Online Supplementary Material to
"Quantile Treatment Effects in Regression Kink Designs"

Heng Chen* Harold D. Chiang† Yuya Sasaki‡§

November 30, 2019

Abstract

This online supplementary material contains additional mathematical details for the article "Quantile Treatment Effects in Regression Kink Designs." This material consists of auxiliary lemmas and their proofs, a proof of Theorem 2, and more details about Remark 2.

*Heng Chen: chhe@bankofcanada.ca. Currency Department, Bank of Canada, 234 Wellington Street, Ottawa, ON, K1A 0G9, Canada.

†Harold D. Chiang: harold.d.chiang@vanderbilt.edu. Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA.

‡Yuya Sasaki: yuya.sasaki@vanderbilt.edu. Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA.

§We thank Peter Phillips (editor), Arthur Lewbel (co-editor), three anonymous referees, Yingying Dong, Robert Moffitt, and participants at New York Camp Econometrics XIII for very useful comments and suggestions. All the remaining errors are ours.

C Auxiliary Lemmas and Proofs

C.1 Auxiliary Lemmas

C.1.1 Uniform Bahadur Representation

The following lemma proposes the uniform UBR for the local slope estimators.

Lemma 1 (Chiang, Hsu, and Sasaki (2019); Lemma 1)**.** *Under Assumption 2, we have the uniform influence function representations (3.4) and (3.5) that hold uniformly on* $\mathscr{Y}_1 \times \mathscr{D}$*.*

C.1.2 Functional Central Limit Theorem

Lemma 2. *Let the triangular array of separable stochastic processes* $\{f_{ni}(\omega,t) : i = 1, ...n, t \in T\}$ *be row independent and write* $X_n(t) = \sum_{i=1}^{n}[f_{ni}(\omega,t) - Ef_{ni}(\cdot,t)]$*, and denote* E^* *to be the outer integral (see e.g., van der Vaart and Welner, 1996, Section 1.2). Suppose that the following conditions are satisfied:*

1. $\{f_{ni}\}$ *are manageable, with envelope* $\{F_{ni}\}$ *which are also independent within rows;*
2. $H(s,t) = \lim_{n\to\infty} EX_n(s)X_n(t)$ *exists for every* $s,t \in T$*;*
3. $\limsup_{n\to\infty} \sum_{i=1}^{n} E^*F_{ni}^2 < \infty$*;*
4. $\lim_{n\to\infty} \sum_{i=1}^{n} E^*F_{ni}^2 \mathbb{1}\{F_{ni} > \epsilon\} = 0$ *for each* $\epsilon > 0$*;*
5. $\rho(s,t) = \lim_{n\to\infty} \rho_n(s,t)$*, where* $\rho_n(s,t) = (\sum_{i=1}^{n} E[f_{ni}(\cdot,s) - f_{ni}(\cdot,t)]^2)^{1/2}$*, exists for every* $s,t \in T$*, and for all deterministic sequences* $\{s_n\}$ *and* $\{t_n\}$ *in* $\mathbb{T}$*, if* $\rho(s_n,t_n) \to 0$ *then* $\rho_n(s_n,t_n) \to 0$.

Then T *is totally bounded under the* ρ *pseudometric, and* X_n *converges weakly to a tight mean zero Gaussian process* $\mathbb{X}$ *concentrated on* $\{z \in l^\infty(T) : z \text{ is uniformly } \rho-\text{continuous}\}$*, with covariance* $H(s,t)$*.*

C.2 Proof of Theorem 2

Before starting to present a proof of the theorem, we introduce additional definitions and notations for the proof of the theorem. Let $\mathcal{F}$ be a class of measurable functions defined on $(\Omega, \mathscr{F})$ with a measurable envelope F. We say that $\mathcal{F}$ is of VC type with envelope F if there exist constants A, $v > 0$ such that $\sup_Q N(\mathcal{F}, L^2(Q), \varepsilon \|F\|_{Q,2}) \leq (A/\varepsilon)^v$, where the supremum is taken over the set of all finite discrete measures Q on $\mathcal{F}$.

To approximate the distribution of the UBR, we define the following Multiplier Processes (MP):

$$\nu^{\pm}_{\xi,n}(y,d,1) = \frac{1}{\sqrt{nh_n} f_X(0)} \sum_{i=1}^{n} \xi_i e_1^\top (\Gamma^{\pm})^{-1} r_3\Big(\frac{X_i}{h_n}\Big) \Big[\mathbb{1}\{Y_i \leq y, D_i = d\} - \mu_1(X_i, y, d)\Big] K\Big(\frac{X_i}{h_n}\Big) \delta_i^{\pm},$$
$$\nu^{\pm}_{\xi,n}(y,d,2) = \frac{1}{\sqrt{nh_n} f_X(0)} \sum_{i=1}^{n} \xi_i e_1^\top (\Gamma^{\pm})^{-1} r_3\Big(\frac{X_i}{h_n}\Big) \Big[\mathbb{1}\{D_i = d\} - \mu_2(X_i, d)\Big] K\Big(\frac{X_i}{h_n}\Big) \delta_i^{\pm}.$$

For ease of writing, we use the following notations for the differences of right and left limits of the UBR, the MP, and the EMP with $k = 1, 2$:

$$\nu_n(y,d,k) = \nu_n^{+}(y,d,k) - \nu_n^{-}(y,d,k),$$
$$\nu_{\xi,n}(y,d,k) = \nu_{\xi,n}^{+}(y,d,k) - \nu_{\xi,n}^{-}(y,d,k),$$
$$\hat{\nu}_{\xi,n}(y,d,k) = \hat{\nu}_{\xi,n}^{+}(y,d,k) - \hat{\nu}_{\xi,n}^{-}(y,d,k).$$

With these preparations, we now start a proof of Theorem 2.

