslant i \leqslant n\}$. For the term (I),

$$\begin{aligned}
\text{(I)} &= E_w\left[\frac{1}{\hat{f}_X(x_0)}E_{\xi|w}\left[\sup_{\theta\in\bar{\Theta}}\left|\sum_{i=1}^n f_{\xi,ni}(\theta)\left(\hat{\varepsilon}^m(Y_i,X_i,\theta) - \varepsilon^m(Y_i,X_i,\theta)\right)\right|\right]\right] \\
&\overset{(1)}{=} E_w\left[\frac{o_{P_w}(1)}{f_X(x_0) + o_{P_w}(1)}E_{\xi|w}\left[\sup_{\theta\in\bar{\Theta}}\left|\sum_{i=1}^n f_{\xi,ni}(\theta)\right|\right]\right] \\
&\overset{(2)}{=} E_w\left[\frac{o_{P_w}(1)}{f_X(x_0) + o_{P_w}(1)}O_{P_w}(1)\right] \\
&= o(1).
\end{aligned}$$

Here, Equality (1) follows from Assumption 6(i) and Lemma 6. For equality (2), by applying analogous arguments to those in the proof of Lemma 4—specifically, applying them to the process $\sum_{i=1}^n f_{\xi,ni}(\cdot)$ conditional on the data (i.e., with respect to $P_{\xi|w}$)—it can be shown that $\sum_{i=1}^n f_{\xi,ni}(\cdot) \underset{\xi}{\overset{P}{\rightsquigarrow}} \widetilde{G}(\cdot)$ uniformly on $\bar{\Theta}$ for some tight Gaussian process $\widetilde{G}$. Consequently, by Prohorov's Theorem (e.g., van der Vaart and Wellner, 2023, Theorem 1.3.9), this process is bounded in probability P_w, i.e., $\sum_{i=1}^n f_{\xi,ni}(\cdot) = O_{P_w}(1)$.

Similarly, for the term (II),

$$\begin{aligned}
\text{(II)} &\overset{(3)}{=} E_w\left[\frac{o_{P_w}(1)}{f_X(x_0) + o_{P_w}(1)}E_{\xi|w}\left[\sup_{\theta\in\bar{\Theta}}\left|\mathbb{G}_\xi^m(\theta)\right|\right]\right] \\
&\overset{(4)}{=} E_w\left[\frac{o_{P_w}(1)}{f_X(x_0) + o_{P_w}(1)}O_{P_w}(1)\right]
\end{aligned}$$

$$= o(1).$$

Here, Equality (3) holds by Assumption 6(i). Equality (4) is a consequence of Theorem 11.19 in Kosorok (2008), which implies the weak convergence $\mathbb{G}_\xi^m(\cdot) \underset{\xi}{\overset{P}{\rightsquigarrow}} \mathbb{G}^m(\cdot)$. By Prohorov's Theorem, this weak convergence ensures that $\mathbb{G}_\xi^m(\cdot) = O_{P_w}(1)$. Combining these steps establishes that $\mathrm{d}\left(\widehat{\mathbb{G}}_\xi^m(\cdot), \mathbb{G}_\xi^m(\cdot)\right) \xrightarrow{P} 0$ uniformly on $\bar{\Theta}$. Finally, applying Lemma S.2.3 yields the desired weak convergence: $\widehat{\mathbb{G}}_\xi^m(\cdot) \underset{\xi}{\overset{P}{\rightsquigarrow}} \mathbb{G}^m(\cdot)$ uniformly on $\bar{\Theta}$.

(ii) Following Chiang and Sasaki (2019) and Qu and Yoon (2015), the pivotal process converges weakly: $\widehat{\mathbb{G}}_u^Q(\cdot) \underset{u}{\overset{P}{\rightsquigarrow}} \mathbb{G}^Q(\cdot)$ uniformly on $\mathcal{T}$. Here, the subscript u indicates that the process is constructed using i.i.d. Uniform$(0, 1)$ random weights, drawn independently of the data. Given that the Hadamard derivative $\psi'_{(\mathrm{RKD}_m,\mathrm{QRKD})}$ is a continuous linear operator, an application of the continuous mapping theorem for bootstrapped empirical processes (e.g., Theorem 10.8 of Kosorok, 2008), combined with the result from Part (i) and the weak convergence of $\widehat{\mathbb{G}}_u^Q$, yields:

$$\begin{aligned}
\widehat{\mathbb{G}}_{\xi,u}^{\psi|m}(\theta') &\overset{(5)}{=} \psi'_{(\mathrm{RKD}_m,\mathrm{QRKD})}\left(\frac{\widehat{\mathbb{G}}_\xi^m(\cdot)}{\sqrt{\varsigma(\cdot)}}, \frac{\widehat{\mathbb{G}}_u^Q(\cdot)}{\sqrt{c(\cdot)}}\right)(\theta') + o_P(1) \\
&\underset{\xi\times u}{\overset{P}{\rightsquigarrow}} \psi'_{(\mathrm{RKD}_m,\mathrm{QRKD})}\left(\frac{\mathbb{G}^m(\cdot)}{\sqrt{\varsigma(\cdot)}}, \frac{\mathbb{G}^Q(\cdot)}{\sqrt{c(\cdot)}}\right)(\theta')
\end{aligned}$$

uniformly in $\theta' \in \bar{\Theta}'$, where the equality (5) follows from Assumption 6(ii). This completes the proof. ■

