SUPPLEMENTARY MATERIAL TO "LOCAL COMPOSITE QUANTILE REGRESSION FOR REGRESSION DISCONTINUITY"[1]

Xiao Huang Zhaoguo Zhan

September 24, 2021

This supplement contains all technical details, lemmas and proofs for the asymptotic and fixed-n results, as well as additional figures and tables.

Contents

[1]Email: xhuang3@kennesaw.edu and zzhan@kennesaw.edu.

S.1 Notation

In a sharp RD, LCQR is applied separately to eqs. (1) and (2), while in a fuzzy RD, LCQR is applied separately to eqs. (1), (2), (7) and (8). Instead of introducing four similar sets of variables, notation and proofs, we will focus on eq. (1). Exactly the same proof holds for the results based on eqs. (2), (7) and (8) with similar notation, where the subscript Y_+ used for eq. (1) becomes Y_-, T_+, T_- for eqs. (2), (7) and (8), respectively.

Consider eq. (1). Let $f_{\epsilon_{Y_+}} = (f_{\epsilon_{Y_+}}(c_1), \cdots, f_{\epsilon_{Y_+}}(c_q))^T$ be a $q \times 1$ vector, $S_{Y_+,11}(c)$ be a $q \times q$ diagonal matrix with diagonal elements $f_{\epsilon_{Y_+}}(c_k)\mu_{+,0}$, $S_{Y_+,12}(c)$ be a $q \times p$ matrix with (k, j) element $f_{\epsilon_{Y_+}}(c_k)\mu_{+,j}$, $S_{Y_+,21}(c)$ be the transpose of $S_{Y_+,12}(c)$, and $S_{Y_+,22}(c)$ be a $p \times p$ matrix with (j, j') element equal to $\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k)\mu_{+,j+j'}(c)$. Let $\Sigma_{Y_+,11}(c)$ be a $q \times q$ matrix with (k, k') element $\nu_{+,0}(c)\tau_{kk'}$, $\Sigma_{Y_+,12}(c)$ be a $q \times p$ matrix with (k, j) element $\sum_{k'=1}^{q} \tau_{kk'}\nu_{+,j}$, $\Sigma_{Y_+,21}(c)$ be the transpose of $\Sigma_{Y_+,12}(c)$, $\Sigma_{Y_+,22}(c)$ be a $p \times p$ matrix with (j, j') element equal to $\sum_{k=1}^{q} \sum_{k'=1}^{q} \tau_{kk'}\nu_{+,j+j'}(c)$. Define

$$S_{Y_+}(c) = \begin{pmatrix} S_{Y_+,11}(c) & S_{Y_+,12}(c) \\ S_{Y_+,21}(c) & S_{Y_+,22}(c) \end{pmatrix}, \quad \Sigma_{Y_+}(c) = \begin{pmatrix} \Sigma_{Y_+,11}(c) & \Sigma_{Y_+,12}(c) \\ \Sigma_{Y_+,21}(c) & \Sigma_{Y_+,22}(c) \end{pmatrix}, \tag{A.1}$$

and the partitioned inverse $S_{Y_+}^{-1}(c)$:

$$S_{Y_+}^{-1}(c) = \begin{pmatrix} (S_{Y_+}^{-1}(c))_{11} & (S_{Y_+}^{-1}(c))_{12} \\ (S_{Y_+}^{-1}(c))_{21} & (S_{Y_+}^{-1}(c))_{22} \end{pmatrix}. \tag{A.2}$$

Let $F_+(c_k, c_{k'})$ be the joint cumulative distribution function of ϵ_{Y_+} and ϵ_{T_+} at $(c_k, c_{k'})$ and assume $h_{Y_+} = h_{T_+}$. Define $\phi_{kk'} = F_+(c_k, c_{k'}) - \tau_k\tau_{k'}$. Also let $\Sigma_{YT_+,11}(c)$ be a $q \times q$ matrix with (k, k') element $\nu_{+,0}(c)\phi_{kk'}$, $\Sigma_{YT_+,12}(c)$ be a $q \times p$ matrix with (k, j) element $\sum_{k'=1}^{q} \phi_{kk'}\nu_{+,j}(c)$, $\Sigma_{YT_+,21}(c)$ be the transpose of $\Sigma_{YT_+,12}(c)$, $\Sigma_{YT_+,22}(c)$ be a $p \times p$ matrix with (j, j') element $\sum_{k=1}^{q} \sum_{k'=1}^{q} \phi_{kk'}\nu_{+,j+j'}(c)$. Define

$$\Sigma_{YT_+}(c) = \begin{pmatrix} \Sigma_{YT_+,11}(c) & \Sigma_{YT_+,12}(c) \\ \Sigma_{YT_+,21}(c) & \Sigma_{YT_+,22}(c) \end{pmatrix}. \tag{A.3}$$

Like (A.1), (A.2) and (A.3) above, a similar set of definitions can be provided to other variables on the boundary, including $S_{Y_-}(c)$, $S_{Y_-}^{-1}(c)$, $\Sigma_{Y_-}(c)$, $S_{T_+}(c)$, $S_{T_+}^{-1}(c)$, $\Sigma_{T_+}(c)$, $S_{T_-}(c)$,

$S_{T_-}^{-1}(c)$, $\Sigma_{T_-}(c)$, $F_-(c_k, c_{k'})$, $\Sigma_{YT_-}(c)$, and $\phi_{kk'}$ can also be redefined with $F_-(c_k, c_{k'})$.