Part (i) (a): We will verify the five conditions in Lemma 2 for the triangular array of stochastic processes $\{f_{ni}\}$ defined by

$$f_{ni}(y,d,1) = \frac{1}{\sqrt{nh_n} f_X(0)} e_1^\top (\Gamma^{+})^{-1} r_3\Big(\frac{X_i}{h_n}\Big) \Big[\mathbb{1}\{Y_i \leq y\}\mathbb{1}\{D_i = d\} - \mu_1(X_i, y, d)\Big] K\Big(\frac{X_i}{h_n}\Big) \delta_i^{+},$$
$$f_{ni}(y,d,2) = \frac{1}{\sqrt{nh_n} f_X(0)} e_1^\top (\Gamma^{+})^{-1} r_3\Big(\frac{X_i}{h_n}\Big) \Big[\mathbb{1}\{D_i = d\} - \mu_2(X_i, y, d)\Big] K\Big(\frac{X_i}{h_n}\Big) \delta_i^{+},$$
$$\nu_n^{+}(y,d,k) = \sum_{i=1}^{n} [f_{ni}(y,d,k) - E f_{ni}(y,d,k)].$$

The separability follows the same argument as in the proof of Theorem 4 of Kosorok (2003) and the left or right continuity of the processes. To show condition 1, define

$$\mathscr{F}_n = \{(y^*, d^*, x^*) \mapsto \mathbb{1}\{x^* \geq 0\}[(\mathbb{1}\{y^* \leq y, d^* = d\} - \mu_1(x^*, y, d))\mathbb{1}\{k = 1\}$$
$$+ (\mathbb{1}\{d^* = d\} - \mu_2(x^*, y, d))\mathbb{1}\{k = 2\}] : (y, d, k) \in \mathscr{Y}_1 \times \{0, 1\} \times \{1, 2\}\}$$
$$\mathscr{F}_n^+ = \{f_{ni}(y, d, k) : (y, d, k) \in \mathscr{Y}_1 \times \{0, 1\} \times \{1, 2\}\}$$

We first claim that $\mathscr{F}_n^+$ is a VC type class with envelope

$$F_n^+(y^*, d^*, x^*) = \frac{C''}{\sqrt{nh_n}} \|K\|_\infty \mathbb{1}\{|x^*/h_n| \in [-1, 1]\}$$

for some constant $C'' > 0$. It is clear $\{(y^*, d^*, x^*) \mapsto \mathbb{1}\{y^* \leq y\} : y \in \mathscr{Y}_1\}$ is of VC-subgraph with VC index ≤ 2, since it is monotone increasing in y, and thus for each pair $(y_1^*, x_1^*, d_1^*, r_1), (y_2^*, x_2^*, d_2^*, r_2) \in \mathscr{Y}_1 \times \mathscr{X} \times \{0, 1\} \times \mathbb{R}$ with $y_1^* \leq y_2^*$, it can never pick out $\{(y_2^*, x_2^*, d_2^*, r_2)\}$. Similarly, $\{(y^*, d^*, x^*) \mapsto \mathbb{1}\{d^* = d\} : d \in \{1, 2\}\}$, $\{(y^*, d^*, x^*) \mapsto \{\mathbb{1}\{k^* = k\} : k \in \{1, 2\}\}$ and $\{(y^*, d^*, x^*) \mapsto \mathbb{1}\{x^* \geq 0\}\}$ are all VC subgraph classes, since they are sub-collections of all half spaces and then by Lemma 9.12 (i) of Kosorok (2008). Each of them is therefore of VC type with envelope 1. Next, Assumption 2(ii) (a) and (ii) (b) imply

$$|\mu_{k_1}(x^*, y_1, d_1) - \mu_{k_2}(x^*, y_2, d_2)| \leq L \|(k_1, y_1, d_1) - (k_2, y_2, d_2)\|$$

for an $L > 0$ and Euclidean norm $\|\cdot\|$. Thus $\{x^* \mapsto \mu_k(x, y, d) : (k, y, d) \in \{1, 2\} \times \mathscr{Y}_1 \times \mathscr{D}\}$ is of VC type with envelope L in light of Example 19.7 of van der Vaart (1998) and Lemma 9.18 of Kosorok (2008). Under Assumption 2 (i) (b), (iii) and (iv), for each n, the collection of a single function

$$\{(y^*, d^*, x^*) \mapsto \frac{e_1^\top(\Gamma^+)^{-1} r_3(x^*/h_n)\mathbb{1}\{|x^*/h_n| \in [-1, 1]\}}{\sqrt{nh_n} f_X(0)}\}$$

is of VC subgraph and therefore VC type with envelope $\frac{C'\mathbb{1}\{|x^*/h_n| \in [-1,1]\}}{\sqrt{nh_n}}$. Example 19.19 of van der Vaart (1998) suggests VC type classes, that are of finite uniform integrals, are closed under element-wise addition and multiplication. Therefore, $\mathscr{F}_n$ is of VC type with envelope constant C'' and thus

$$\mathscr{F}_n^+ = \{\frac{e_1^\top(\Gamma^+)^{-1} r_3(\cdot/h_n) K(\cdot/h_n)}{\sqrt{nh_n} f_X(0)} \cdot f : f \in \mathscr{F}_n\}$$

is of VC type with envelope $F_n^+(y^*,d^*,x^*)=\frac{C''}{\sqrt{nh_n}}\|K\|_\infty \mathbb{1}\{x^*/h_n\in[-1,1]\}$. Finally, standard calculations show for each n and for any $\delta\in(0,1)$ the uniform entropy integral bound

$$\int_0^\delta \sup_Q \sqrt{1+\log N(\mathscr{F}_n, L_2(Q), \varepsilon\|F_n\|_{Q,2})}d\varepsilon \lesssim \delta\sqrt{v\log(A/\delta)}.$$

Equation (A.1) in the proof of Theorem 1 in Andrews (1994) then implies that $\mathscr{F}_n^+$ is a manageable class of functions, as defined in Section 11.4.1 of Kosorok (2008). To check condition 2, notice

$$\begin{aligned}E\nu_n^+(y_1,d_1,k_1)\nu_n^+(y_2,d_2,k_2)=&\sum_{i=1}^n Ef_{ni}(y_1,d_1,k_1)f_{ni}(y_2,d_2,k_2)\\&-(\sum_{i=1}^n Ef_{ni}(y_1,d_1,k_1))(\sum_{i=1}^n Ef_{ni}(y_2,d_2,k_2)).\end{aligned}$$