S.2.5 Proof of Corollary 1

Proof. **(i)** Since the LTE $\Delta_m(\theta)$ is identified by the estimand $\mathrm{RKD}_m(\theta)$, the null hypothesis $\mathscr{H}_{0,m}^S : \Delta_m(\theta) = 0$ for all $\theta \in \bar{\Theta}$ is equivalent to $\mathscr{H}_{0,m}^S : \mathrm{RKD}_m(\theta) = 0$ for all $\theta \in \bar{\Theta}$. Under this null hypothesis, Theorem 2(i) implies:

$$W_m^s = \sup_{\theta\in\bar{\Theta}} \left|\sqrt{nh_{n,\theta}^3}\widehat{\mathrm{RKD}}_m(\theta)\right| \rightsquigarrow \sup_{\theta\in\bar{\Theta}} |\mathbb{G}^m(\theta)|\,.$$

Next, consider the functional $\tilde{\psi}(h) := h - \frac{1}{|\bar{\Theta}|}\int_{\bar{\Theta}} h(\theta)\, d\theta$. As this functional is linear, it is Hadamard differentiable at RKD_m tangentially to $\ell^\infty(\bar{\Theta})$. Therefore, under the null hypothesis $\mathscr{H}_{0,m}^H : \Delta_m(\theta_1) = \Delta_m(\theta_2)$ for all $\theta_1, \theta_2 \in \bar{\Theta}$, it follows that:

$$\begin{aligned}
W_m^H &= \sup_{\theta\in\bar{\Theta}} \left|\sqrt{nh_{n,\theta}^3}\left(\widehat{\mathrm{RKD}}_m(\theta) - \frac{1}{|\bar{\Theta}|}\int \widehat{\mathrm{RKD}}_m(\theta)\, d\theta\right)\right| \\
&\rightsquigarrow \sup_{\theta\in\bar{\Theta}} \left|\tilde{\psi}'_{\mathrm{RKD}_m}(\mathbb{G}^m)(\theta)\right| \\
&= \sup_{\theta\in\bar{\Theta}} \left|\mathbb{G}^m - \frac{1}{|\bar{\Theta}|}\int_{\bar{\Theta}} \mathbb{G}^m(\theta)\, d\theta\right|.
\end{aligned}$$

The proof of Part (ii) follows by analogous reasoning, substituting $\psi|m$ for m and applying Theorem 2(ii). This completes the proof. ■

REFERENCES

Alt, H. W. (2016). *Linear Functional Analysis: An Application-Oriented Introduction.* Springer.

Andrews, D. W. (1994). Empirical process methods in econometrics. *Handbook of Econometrics*, 4:2247–2294.

Bhattacharya, D. (2007). Inference on inequality from household survey data. *Journal of Econometrics*, 137(2):674–707.

Card, D., Lee, D. S., Pei, Z., and Weber, A. (2015). Inference on causal effects in a generalized regression kink design. *Econometrica*, 83(6):2453–2483.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2014). Gaussian approximation of suprema of empirical processes.

Chiang, H. D., Hsu, Y.-C., and Sasaki, Y. (2019). Robust uniform inference for quantile treatment effects in regression discontinuity designs. *Journal of Econometrics*, 211(2):589–618.

Chiang, H. D. and Sasaki, Y. (2019). Causal inference by quantile regression kink designs. *Journal of Econometrics*, 210(2):405–433.

Gin, E. and Nickl, R. (2015). Mathematical foundations of infinite-dimensional statistical models.

Kosorok, M. R. (2003). Bootstraps of sums of independent but not identically distributed stochastic processes. *Journal of Multivariate Analysis*, 84(2):299–318.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference.* Springer.

Pollard, D. (1990). Empirical processes: theory and applications. Ims.

Qu, Z. and Yoon, J. (2015). Nonparametric estimation and inference on conditional quantile processes. *Journal of Econometrics*, 185(1):1–19.

Sasaki, Y. (2015). What do quantile regressions identify for general structural functions? *Econometric Theory*, 31(5):1102–1116.

van der Vaart, A. and Wellner, J. A. (2023). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer Nature.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.