Let $x_{+,i} = (X_{+,i} - x)/h_{Y_+}$ and $K_{+,i} = K(x_{+,i})$ with $x = 0$. Define

$$\begin{aligned}
u_k &= \sqrt{n_+ h_{Y_+}}(a_k - m_{Y_+}(x) - \sigma_{\epsilon_{Y_+}} c_k), \;\; k = 1, \cdots, q, \\
v_j &= h_{Y_+}^j \sqrt{n_+ h_{Y_+}}(j! b_j - m_{Y+}^{(j)}(x))/j!, \;\; j = 1, \cdots, p, \\
\Delta_{i,k} &= \frac{u_k}{\sqrt{n_+ h_{Y_+}}} + \sum_{j=1}^{p} \frac{v_j x_{+,i}^j}{\sqrt{n_+ h_{Y_+}}}, \\
r_{i,p} &= m_{Y_+}(X_{+,i}) - \sum_{j=0}^{p} m_{Y_+}^{(j)}(x)(X_{+,i} - x)^j / j!, \\
d_{i,k} &= c_k[\sigma_{\epsilon_{Y_+}}(X_{+,i}) - \sigma_{\epsilon_{Y_+}}(x)] + r_{i,p}.
\end{aligned} \tag{A.4}$$

Let $W_{Y_+,n_+}^* = (w_{Y_+,11}^*, \cdots, w_{Y_+,1q}^*, w_{Y_+,21}^*, \cdots, w_{Y_+,2p}^*)^T = (w_{Y_+,1n}^*, w_{Y_+,2n}^*)^T$, where

$$\begin{aligned}
w_{Y_+,1k}^* &= \frac{1}{\sqrt{n_+ h_{Y_+}}} \sum_{i=1}^{n_+} K(x_{+,i}) \eta_{Y_+,i,k}^*, \\
w_{Y_+,2j}^* &= \frac{1}{\sqrt{n_+ h_{Y_+}}} \sum_{k=1}^{q} \sum_{i=1}^{n_+} K(x_{+,i}) x_{+,i}^j \eta_{Y_+,i,k}^*, \\
\eta_{Y_+,i,k}^* &= I(\epsilon_{Y_+,i} \le c_k - \frac{d_{i,k}}{\sigma_{\epsilon_{Y_+,i}}}) - \tau_k.
\end{aligned} \tag{A.5}$$

Also let $W_{Y_+,n_+} = (w_{Y_+,11}, \cdots, w_{Y_+,1q}, w_{Y_+,21}, \cdots, w_{Y_+,2p})^T = (w_{Y_+,1n}, w_{Y_+,2n})^T$, where

$$\begin{aligned}
w_{Y_+,1k} &= \frac{1}{\sqrt{n_+ h_{Y_+}}} \sum_{i=1}^{n_+} K(x_{+,i}) \eta_{Y_+,i,k}, \\
w_{Y_+,2j} &= \frac{1}{\sqrt{n_+ h_{Y_+}}} \sum_{k=1}^{q} \sum_{i=1}^{n_+} K(x_{+,i}) x_{+,i}^j \eta_{Y_+,i,k}, \\
\eta_{Y_+,i,k} &= I(\epsilon_{Y_+,i} \le c_k) - \tau_k.
\end{aligned} \tag{A.6}$$

Similarly, we define $W_{T_+,n_+}^*, w_{T_+,1k}^*, w_{T_+,2j}^*, \eta_{T_+,i,k}^*, W_{T_+,n_+}, w_{T_+,1k}, w_{T_+,2j}$ and $\eta_{T_+,i,k}$.

Consider the case $p = 1$ and define $\theta = (u_1, \cdots, u_q, v_1)^T$. Let $\hat{\theta}_{n_+} = (\hat{u}_1, \cdots, \hat{u}_q, \hat{v}_1)^T$ be the transformed minimizer of (13). It can be shown that minimizing (13) is equivalent to minimizing

$$L_{n_+}(\theta) = \sum_{i=1}^{n_+} \left(K(x_{+,i}) \sum_{k=1}^{q} (\rho_{\tau_k}(\sigma_{\epsilon_{Y_+,i}}(\epsilon_{Y_+,i} - c_k) + d_{i,k} - \Delta_{i,k}) - \rho_{\tau_k}(\sigma_{\epsilon_{Y_+,i}}(\epsilon_{Y_+,i} - c_k) + d_{i,k})) \right).$$