It suffices to check $\sum_{i=1}^n Ef_{ni}(y_1,d_1,k_1)f_{ni}(y_2,d_2,k_2)<\infty$ since $Ef_{ni}(y,d,k)=0$ due to the law of iterated expectations, and thus the second term is 0. When $k_1=k_2=1$, under Assumption 2 (i) (a), (i) (b), (ii) (c), (iii), and (iv) (a),

$$\begin{aligned}&\sum_{i=1}^n Ef_{ni}(y_1,d_1,1)f_{ni}(y_2,d_2,1)\\=&E[\frac{e_1^\top(\Gamma^+)^{-1}r_3(\frac{X_i}{h_n})r_3^\top(\frac{X_i}{h_n})(\Gamma^+)^{-1}e_1}{h_nf_X^2(0)}[\mathbb{1}\{Y_i\le y_1,D_i=d_1\}-\mu_1(X_i,y_1,d_1)]\\&\times[\mathbb{1}\{Y_i\le y_2,D_i=d_2\}-\mu_1(X_i,y_2,d_2)]K^2(\frac{X_i}{h_n})\delta_i^+]\\=&\int_{\mathbb{R}_+}\frac{e_1^\top(\Gamma^+)^{-1}r_3(u)r_3^\top(u)(\Gamma^+)^{-1}e_1}{f_X^2(0)}K^2(u)\left(\sigma_{11}((y_1,d_1),(y_2,d_2)|0^+)+O\left(uh_n\right)\right)(f_X(0)+O(uh_n))du\\=&\int_{\mathbb{R}_+}\frac{e_1'(\Gamma^+)^{-1}r(u)r'^+)^{-1}e_1}{f_X(0)}K^2(u)\sigma_{11}((y_1,d_1),(y_2,d_2)|0^+)du+O(h_n)<\infty\end{aligned}$$

where the second to the last equality is due to mean value expansions under Assumption 2 (i) (b) and (ii) (c). Notice that n enters only through the $O(h_n)$ term, and thus

$$\lim_{n\to\infty}\sum_{i=1}^n Ef_{ni}(y_1,d_1,1)f_{ni}(y_2,d_2,1)$$

exists. Similar calculations hold for $k_1=k_2=1$ and $k_1=1$, $k_2=2$. This shows condition 2.

Condition 3 is clear since

$$\lim_{n\to\infty}\sum_{i=1}^{n} E\left[F_n^{+}(y^*,d^*,x^*)\right]^2 = \lim_{n\to\infty}\int \frac{(C''^2}{h_n}\|K\|_\infty^2 \mathbb{1}\{|x/h_n|\in[-1,1]\}f_X(x)dx$$
$$= f(0)(C''^2\|K\|_\infty^2 < \infty$$

under Assumption 2 (i) (a), (iii) and (iv) (a). To show condition 4, note that for each $\varepsilon > 0$,

$$\lim_{n\to\infty}\sum_{i=1}^{n} E[\big(F_n^{+}(y^*,d^*,x^*)\big)^2 \mathbb{1}\{F_n^{+}(y^*,d^*,x^*) > \varepsilon\}]$$
$$= \lim_{n\to\infty}\int_{\mathbb{R}} \frac{(C''^2}{h_n}\|K\|_\infty^2 \mathbb{1}\{x/h_n\in[-1,1]\}\mathbb{1}\{\frac{(C'')}{nh_n}\|K\|_\infty \mathbb{1}\{x/h_n\in[-1,1]\} > \varepsilon\}f_X(x)dx$$
$$\leq \int_{\mathbb{R}} (C''^2\|K\|_\infty^2 \mathbb{1}\{u\in[-1,1]\}\lim_{n\to\infty}\mathbb{1}\{\frac{(C'')}{nh_n}\|K\|_\infty \mathbb{1}\{u\in[-1,1]\} > \varepsilon\}f_X(0)du + O(h_n) = 0$$

under Assumption 2 (i) (a), (iii) and (iv) (a). This shows condition 4. To show condition 5, note that we can write

$$\rho_n^2((y_1,d_1,k_1),(y_2,d_2,k_2)) = \sum_{i=1}^{n} E[f_{ni}(y_1,d_1,k_1) - f_{ni}(y_2,d_2,k_2)]^2$$
$$= \sum_{i=1}^{n} E[f_{ni}^2(y_1,d_1,k_1) + f_{ni}^2(y_2,d_2,k_2) - 2f_{ni}(y_1,d_1,k_1)f_{ni}(y_2,d_2,k_2)].$$

From our calculations on the way to show condition 2, we know that each term on the right-hand side exists under Assumption 2 (i) (a), (i) (b), (ii) (c), (iii), and (iv) (a). Since n enters the expression only through the $O(h_n)$ part, for all deterministic sequences $s_n \in \mathscr{Y}_1\times\{0,1\}\times\{1,2\}$ and $t_n \in \mathscr{Y}_1\times\{0,1\}\times\{1,2\}$, $\rho^2(s_n,t_n)\to 0$ implies $\rho_n^2(s_n,t_n)\to 0$. By Lemma 4, we have $\nu_n^{+} \rightsquigarrow \mathbb{G}_{+}$ and similarly for $\nu_n^{-}\rightsquigarrow \mathbb{G}_{-}$. Assumption 2 (i) (a) then implies $\nu_n = \nu_n^{+} - \nu_n^{-} \rightsquigarrow \mathbb{G} := \mathbb{G}_{+} - \mathbb{G}_{-}$.

Part (i) (b): We apply the FCLT and the functional delta method. Notice that $\nu_n \rightsquigarrow \mathbb{G}$ suggests

$$\sqrt{nh_n^3}\begin{bmatrix}(\hat{\mu}_1'(0^+,y,d)-\hat{\mu}_1'(0^-,y,d)) - (\mu_1'(0^+,y,d)-\mu_1'(0^-,y,d))\\ (\hat{\mu}_2'(0^+,d)-\hat{\mu}_2'(0^-,d)) - (\mu_2'(0^+,d)-\mu_2'(0^-,d))\end{bmatrix} = \begin{bmatrix}\mathbb{G}(y,d,1)\\ \mathbb{G}(y,d,2)\end{bmatrix}.$$