Next, in a similar fashion we introduce the notation used for fixed-n approximations in Section S.3. The notation of $K_{+,i}$ as well as $x_{+,i}$ has been provided for (A.4). Let $S_{nY_+,11}$ be a $q\times q$ diagonal matrix with diagonal elements $f_{\epsilon_{Y_+}}(c_k)\frac{1}{n_+h_{Y+}}\sum_{i=1}^{n+}\frac{K_{+,i}}{\sigma_{\epsilon_{Y_+,i}}}$, $S_{nY_+,12}$ be a $q\times p$ matrix with (k,j) element $f_{\epsilon_{Y_+}}(c_k)\frac{1}{n_+h_{Y+}}\sum_{i=1}^{n+}\frac{K_{+,i}x_{+,i}^j}{\sigma_{\epsilon_{Y_+,i}}}$, $S_{nY_+,21}$ be the transpose of $S_{nY_+,12}$, and $S_{nY_+,22}$ be a $p\times p$ matrix with (j,j') element equal to $\sum_{k=1}^q f_{\epsilon_{Y_+}}(c_k)\frac{1}{n_+h_{Y+}}\sum_{i=1}^{n+}\frac{K_{+,i}x_{+,i}^{j+j'}}{\sigma_{\epsilon_{Y_+,i}}}$. Let $\Sigma_{nY_+,11}$ be a $q\times q$ matrix with (k,k') element $\frac{1}{n_+h_{Y+}}\sum_{i=1}^{n+}K_{+,i}^2\tau_{kk'}$, $\Sigma_{nY_+,12}$ be a $q\times p$ matrix with (k,j) element $\sum_{k'=1}^q\tau_{kk'}\frac{1}{n_+h_{Y+}}\sum_{i=1}^{n+}K_{+,i}^2x_{+,i}^j$, $\Sigma_{nY_+,21}$ be the transpose of $\Sigma_{nY_+,12}(c)$, $\Sigma_{nY_+,22}$ be a $p\times p$ matrix with (j,j') element equal to $\sum_{k=1}^q\sum_{k'=1}^q\tau_{kk'}\frac{1}{n_+h_{Y+}}\sum_{i=1}^{n}K_{+,i}^2x_{+,i}^{j+j'}$.

Similar to (A.1) and (A.2), define

$$S_{nY_+}=\begin{pmatrix} S_{nY_+,11} & S_{nY_+,12}\\ S_{nY_+,21} & S_{nY_+,22}\end{pmatrix},\quad \Sigma_{nY_+}=\begin{pmatrix} \Sigma_{nY_+,11} & \Sigma_{nY_+,12}\\ \Sigma_{nY_+,21} & \Sigma_{nY_+,22}\end{pmatrix}, \tag{A.7}$$

$$S_{nY_+}^{-1}=\begin{pmatrix} (S_{nY_+}^{-1})_{11} & (S_{nY_+}^{-1})_{12}\\ (S_{nY_+}^{-1})_{21} & (S_{nY_+}^{-1})_{22}\end{pmatrix}. \tag{A.8}$$

Similar to (A.3), let $\Sigma_{nYT_+,11}$ be a $q\times q$ matrix with (k,k') element $\phi_{kk'}\frac{1}{n_+\sqrt{h_{Y+}h_{T+}}}\sum_{i=1}^{n+}K_{+,i}^2$, $\Sigma_{nYT_+,12}$ be a $q\times p$ matrix with (k,j) element $\sum_{k'=1}^q\phi_{kk'}\frac{1}{n_+\sqrt{h_{Y+}h_{T+}}}\sum_{i=1}^{n+}K_{+,i}^2x_{+,i}^j$, $\Sigma_{nYT_+,21}$ be the transpose of $\Sigma_{nYT_+,12}$, and $\Sigma_{nYT_+,22}$ be a $p\times p$ matrix with (j,j') element equal to $\sum_{k=1}^q\sum_{k'=1}^q\phi_{kk'}\frac{1}{n_+\sqrt{h_{Y+}h_{T+}}}\sum_{i=1}^{n+}K_{+,i}^2x_{+,i}^{j+j'}$. Define

$$\Sigma_{nYT_+}=\begin{pmatrix} \Sigma_{nYT_+,11} & \Sigma_{nYT_+,12}\\ \Sigma_{nYT_+,21} & \Sigma_{nYT_+,22}\end{pmatrix}. \tag{A.9}$$

A similar set of definitions can be provided to other fixed-n variables on the boundary, including S_{nY_-}, $S_{nY_-}^{-1}$, Σ_{nY_-}, S_{nT_+}, $S_{nT_+}^{-1}$, Σ_{nT_+} , S_{nT_-} , $S_{nT_-}^{-1}$, Σ_{nT_-} and Σ_{nYT_-}.

Given the above fixed-n definitions and let $x=0$, it can be verified that, as $n_+\to\infty$,

$$S_{nY_+}\to\frac{f_{X_+}(x)}{\sigma_{\epsilon_{Y_+}}(x)}S_{Y_+}(c),\ \Sigma_{nY_+}\to f_{X_+}(x)\Sigma_{Y_+}(c)\text{ and }\Sigma_{nYT_+}\to f_{X_+}(x)\Sigma_{YT_+}(c).$$

S.2 Lemmas and proofs for Theorems 1 to 5

Lemma 1. Under Assumptions 1 to 6, as $n_+ \to \infty$, we have

$$\hat{\theta}_{n_+} + \frac{\sigma_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)} S_{Y_+}^{-1}(c) E(W_{n_+}^* | \mathbf{X}) \xrightarrow{L} MVN\left(\mathbf{0}, \frac{\sigma_{\epsilon_{Y_+}}^2(0)}{f_{X_+}(0)} S_{Y_+}^{-1}(c) \Sigma_{Y_+}(c) S_{Y_+}^{-1}(c)\right). \tag{A.10}$$

Proof of Lemma 1. See the proof of Theorem 2.1 in Kai *et al.* (2009). □

Lemma 2. Under Assumptions 1 to 6, as $n_+ \to \infty$, the asymptotic bias and variance for the LCQR estimator in eq. (1) are given by

$$\begin{aligned}
\text{Bias}(\hat{m}_{Y_+}(0)|\mathbf{X}) &= \frac{1}{2} a_{Y_+}(c) m_{Y_+}^{(2)}(0) h_{Y_+}^2 + o_p(h_{Y_+}^2), \\
\text{Var}(\hat{m}_{Y_+}(0)|\mathbf{X}) &= \frac{1}{n_+ h_{Y_+}} \frac{b_{Y_+}(c) \sigma_{\epsilon_{Y_+}}^2(0)}{f_{X_+}(0)} + o_p(\frac{1}{n_+ h_{Y_+}}), \\
a_{Y_+}(c) &= \frac{\mu_{+,2}^2(c) - \mu_{+,1}(c)\mu_{+,3}(c)}{\mu_{+,0}(c)\mu_{+,2}(c) - \mu_{+,1}^2(c)}, \\
b_{Y_+}(c) &= e_q^T (S_{Y_+}^{-1}(c) \Sigma_{Y_+}(c) S_{Y_+}^{-1}(c))_{11} e_q / q^2.
\end{aligned} \tag{A.11}$$

Proof of Lemma 2. The bias result follows that in Theorem 2.2 in Kai *et al.* (2009). The variance result also largely follows that in Kai *et al.* (2009). Given

$$\text{Var}(\hat{m}_{Y_+}(0)|\mathbf{X}) = \frac{1}{n_+ h_{Y_+}} \frac{\sigma_{\epsilon_{Y_+}}^2}{q^2 f_{X_+}(0)} e_q^T (S_{Y_+}^{-1}(c) \Sigma_{Y_+}(c) S_{Y_+}^{-1}(c))_{11} e_q + o_p(\frac{1}{n_+ h_{Y_+}}), \tag{A.12}$$

It is easy to verify that when $q = 1$, eq. (A.12) can be written as

$$\begin{aligned}
\text{Var}(\hat{m}_{Y_+}(0)|\mathbf{X}) = &\frac{1}{n_+ h_{Y_+}} \frac{\sigma_{\epsilon_{Y_+}}^2}{f_{X_+}(0)} \frac{\mu_{+,2}^2(c)\nu_{+,0}(c) - 2\mu_{+,1}(c)\mu_{+,2}(c)\nu_{+,1}(c) + \mu_{+,1}^2(c)\nu_{+,2}(c)}{(\mu_{+,0}(c)\mu_{+,2}(c) - \mu_{+,1}^2(c))^2} R_1(q) \\
&+ o_p(\frac{1}{n_+ h_{Y_+}}),
\end{aligned} \tag{A.13}$$

where $R_1(q) = \frac{1}{q^2} \sum_{k=1}^q \sum_{k'=1}^q \frac{\tau_{kk'}}{f_{\epsilon_{Y_+}(c_k)} f_{\epsilon_{Y_+}(c_{k'})}}$. However, for $q \geq 2$, the result in eq. (A.13) no longer holds and we use eq. (A.12) instead. □

Lemma 3. Under Assumptions 1 to 6, as $n_+ \to \infty$, the covariance between $\hat{m}_{Y_+}(x)$ and $\hat{m}_{T_+}(x)$ at the boundary point 0 is given by

$$\mathrm{Cov}(\hat{m}_{Y_+}(0),\hat{m}_{T_+}(0)|\mathbf{X}) = \frac{1}{n_+\sqrt{h_{Y_+}h_{T_+}}}\frac{\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{f_{X_+}(0)}b_{YT_+} + o_p(\frac{1}{n_+h_{Y_+}}+\frac{1}{n_+h_{T_+}}), \quad \text{(A.14)}$$

where

$$b_{YT_+} = \frac{1}{q^2}e_q^T\left(S_{Y_+}^{-1}(c)\Sigma_{YT_+}S_{T_+}^{-1}(c)\right)_{11}e_q. \quad \text{(A.15)}$$