Let $(A(\cdot),B(\cdot)) \in \ell^\infty(\mathscr{Y}_1\times\{0,1\})\times\ell^\infty(\mathscr{Y}_1)$, if $B(\cdot) > C > 0$, then $(G,H)\overset{\Psi}{\mapsto} G/H$ is Hadamard differentiable at (A,B) tangentially to ℓ^∞ with the Hadamard derivative $\Psi'_{(A,B)}$

given by $\Psi'_{(A,B)}(g,h) = (Bg - Ah)/B^2$. Therefore, under Assumption 1 (ii), we know that $\mu'_2(0^+,d) - \mu'_2(0^-,d)$ is bounded away from 0. Also, $f_{Y^d|VX}(\cdot|h(0),0)$ is bounded away from zero under Assumption 2 (i) (c). The functional delta method then yields

$$\begin{aligned}
&\sqrt{nh_n^3}[\widehat{F}_{Y^d|VX}(\cdot|h(0),0) - F_{Y^d|VX}(\cdot|h(0),0)] \\
=&\sqrt{nh_n^3}[\frac{\hat{\mu}'_1(0^+,\cdot,d) - \hat{\mu}'_1(0^-,\cdot,d)}{\hat{\mu}'_2(0^+,d) - \hat{\mu}'_2(0^-,d)} - \frac{\mu'_1(0^+,\cdot,d) - \mu'_1(0^-,\cdot,d)}{\mu'_2(0^+,d) - \mu'_2(0^-,d)}] \\
\rightsquigarrow&\mathbb{G}_F(\cdot,d)
\end{aligned}$$

where

$$\mathbb{G}_F(y,d) := \frac{[\mu'_2(0^+,d) - \mu'_2(0^-,d)]\mathbb{G}(y,d,1) - [\mu'_1(0^+,y,d) - \mu'_1(0^-,y,d)]\mathbb{G}(y,d,2)}{[\mu'_2(0^+,d) - \mu'_2(0^-,d)]^2}.$$

Part (i) (c): Define operator $\Upsilon : \mathbb{D}_\Upsilon(\mathscr{Y}_1 \times \{0,1\}) \to \ell^\infty([a,1-a])$ as

$$F(\cdot,\cdot) \overset{\Upsilon}{\mapsto} \Phi(F(\cdot,1))(\cdot) - \Phi(F(\cdot,0))(\cdot) = Q(\cdot,1) - Q(\cdot,0)$$

where $\Phi(F)(\theta) = Q(\theta) = \inf\{y \in \mathscr{Y}_1 : F(y) \geq \theta\}$. By Hadamard differentiability from Lemma 3.9.23 (ii) of van der Vaart and Welner (1996) and the chain rule van der Vaart (1998, Theorem 20.9), under Assumption 2 (i) (c), (ii) (a), and (ii) (b), Υ is Hadamard differentiable at $F_{Y^\cdot|VX}(\cdot|h(0),0)$ tangentially to $\mathcal{C}(\mathscr{Y}_1 \times \mathscr{D})$ and the derivative (Kosorok, 2008, Section 2.2.4) is

$$\begin{aligned}
&\Upsilon'_{F_{Y^\cdot|VX}(\cdot|h(0),0)}(g(\cdot,\cdot)) \\
=&-\frac{g(Q_{Y^1|VX}(\cdot|h(0),0),1)}{f_{Y^1|VX}(Q_{Y^1|VX}(\cdot|h(0),0)|h(0),0)} + \frac{g(Q_{Y^0|VX}(\cdot|h(0),0),0)}{f_{Y^0|VX}(Q_{Y^0|VX}(\cdot|h(0),0)|h(0),0)}
\end{aligned}$$

is tangential to $C(\mathscr{Y}_1 \times \mathscr{D})$. The functional delta method then yields

$$\sqrt{nh_n^3}[\hat{\tau} - \tau] \rightsquigarrow \mathbb{G}_\tau$$

where

$$\mathbb{G}_\tau(\theta) = -\Big[\frac{\mathbb{G}_F(Q_{Y^1|VX}(\theta|h(0),0),1)}{f_{Y^1|VX}(Q_{Y^1|VX}(\theta|h(0),0)|h(0),0)} - \frac{\mathbb{G}_F(Q_{Y^0|VX}(\theta|h(0),0),1)}{f_{Y^0|VX}(Q_{Y^0|VX}(\theta|h(0),0)|h(0),0)}\Big].$$

Part (ii): This part of the proof consists of two steps. We first show the convergence result for the EMP, and then show the convergence result for $\widehat{\Xi}(\cdot)$.

Step 1 We claim $\nu_{\widehat{\xi},n} \underset{\xi}{\overset{p}{\rightsquigarrow}} \mathbb{G}$. Applying Theorem 11.19 of Kosorok (2008), which is applicable under the five conditions verified in (i), we have $\nu_{\xi,n} = \nu^{+}_{\xi,n} - \nu^{-}_{\xi,n} \underset{\xi}{\overset{p}{\rightsquigarrow}} \mathbb{G}$. In light of Lemma 2 of Chiang, Hsu, and Sasaki (2019), it then suffices to show

$$\sup_{(y,d,k)\in\mathscr{Y}_1\times\{0,1\}\times\{1,2\}} |\hat{\nu}^{\pm}_{\xi,n}(y,d,k) - \nu^{\pm}_{\xi,n}(y,d,k)| \underset{x\times\xi}{\overset{p}{\to}} 0.$$