Proof of Lemma 3. Assume $p=1$. From Lemma 1, we write

$$\begin{aligned}
&\hat{m}_{Y_+}(0) - E(\hat{m}_{Y_+}(0)|\mathbf{X}) \\
&= -\frac{1}{q\sqrt{n_+h_{Y_+}}}\frac{\sigma_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)}e_q^T\begin{pmatrix}(S_{Y_+}^{-1}(c))_{11} & (S_{Y_+}^{-1}(c))_{12}\end{pmatrix}\begin{pmatrix}w^*_{Y_+,1n} - E(w^*_{Y_+,1n}|\mathbf{X})\\ w^*_{Y_+,2n} - E(w^*_{Y_+,2n}|\mathbf{X})\end{pmatrix} + o_p(1) \\
&= -\frac{1}{q\sqrt{n_+h_{Y_+}}}\frac{\sigma_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)}e_q^T\begin{pmatrix}(S_{Y_+}^{-1}(c))_{11} & (S_{Y_+}^{-1}(c))_{12}\end{pmatrix}\begin{pmatrix}w_{Y_+,1n} - E(w_{Y_+,1n}|\mathbf{X})\\ w_{Y_+,2n} - E(w_{Y_+,2n}|\mathbf{X})\end{pmatrix} + o_p(1),
\end{aligned}$$

where the last equality follows by the result that $\mathrm{Var}(w^*_{Y_+,1n} - w_{Y_+,1n}|\mathbf{X}) = o_p(1)$ and $\mathrm{Var}(w^*_{Y_+,21} - w_{Y_+,21}|\mathbf{X}) = o_p(1)$. See Kai *et al.* (2010) for a proof. Similarly, we have

$$\begin{aligned}
&\hat{m}_{T_+}(0) - E(\hat{m}_{T_+}(0)|\mathbf{X}) \\
&= -\frac{1}{q\sqrt{n_+h_{T_+}}}\frac{\sigma_{\epsilon_{T_+}}(0)}{f_{X_+}(0)}e_q^T\begin{pmatrix}(S_{T_+}^{-1}(c))_{11} & (S_{T_+}^{-1}(c))_{12}\end{pmatrix}\begin{pmatrix}w_{T_+,1n} - E(w_{T_+,1n}|\mathbf{X})\\ w_{T_+,2n} - E(w_{T_+,2n}|\mathbf{X})\end{pmatrix} + o_p(1).
\end{aligned}$$

$$\begin{aligned}
&\mathrm{Cov}(\hat{m}_{Y_+}(0),\hat{m}_{T_+}(0)|\mathbf{X}) \\
&= E\left((\hat{m}_{Y_+}(0) - E(\hat{m}_{Y_+}(0)|\mathbf{X}))(\hat{m}_{T_+}(0) - E(\hat{m}_{T_+}(0)|\mathbf{X}))\right) \\
&= \frac{1}{q^2n_+\sqrt{h_{Y_+}h_{T_+}}}\frac{\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{f^2_{X_+}(0)}e_q^T\begin{pmatrix}(S_{Y_+}^{-1}(c))_{11} & (S_{Y_+}^{-1}(c))_{12}\end{pmatrix} \\
&\quad \times E\left[\begin{pmatrix}w_{Y_+,1n} - E(w_{Y_+,1n}|\mathbf{X})\\ w_{Y_+,2n} - E(w_{Y_+,2n}|\mathbf{X})\end{pmatrix}\begin{pmatrix}w_{T_+,1n} - E(w_{T_+,1n}|\mathbf{X})\\ w_{T_+,2n} - E(w_{T_+,2n}|\mathbf{X})\end{pmatrix}^T\right] \times \begin{pmatrix}(S_{T_+}^{-1}(c))_{11}\\ (S_{T_+}^{-1}(c))_{12}\end{pmatrix}e_q \\
&= \frac{1}{q^2n_+\sqrt{h_{Y_+}h_{T_+}}}\frac{\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{f_{X_+}(0)}e_q^T\left(S_{Y_+}^{-1}(c)\Sigma_{YT_+}S_{T_+}^{-1}(c)\right)_{11}e_q + o_p(\frac{1}{n_+h_{Y_+}}+\frac{1}{n_+h_{T_+}}),
\end{aligned}$$

where $\mathrm{Cov}(\eta_{Y_+,i,k},\eta_{T_+,j,k'}) = \phi_{kk'}$ if $i=j$, and $\mathrm{Cov}(\eta_{Y_+,i,k},\eta_{T_+,j,k'}) = 0$ if $i \neq j$. $\square$

Proof of Theorem 1. See Lemma 2 above. $\square$

Proof of Theorem 2. Consider the approximation

$$
\begin{aligned}
\hat{\tau}_{\text{fuzzy}} - \tau_{\text{fuzzy}} = & \frac{1}{m_{T_+}(0) - m_{T_-}(0)} \left[\hat{m}_{Y_+}(0) - m_{Y_+}(0) - (\hat{m}_{Y_-}(0) - m_{Y_-}(0))\right] \\
& - \frac{m_{Y_+}(0) - m_{Y_-}(0)}{\left[m_{T_+}(0) - m_{T_-}(0)\right]^2} \left[\hat{m}_{T_+}(0) - m_{T_+}(0) - (\hat{m}_{T_-}(0) - m_{T_-}(0))\right] \\
& + o_p(h_{Y_+}^2 + h_{Y_-}^2 + h_{T_+}^2 + h_{T_-}^2).
\end{aligned}
$$