Indeed, for $k=1$, by Assumption 2 (i) (b) and (v), we have

$$\begin{aligned}
&\hat{\nu}^{+}_{\xi,n}(y,d,1) - \nu^{+}_{\xi,n}(y,d,1)\\
=&\frac{1}{f_X(0)\hat{f}_X(0)}\sum_{i=1}^{n}\xi_i\frac{e_1^\top(\Gamma^+)^{-1}r_3(\frac{X_i}{h_n})K(\frac{X_i}{h_n})\delta_i^+}{\sqrt{nh_n}}[\mathbb{1}\{Y_i\le y, D_i=d\}f_X(0) - \tilde{\mu}_1(0^+,y,d)f_X(0)\\
&-\mathbb{1}\{Y_i\le y, D_i=d\}\hat{f}_X(0) + \mu_1(0^+,y,d)\hat{f}_X(0)]\\
=&\frac{1}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+[-\tilde{\mu}_1(0^+,y,d)f_X(0) + \mu_1(0^+,y,d)\hat{f}_X(0) + o_p^{x\times\xi}(1)]\\
=&\frac{1}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+[-\tilde{\mu}_1(0^+,y,d)f_X(0) + \mu_1(0^+,y,d)f_X(0)\\
&-\mu_1(0^+,y,d)f_X(0) + \mu_1(0^+,y,d)\hat{f}_X(0) + o_p^{x\times\xi}(1)]\\
=&\frac{f_X(0)}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+[-\tilde{\mu}_1(0^+,y,d) + \mu_1(0^+,y,d)]\\
&+\frac{\mu_1(0^+,y,d)}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+[-f_X(0)+\hat{f}_X(0)] + \frac{o_p^{x\times\xi}(1)}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+\\
=&\frac{f_X(0)}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+o_p^{x\times\xi}(1) + \frac{\mu_1(0^+,y,d)}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+o_p^{x\times\xi}(1)\\
&+\frac{o_p^{x\times\xi}(1)}{f_X^2(0)+o_p^{x\times\xi}(1)}\sum_{i=1}^{n}T_i^+\\
=&o_p^{x\times\xi}(1)\sum_{i=1}^{n}T_i^+ \qquad\qquad (C.1)
\end{aligned}$$

where $T_i^+ = \xi_i\frac{e_1^\top(\Gamma^+)^{-1}r_3(\frac{X_i}{h_n})K(\frac{X_i}{h_n})\delta_i^+}{\sqrt{nh_n}}$. It can be shown that the array of zero mean random variables $\{\sum_{i=1}^n T_i^+\}_{i=1}^n$ satisfies Lindeberg-Feller conditions (Proposition 2.27 of van der

Vaart, 1998) under Assumption 2 (i) (a), (iii), (iv) (a), and (iv) (c) and therefore converges in distribution to a normal distribution. Therefore, the asymptotic tightness then implies $\sum_{i=1}^{n} T_i = O_p^{x\times\xi}(1)$. Thus we conclude that equation C.1 is $o_p^{x\times\xi}(1)$.

Step 2 We will show

$$-[\frac{\hat{Z}_{\xi,n}(\hat{Q}_{Y^1|VX}(\cdot|h(0),0),1)}{\hat{f}_{Y^1|VX}(\hat{Q}_{Y^1|VX}(\cdot)|h(0),0)} - \frac{\hat{Z}_{\xi,n}(\hat{Q}_{Y^0|VX}(\cdot|h(0),0),0)}{\hat{f}_{Y^0|VX}(\hat{Q}_{Y^0|VX}(\cdot)|h(0),0)}] \underset{\xi}{\overset{p}{\rightsquigarrow}} \mathbb{G}_\tau(\cdot)$$

where

$$\hat{Z}_{\xi,n}(y,d) = \frac{[\hat{\mu}_2'(0^+,d) - \hat{\mu}_2'(0^-,d)]\hat{\nu}_{\xi,n}(y,d,1) - [\hat{\mu}_1'(0^+,y,d) - \hat{\mu}_1'(0^-,y,d)]\hat{\nu}_{\xi,n}(y,d,2)}{[\hat{\mu}_2'(0^+,d) - \hat{\mu}_2'(0^-,d)]^2}.$$

We first use Theorem 12.1 of Kosorok (2008) (the functional delta for bootstrap) along with the conclusion of Step 1 to get

$$\tilde{Z}_{\xi,n}(\cdot,\cdot) := \frac{[\mu_2'(0^+,\cdot) - \mu_2'(0^-,\cdot)]\hat{\nu}_{\xi,n}(\cdot,\cdot,1) - [\mu_1'(0^+,\cdot,\cdot) - \mu_1'(0^-,\cdot,\cdot)]\hat{\nu}_{\xi,n}(\cdot,\cdot,2)}{[\mu_2'(0^+,\cdot) - \mu_2'(0^-,\cdot)]^2} \underset{\xi}{\overset{p}{\rightsquigarrow}} \mathbb{G}_F(\cdot,\cdot).$$

Since the denominator is bounded away from 0 under Assumption 2 (i) (iv), uniform consistency of $\hat{\mu}_1'$, $\hat{\mu}_2'$ from Theorem 2 gives $\left\|\tilde{Z}_{\xi,n} - \hat{Z}_{\xi,n}\right\|_{\mathscr{Y}_1\times\{0,1\}} \underset{x\times\xi}{\overset{p}{\to}} 0$, and Lemma 2 of Chiang, Hsu, and Sasaki (2019) implies $\hat{Z}_{\xi,n} \underset{\xi}{\overset{p}{\rightsquigarrow}} \mathbb{G}_F$. Using the functional delta method for bootstrap again, we obtain

$$-[\frac{\hat{Z}_{\xi,n}(Q_{Y^1|VX}(\cdot|h(0),0),1)}{f_{Y^1|VX}(Q_{Y^1|VX}(\cdot|h(0),0)|h(0),0)} - \frac{\hat{Z}_{\xi,n}(Q_{Y^0|VX}(\cdot|h(0),0),0)}{f_{Y^0|VX}(Q_{Y^0|VX}(\cdot|h(0),0)|h(0),0)}] \underset{\xi}{\overset{p}{\rightsquigarrow}} \mathbb{G}_\tau(\cdot).$$

Since $f_{Y^d|VX}(\cdot|h(0),0)$ are bounded away from zero, using asymptotic $\rho-$equicontinuity of $\hat{Z}_{\xi,n}(\cdot,\cdot)$ following its (conditional) weak convergence and Theorem 3.7.23 of Giné and Nickl (2016), and the uniform consistency of $\hat{f}_{Y^d|VX}(\cdot|h(0),0)$ and $\hat{Q}_{Y^d|VX}(\cdot)$ with $d = 1,2$ along with Lemma 2 of Chiang, Hsu, and Sasaki (2019), we conclude part (ii) of the theorem.