The bias expression follows from Lemma 2: use it four times for $\hat{m}_{Y_+}(0)$, $\hat{m}_{Y_-}(0)$, $\hat{m}_{T_+}(0)$, and $\hat{m}_{T_-}(0)$. For the variance expression, note that the approximation above leads to

$$
\begin{aligned}
& \operatorname{Var}(\hat{\tau}_{\text{fuzzy}}) \\
= \;& \frac{\operatorname{Var}(\hat{m}_{Y_+}(0)) + \operatorname{Var}(\hat{m}_{Y_-}(0))}{\left(m_{T_+}(0) - m_{T_-}(0)\right)^2} + \frac{\left(m_{Y_+}(0) - m_{Y_-}(0)\right)^2}{\left(m_{T_+}(0) - m_{T_-}(0)\right)^4} \left[\operatorname{Var}(\hat{m}_{T_+}(0)) + \operatorname{Var}(\hat{m}_{T_-}(0))\right] \\
& -2 \frac{m_{Y_+}(0) - m_{Y_-}(0)}{\left(m_{T_+}(0) - m_{T_-}(0)\right)^3} \left[\operatorname{Cov}(\hat{m}_{Y_+}(0), \hat{m}_{T_+}(0)) + \operatorname{Cov}(\hat{m}_{Y_-}(0), \hat{m}_{T_-}(0))\right] + \text{s.o.} \quad \text{(A.16)}
\end{aligned}
$$

where $s.o.$ denotes a small order term.

Plugging the variance and covariance expressions in Lemmas 2 and 3 to (A.16) leads to the asymptotic variance expression of $\hat{\tau}_{\text{fuzzy}}$. $\square$

For convenience we write $\hat{m}_{Y_+}(0)$ and $\hat{m}_{Y_-}(0)$ as $\hat{m}_{Y_+}$ and $\hat{m}_{Y_-}$, respectively. Equation (30) suggests that we need the expressions for $\operatorname{Var}(\operatorname{Bias}(\hat{m}_{Y_+}))$ and $\operatorname{Cov}(\hat{m}_{Y_+}, \operatorname{Bias}(\hat{m}_{Y_+}))$ to adjust the variance. The next lemma provides results for computing $\operatorname{Var}(\operatorname{Bias}(\hat{m}_{Y_+}))$. In deriving the results, we also present the bias of $\operatorname{Bias}(\hat{m}_{Y_+})$. Let e_r be a $p \times 1$ unit vector with the r-th element equal to one. Let $p = 3$ in the following proof.

Lemma 4. Under Assumptions 1 to 6, as $n_+ \to \infty$, the asymptotic bias and variance of $\hat{m}_{Y_+}^{(2)}$ are given by

$$
\operatorname{Bias}(\hat{m}_{Y_+}^{(2)} | \mathbf{X}) = \frac{1}{12} a_{Y_+}^*(c) m_{Y_+}^{(4)} h_{Y_+}^2 + o_p(h_{Y_+}^2), \quad \text{(A.17)}
$$

$$
\operatorname{Var}\,(\hat{m}_{Y_+}^{(2)} | \mathbf{X}) = \frac{4}{n_+ h_{Y_+}^5} \frac{\sigma_{\epsilon_{Y_+}}^2(0) b_{Y_+}^*(c)}{f_{X_+}(0)} + o_p(\frac{1}{n_+ h_{Y_+}^5}), \quad \text{(A.18)}
$$

where

$$a^*_{Y_+}(c) = \mu_{+,4} e_2^T (S^{-1}_{Y_+}(c))_{21} f_{\epsilon_{Y_+}} + \sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k) e_2^T (S^{-1}_{Y_+}(c))_{22} (\mu_{+,5}, \mu_{+,6}, \mu_{+,7})^T \tag{A.19}$$

$$b^*_{Y_+}(c) = e_2^T (S^{-1}_{Y_+}(c) \Sigma_{Y_+}(c) S^{-1}_{Y_+})_{22} e_2. \tag{A.20}$$

Proof of Lemma 4. From the definition of v_j, we have

$$\hat{m}^{(2)}_{Y_+} = m^{(2)}_{Y_+} + \frac{2\hat{v}_2}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}}}. \tag{A.21}$$

Hence the bias becomes

$$\begin{aligned}
E(\hat{m}^{(2)}_{Y_+}) - m^{(2)}_{Y_+} &= -\frac{2\sigma_{\epsilon_{Y_+}}(0)}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}} f_{X_+}(0)} e_2^T ((S^{-1}_{Y_+}(c))_{21}, (S^{-1}_{Y_+}(c))_{22}) E(W^*_{Y_+,n}) \\
&= -\frac{2\sigma_{\epsilon_{Y_+}}(0)}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}} f_{X_+}(0)} e_2^T (S^{-1}_{Y_+}(c))_{21} E(W^*_{Y_+,1n}) - \frac{2\sigma_{\epsilon_{Y_+}}(0)}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}} f_{X_+}(0)} e_2^T (S^{-1}_{Y_+}(c))_{22} E(W^*_{Y_+,2n}) \\
&= \mathrm{I} + \mathrm{II}.
\end{aligned}$$