C.3 Proof of Lemma 1

We prove the lemma by two steps: for each $(y,d,x) \in \mathscr{Y} \times \mathscr{D} \times ([\underline{x},\overline{x}] \setminus \{0\})$, Step 1 shows

$$\frac{\partial}{\partial x}\mu(0^\pm,y,d) = \frac{\partial}{\partial x}(f_{Y|XD}(y|x,d)P_{D|X}(d|x))$$

and Step 2 shows

$$\frac{\partial}{\partial y}\mu_1'(0^{\pm}, y, d) = \frac{\partial}{\partial y}\frac{\partial}{\partial x}E[\mathbb{1}\{Y_i \leq y, D_i = d\}|X_i = x] = \frac{\partial}{\partial x}(f_{Y|XD}(y|x,d)P_{D|X}(d|x)).$$

Step 1 For $d = 1$, under Assumptions 2 (i) (b), (ii) (a) (b), (iv) (a) and 3, for each $(y, x) \in \mathscr{Y} \times ([\underline{x}, \overline{x}] \setminus \{0\})$, for $d = 1$, applying the dominated convergence theorem, we have

$$\begin{aligned}
&\frac{\partial}{\partial x}\lim_{n\to\infty} E\Big[\frac{1}{b_n}K\Big(\frac{Y_i - y}{b_n}\Big)\mathbb{1}\{D_i = 1\}|X_i = x\Big] \\
=&\frac{\partial}{\partial x}\lim_{n\to\infty}(E\Big[\frac{1}{b_n}K\Big(\frac{Y_i - y}{b_n}\Big)\Big|X_i = x\Big]P_{D|X}(1|x) + 0) \\
=&\frac{\partial}{\partial x}\lim_{n\to\infty}\Big(\int_{\mathbb{R}} K(u)f_{Y|XD}(ub_n + y|x,1)duP_{D|X}(1|x)\Big) \\
=&\frac{\partial}{\partial x}\lim_{n\to\infty}\Big(\int_{\mathbb{R}} K(u)(f_{Y|XD}(y|x,1) + \frac{\partial}{\partial y}f_{Y|XD}(y|x,1)ub_n + \frac{\partial^2}{\partial y^2}f_{Y|XD}(y^*|x,1)\frac{u^2b_n^2}{2})duP_{D|X}(1|x)\Big) \\
=&\frac{\partial}{\partial x}\lim_{n\to\infty}((f_{Y|XD}(y|x,1) + O(b_n^2))P_{D|X}(1|x)) = \frac{\partial}{\partial x}(f_{Y|XD}(y|x,1)P_{D|X}(1|x)).
\end{aligned}$$

where $y*$ lies between y and $y + ub_n$. A similar result holds for $d = 0$.

Step 2 Under Assumptions 2 (i) (b), (ii) (a) (b), (iv) (a) and 3, for each $(y, x) \in \mathscr{Y} \times ([\underline{x}, \overline{x}] \setminus \{0\})$, for $d = 1$, an application of the dominated convergence theorem yields

$$\begin{aligned}
\frac{\partial}{\partial y}\frac{\partial}{\partial x}E\Big[\mathbb{1}\{Y_i \leq y, D_i = 1\}\Big|X_i = x\Big] =&\frac{\partial}{\partial y}\frac{\partial}{\partial x}\Big(E\Big[\mathbb{1}\{Y_i \leq y\}\Big|X_i = x\Big]P_{D|X}(1|x) + 0\Big) \\
=&\frac{\partial}{\partial y}\frac{\partial}{\partial x}F_{Y|XD}(y|x,1)P_{D|X}(1|x) \\
=&\frac{\partial}{\partial x}\frac{\partial}{\partial y}F_{Y|XD}(y|x,1)P_{D|X}(1|x) = \frac{\partial}{\partial x}f_{Y|XD}(y|x,1)P_{D|X}(1|x).
\end{aligned}$$

A similar result holds for $d = 0$. □

C.4 Proof of Lemma 2

The proof makes use of a maximal inequality from Chernozhukov, Chetverikov, & Kato (2014). Under Assumptions 2 (ii) (a) and (ii) (b) and 3, as in Section 1.6 of Tsybakov (2008), the solution to equation (B.3) can be written as

$$
\begin{aligned}
&\tilde{\alpha}(0^+, y, d)\\
=&\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big) r_3^\top\Big(\frac{X_i}{a_n}\Big)\Big]^{-1}\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\Big(\frac{1}{b_n}K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\}\Big)\Big]\\
=&\alpha(0^+, y, d) + \Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big) r_3^\top\Big(\frac{X_i}{a_n}\Big)\Big]^{-1}\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\frac{\mu^{(4)}(x_{ni}^*, y, d)}{4!}a_n^4\\
+&\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big) r_3^\top\Big(\frac{X_i}{a_n}\Big)\Big]^{-1}\\
&\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\Big(\frac{1}{b_n}K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\} - \mu(X_i, y, d)\Big)\Big]
\end{aligned}
$$

where $\alpha(0^+, y, d) = \Big[\mu(0^\pm, y, d), \mu'(0^\pm, y, d)a_n, \mu''(0^\pm, y, d)a_n^2/2!, \mu'''(0^\pm, y, d)a_n^3/3!\Big]^\top$. Multiply both sides by $e_1^\top$ to get

$$
\tilde{\mu}'(0^+, y, d) = \mu'(0^+, y, d) + (1) + (2)
$$

where

$$
\begin{aligned}
(1) =& e_1^\top\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big) r_3^\top\Big(\frac{X_i}{a_n}\Big)\Big]^{-1}\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\frac{\mu^{(4)}(x_{ni}^*, y, d)}{4!}a_n^4\\
(2) =& e_1^\top\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big) r_3^\top\Big(\frac{X_i}{a_n}\Big)\Big]^{-1}\\
&\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\Big(\frac{1}{b_n}K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\} - \mu(X_i, y, d)\Big).
\end{aligned}
$$

From Step 1 of Proof of Lemma 1 in Chiang, Hsu, and Sasaki (2019), with Assumption 2 (i) (a), (i) (b), (iii), and (iv) and 3, we have the common inverse factor

$$
\Big[\frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big) r_3^\top\Big(\frac{X_i}{a_n}\Big)\Big]^{-1} \underset{x\times\xi}{\overset{p}{\to}} \frac{(\Gamma^+)^{-1}}{f_X(0)}
$$

uniformly in (y, d). It suffices to show that each of

$$
\begin{aligned}
(3) =& \frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\frac{\mu^{(4)}(x_{ni}^*, y, d)}{4!}a_n^4\\
(4) =& \frac{1}{na_n}\sum_{i=1}^n \delta_i^+ K\Big(\frac{X_i}{a_n}\Big) r_3\Big(\frac{X_i}{a_n}\Big)\Big(\frac{1}{b_n}K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\} - \mu(X_i, y, d)\Big)
\end{aligned}
$$

converges in probability to zero uniformly. We will divide the argument into the following four steps.

Step 1 Under Assumption 2 (i) (a), (ii) (a), (ii) (b), (iii) and (iv) (a), it holds that

$$\Big|\frac{1}{na_n}\sum_{i=1}^{n}\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)\frac{\mu^{(4)}(x^*_{ni},y,d)}{4!}a_n^4\Big| \leq \frac{1}{na_n}\sum_{i=1}^{n}\Big|K\Big(\frac{X_i}{a_n}\Big)\Big|\Big|r_3\Big(\frac{X_i}{a_n}\Big)\Big|\Big|\frac{\mu^{(4)}(x^*_{ni},y,d)}{4!}a_n^4\Big|$$
$$\lesssim \frac{n}{na_n}\|K\|_\infty Ma_n^4 \to 0.$$

Step 2 We first bound the difference

$$\frac{1}{na_nb_n}\sum_{i=1}^{n}\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\}$$
$$-E\Big[\frac{1}{na_nb_n}\sum_{i=1}^{n}\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\}\Big].$$

It suffices to show that each term converges in probability uniformly. Define for each $t = 0, 1, ..., 3$

$$\mathscr{F}_t = \{(y^*,d^*,x^*)\longmapsto \delta_x^+(ax^*)^tK(ax^*)K(by^*+c)\{d^*=d\} : d\in\mathscr{D}, a,b\geq 0, c\in\mathbb{R}\} \qquad \text{and}$$
$$\mathscr{F}_{t,n} = \{(y^*,d^*,x^*)\longmapsto \delta_x^+(x/a_n)^tK(x/a_n)K((y^*-y)/b_n)\{d^*=d\} : d\in\mathscr{D}, y\in\mathscr{Y}_1\}.$$

where $\delta_x^+ = \mathbb{1}\{x\geq 0\}$ and $\delta_x^- = \mathbb{1}\{x<0\}$. Note that for a fixed t, $\mathscr{F}_{t,n}\subset\mathscr{F}_t$ for all n. Fix any t, under Assumption 2 (iv), $\{x^*\mapsto K(ax^*) : a\in\mathbb{R}\}$ is of VC Type class with measurable envelope $\|K\|_\infty$. By Proposition 3.6.12 of Giné and Nickl (2016), $x\mapsto (ax)^t\mathbb{1}\{ax\leq 1\}$ is of VC type class with measurable envelope 1 since $z\mapsto z^t\mathbb{1}\{z\leq 1\}$ is a mapping of bounded variations. Furthermore, $\{\mathbb{1}\{d^*=d\} : d\in\mathscr{D}\}$ is of VC-subgraph class and therefore of VC type. Lemma A.6 of Chernozhukov, Chetverikov, & Kato (2014) then implies that the class of their element-wise product $\mathscr{F}_t$ is of VC type with envelope $F_t = \|K\|_\infty^2$, i.e., there exist positive constants k, $v<\infty$ such that $\sup_Q N(\mathscr{F}_t, \|\cdot\|_{Q,2}, \varepsilon\|F_t\|_{Q,2}) \leq (\frac{k}{\varepsilon})^v$ for $0<\varepsilon\leq 1$ and the supremum is taken over the set of all probability measures on $(\Omega^x,\mathcal{F}^x)$. Corollary 5.1 in Chernozhukov, Chetverikov, & Kato (2014) then gives

$$E\Bigg[\Bigg\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f(Y_i,D_i,X_i)-Ef(Y_i,D_i,X_i))\Bigg\|_{\mathscr{F}_t}\Bigg] = O_p^x(1).$$

Multiplying both sides by $(\sqrt{n}a_nb_n)^{-1}$, we have

$$E\Bigg[\sup_{(y,d)\in\mathscr{Y}_1\times\mathscr{D}}\Bigg|\frac{1}{na_nb_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\}$$
$$-E\Big[\frac{1}{na_nb_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\}\Big]\Bigg|\Bigg]=O(\frac{1}{\sqrt{n}a_nb_n}).$$

The result then follows from Markov's inequality and Assumption 3.

Step 3 We now want to control

$$\frac{1}{na_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)\mu(X_i,y,d)-E\Big[\frac{1}{na_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)\mu(X_i,y,d)\Big].$$

Since under Assumption 2 (ii) (a) and (ii) (b), for any (y_1,d_1), $(y_2,d_2)\in\mathscr{Y}\times\mathscr{D}$, $|\mu(x,y_1,d_1)-\mu(x,y_2,d_2)|\le M(x)(|y_1-y_2|+|d_1-d_2|)$, this implies that $\{\mu(\cdot,y,d):y\in\mathscr{Y}_1,d\in\mathscr{D}\}$ is of VC type class in lieu of Example 19.7 of van der Vaart (1998) and Lemma 9.18 of Kosorok (2008). We can then follow the same steps as in Step 2 to show

$$E\Bigg[\sup_{(y,d)\in\mathscr{Y}_1\times\mathscr{D}}\Bigg|\frac{1}{na_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)\mu(X_i,y,d)-E\Big[\frac{1}{na_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)\mu(X_i,y,d)\Big]\Bigg|\Bigg]$$
$$=O\Big(\frac{1}{\sqrt{n}a_n}\Big).$$

The desired result of the current step then follows from Markov's inequality and Assumption 3.