$$\begin{aligned}
\mathrm{I} &= -\frac{2\sigma_{\epsilon_{Y_+}}(0)}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}} f_{X_+}(0)} e_2^T (S^{-1}_{Y_+}(c))_{21} \\
&\quad \times \left[-\frac{f_{\epsilon_{Y_+}}}{\sqrt{n_+ h_{Y_+}}} \sum_{i=1}^{n_+} K_i c_k \frac{\sigma_{\epsilon_{Y_+,i}} - \sigma_{\epsilon_{Y_+}}(0)}{\sigma_{\epsilon_{Y_+,i}}} - \frac{f_{\epsilon_{Y_+}}}{\sqrt{n_+ h_{Y_+}}} \sum_{i=1}^{n_+} K_i \frac{r_{i,3}}{\sigma_{\epsilon_{Y_+,i}}} \right] \\
&= \frac{1}{12} m^{(4)}_{Y_+} \mu_{+,4}(c) e_2^T (S^{-1}_{Y_+}(c))_{21} f_{\epsilon_{Y_+}} h^2_{Y_+} + o_p(h^2_{Y_+}), \\
\mathrm{II} &= -\frac{2\sigma_{\epsilon_{Y_+}}(0)}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}} f_{X_+}(0)} e_2^T (S^{-1}_{Y_+}(c))_{22} \\
&\quad \times \left[-\frac{\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k)}{\sqrt{n_+ h_{Y_+}}} \sum_{i=1}^{n_+} K_i c_k \frac{\sigma_{\epsilon_{Y_+,i}} - \sigma_{\epsilon_{Y_+}}(0)}{\sigma_{\epsilon_{Y_+,i}}} \begin{pmatrix} x_{+,i} \\ x^2_{+,i} \\ x^3_{+,i} \end{pmatrix} - \frac{\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k)}{\sqrt{n_+ h_{Y_+}}} \sum_{i=1}^{n_+} K_i \frac{r_{i,3}}{\sigma_{\epsilon_{Y_+,i}}} \begin{pmatrix} x_{+,i} \\ x^2_{+,i} \\ x^3_{+,i} \end{pmatrix} \right] \\
&= \frac{1}{12} m^{(4)}_{Y_+} \sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k) e_2^T (S^{-1}_{Y_+}(c))_{22} \begin{pmatrix} \mu_{+,5} \\ \mu_{+,6} \\ \mu_{+,7} \end{pmatrix} h^2_{Y_+} + o_p(h^2_{Y_+}).
\end{aligned}$$

The bias result is proved by combining the two terms I and II. One would expect a number of $4! = 24$ instead of 12 on the denominator. This is due to the extra number 2 in eq. (A.21). Because of the way $\hat{v}_2$ is defined, the "effective" constant on the denominator is still 24, in line with the standard results for nonparametric derivatives. Similarly, the number 4 appearing on the numerator of the variance is also a result of the number 2 in eq. (A.21). The variance results from eq. (A.21) and Lemma 1. $\square$

Proof of Theorem 3. Following Theorem 1, we have

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_{Y_+}) &= \frac{1}{n_+ h_{Y_+}} \frac{b_{Y_+}(c)\sigma^2_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)} + o_p(\frac{1}{n_+ h_{Y_+}}), \\
\mathrm{Var}(\hat{m}_{Y_-}) &= \frac{1}{n_- h_{Y_-}} \frac{b_{Y_-}(c)\sigma^2_{\epsilon_{Y_-}}(0)}{f_{X_-}(0)} + o_p(\frac{1}{n_- h_{Y_-}}).
\end{aligned}
$$

Use the bias expression in Theorem 1 and the variance result in Lemma 4, we have

$$
\begin{aligned}
\mathrm{Var}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})) &= \frac{\sigma^2_{\epsilon_{Y_+}}(0)}{n_+ h_{Y_+} f_{X_+}(0)} a^2(c) b^*_{Y_+}(c) + o_p(\frac{1}{n_+ h_{Y_+}}), \\
\mathrm{Var}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_-})) &= \frac{\sigma^2_{\epsilon_{Y_-}}(0)}{n_- h_{Y_-} f_{X_-}(0)} a^2(c) b^*_{Y_-}(c) + o_p(\frac{1}{n_- h_{Y_-}}).
\end{aligned}
$$