Step 4 Finally, we show that the two expectations above are asymptotically equivalent uniformly in y and d. Under Assumption 2 (i) (b), (ii) (a), (ii) (b), (iii), and (iv) (a), calculations yield

$$E\Big[\frac{1}{na_nb_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)K\Big(\frac{Y_i-y}{b_n}\Big)\mathbb{1}\{D_i=d\}\Big]=$$
$$E\Big[\frac{1}{na_n}\sum_{i=1}^n\delta_i^+K\Big(\frac{X_i}{a_n}\Big)r_3\Big(\frac{X_i}{a_n}\Big)\mu(X_i,y,d)\Big]$$

by the law of iterated expectations under Assumption 3. This result, along with results from Steps 2 and 3, concludes the proof. □

C.5 On Remark 2

This appendix section proves the statement in Remark 2. We mostly follow the proof of Proposition 6 of Card, Lee, Pei, and Weber (2015). Let $\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\}$ be the "stacked" $n \times 1$ outcome variable $\{\mathbb{1}\{Y_i \le y\}\mathbb{1}\{D_i = d\}\}_{i=1}^n$, where the first n^- entries are observations to the left of x_0 and the last n^+ entries are those to the right of x_0. Let $\mathbf{Z}$ be the $n \times 8$ matrix whose i^{th} row is

$$\left(\delta_i^-, \frac{X_i}{h_n}\delta_i^-, \left(\frac{X_i}{h_n}\right)^2 \delta_i^-, \left(\frac{X_i}{h_n}\right)^3 \delta_i^-, \delta_i^+, \frac{X_i}{h_n}\delta_i^+, \left(\frac{X_i}{h_n}\right)^2 \delta_i^+, \left(\frac{X_i}{h_n}\right)^3 \delta_i^+\right).$$

Also let

$$\mathbf{W}_K = \begin{pmatrix} \mathbf{W}_K^- & 0 \\ 0 & \mathbf{W}_K^+ \end{pmatrix}$$

with $\mathbf{W}_K^\pm$ being the diagonal matrices

$$\mathbf{Diag}\left(K\left(\frac{X_1}{h_n}\right)\delta_1^\pm, ..., K\left(\frac{X_n}{h_n}\right)\delta_n^\pm\right).$$

The constrained estimator can be obtained from

$$\min_{\beta^R \in \mathcal{R}^8} \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta^R\right)^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta^R\right)$$

subject to $\mathbf{R}\beta^R = 0$ where $\mathbf{R} = (1, 0, 0, 0, -1, 0, 0, 0)$. Denote the resulting estimator by

$$\widehat{\beta}^R = \begin{pmatrix} \hat{\mu}_1^R(0^+, y, d), \hat{\mu}_1'^R(0^+, y, d)h_n, \hat{\mu}_1''^R(0^+, y, d)h_n^2/2!, \hat{\mu}_1'''^R(0^+, y, d)h_n^3/3!, \\ \hat{\mu}_1^R(0^-, y, d), \hat{\mu}_1'^R(0^-, y, d)h_n, \hat{\mu}_1''^R(0^-, y, d)h_n^2/2!, \hat{\mu}_1'''^R(0^-, y, d)h_n^3/3! \end{pmatrix}.$$

From equation (1.4.5) of Amemiya (1985), we have

$$
\begin{aligned}
&\widehat{\beta}^R - \beta \\
&= \begin{bmatrix} \left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \\ - \left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{R}^\top \left(\mathbf{R}\left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{R}^\top\right)^{-1} \mathbf{R}\left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \end{bmatrix} \mathbf{Z}^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta\right) \\
&= \left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{Z}^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta\right) \\
&- \left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{R}^\top \left(\mathbf{R}\left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{R}^\top\right)^{-1} \mathbf{R}\left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \cdot \mathbf{Z}^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta\right) \\
&= \left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{Z}^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta\right) \\
&- \Pi^{-1}\mathbf{R}^\top \left(\mathbf{R}\Pi^{-\mathbf{1}}\mathbf{R}^\top\right)^{-1} \mathbf{R}\Pi^{-1} \cdot \mathbf{Z}^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta\right) \frac{1}{nh_n} + o_p\left(\frac{1}{nh_n}\right),
\end{aligned}
$$

where the first term on the RHS is the unconstrained version and Π^{-1} is

$$
\Pi^{-1} = \begin{pmatrix} \Gamma^- & 0 \\ 0 & \Gamma^+ \end{pmatrix}.
$$

Since $\hat{\mu}_1'^R(0^+, y, d)h_n - \hat{\mu}_1'^R(0^-, y, d)h_n = \mathbf{E}\widehat{\beta}^R$, where $\mathbf{E} = (0, 1, 0, 0, 0, -1, 0, 0)$ and K is the uniform kernel, we have $\mathbf{E}\cdot\Pi^{-1} \cdot \mathbf{R}^\top = 0$. Therefore,

$$
\hat{\mu}_1'^R(0^+, y, d)h_n - \hat{\mu}_1'^R(0^-, y, d)h_n = \mathbf{E}\cdot \left(\mathbf{Z}^\top \mathbf{W}_K \mathbf{Z}\right)^{-1} \mathbf{Z}^\top \mathbf{W}_K \left(\mathbb{1}\{\mathbf{Y} \le y\}\mathbb{1}\{\mathbf{D} = d\} - \mathbf{Z}\beta\right) + o_p\left(\frac{1}{nh_n}\right),
$$

where the constrained estimator has the same asymptotic distribution as the unconstrained one. □

References

Amemiya, T. (1985) *Advanced Econometrics.* Harvard University Press.

Card, D., Lee, D., Pei, Z., and Weber, A. (2015) "Inference on Causal Effects in a Generalized Regression Kink Design,"*Econometrica,* Vol. 83, No. 6, pp. 2453–2483.

Chernozhukov, V., Chetverikov, D., and Kato, K. (2014) "Gaussian Approximation of Suprema of Empirical Processes," *The Annals of Statistics*, Vol. 42, No. 4, pp. 1564-1597.

Chiang, H., Hsu, Y.-C., and Sasaki, Y. (2019) "Robust Uniform Inference for Quantile Treatment Effects in Regression Discontinuity Designs," *Journal of Econometrics,* Vol. 211, No. 2, pp. 589–618.

Giné, E. and Nickl, R. (2016) *Mathematical Foundations of Infinite-Dimensional Statistical Models.* Cambridge University Press: Cambridge.

Kosorok, M.R. (2003) "Bootstraps of Sums of Independent but Not Identically Distributed Stochastic Processes." *Journal of Multivariate Analysis,* Vol. 84, No. 2, pp. 299–318.

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

Tsybakov, A.B. (2008) *Introduction to Nonparametric Estimation.* Springer Series in Statistics.

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

Van der Vaart, A.W. and Wellner, J.A. (1996) *Weak Convergence and Empirical Processes.* Springer.