For the covariances, we have

$$
\begin{aligned}
\mathrm{Cov}(\hat{m}_{Y_+}, \widehat{\mathrm{Bias}}(\hat{m}_{Y_+})) &= \mathrm{Cov}(m_{Y_+} + \frac{1}{q\sqrt{n_+ h_{Y_+}}} \sum_{k=1}^{q} \hat{u}_k, \frac{1}{2} a_{Y_+}(c) h^2_{Y_+} (m^{(2)}_{Y_+} + \frac{2\hat{v}_2}{h^2_{Y_+}\sqrt{n_+ h_{Y_+}}}) \\
&= \frac{a_{Y_+}(c)}{n_+ h_{Y_+} q} \sum_{k=1}^{q} \mathrm{Cov}(\hat{u}_k, \hat{v}_2) \\
&= \frac{a_{Y_+}(c)\sigma^2_{\epsilon_{Y_+}}(0)}{n_+ h_{Y_+} q f_{X_+}(0)} e_q^T (S^{-1}_{Y_+} \Sigma_{Y_+} S^{-1}_{Y_+})_{12,2} + o_p(\frac{1}{n_+ h_{Y_+}}),
\end{aligned}
$$

where $(S^{-1}_{Y_+} \Sigma_{Y_+} S^{-1}_{Y_+})_{12,2}$ is the second column of the matrix $(S^{-1}_{Y_+} \Sigma_{Y_+} S^{-1}_{Y_+})_{12}$ and the last line follows from Lemma 1. Similarly, for data below the cutoff, we have

$$
\mathrm{Cov}(\hat{m}_{Y_-}, \widehat{\mathrm{Bias}}(\hat{m}_{Y_-})) = \frac{a_{Y_-}(c)\sigma^2_{\epsilon_{Y_-}}(0)}{n_- h_{Y_-} q f_{X_-}(0)} e_q^T (S^{-1}_{Y_-} \Sigma_{Y_-} S^{-1}_{Y_-})_{12,2} + o_p(\frac{1}{n_- h_{Y_-}}).
$$

The expression for $\mathrm{Var}(\hat{\tau}_{\mathrm{sharp}} - \widehat{\mathrm{Bias}}(\hat{\tau}_{\mathrm{sharp}}))$ is obtained by substituting the six variance and covariance results into eq. (30),

$$\text{Var}(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})) = \frac{1}{n_+ h_{Y_+}} V_{\text{sharp},+} + \frac{1}{n_- h_{Y_-}} V_{\text{sharp},-},$$

where

$$V_{\text{sharp},+} = \frac{b_{Y_+}(c)\sigma^2_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)} + \frac{\sigma^2_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)} a^2(c) b^*_{Y_+}(c) - 2\frac{a_{Y_+}(c)\sigma^2_{\epsilon_{Y_+}}(0)}{q f_{X_+}(0)} e_q^T (S^{-1}_{Y_+} \Sigma_{Y_+} S^{-1}_{Y_+})_{12,2},$$
$$V_{\text{sharp},-} = \frac{b_{Y_-}(c)\sigma^2_{\epsilon_{Y_-}}(0)}{f_{X_-}(0)} + \frac{\sigma^2_{\epsilon_{Y_-}}(0)}{f_{X_-}(0)} a^2(c) b^*_{Y_-}(c) - 2\frac{a_{Y_-}(c)\sigma^2_{\epsilon_{Y_-}}(0)}{q f_{X_-}(0)} e_q^T (S^{-1}_{Y_-} \Sigma_{Y_-} S^{-1}_{Y_-})_{12,2}.$$

Next, we establish the asymptotic normality of the adjusted t-statistic. From Lemma 2, we have

$$\begin{aligned}
\frac{\hat{m}_{Y_+} - \frac{1}{2} a_{Y_+}(c) m^{(2)}_{Y_+} h^2_{Y_+} - m_{Y_+}}{\sqrt{\text{Var}(\hat{m}_{Y_+})}} &= \frac{\hat{m}_{Y_+} - E(\hat{m}_{Y_+})}{\sqrt{\text{Var}(\hat{m}_{Y_+})}} + \frac{E(\hat{m}_{Y_+}) - m_{Y_+} - \frac{1}{2} a_{Y_+}(c) m^{(2)}_{Y_+} h^2_{Y_+}}{\sqrt{\text{Var}(\hat{m}_{Y_+})}} \\
&= \frac{\hat{m}_{Y_+} - E(\hat{m}_{Y_+})}{\sqrt{\text{Var}(\hat{m}_{Y_+})}} + \frac{O_p(h^3_{Y_+})}{O_p(\sqrt{1/n_+ h_{Y_+}})}. && \text{(A.22)} \\
&\xrightarrow{d} N(0,1). && \text{(A.23)}
\end{aligned}$$

The second term in eq. (A.22) converges to 0 under Assumption 6. In the first term, given the definition of u_k in eq. (A.4) and since $\hat{m}_{Y_+}$ is a linear function of $\hat{u}_k$ in eq. (14), Lemma 1 and the Delta method lead to the normality result in eq. (A.23).

Similarly, we have

$$\frac{\hat{m}_{Y_-} - \frac{1}{2} a_{Y_-}(c) m^{(2)}_{Y_-} h^2_{Y_-} - m_{Y_-}}{\sqrt{\text{Var}(\hat{m}_{Y_-})}} \xrightarrow{d} N(0,1). \tag{A.24}$$

Let $\tau_0 = m_{Y_+} - m_{Y_-}$. Using the proof for eqs. (A.23) and (A.24), we can show

$$\frac{\hat{\tau}_{\text{sharp}} - \left[\frac{1}{2} a_{Y_+}(c) m^{(2)}_{Y_+} h^2_{Y_+} - \frac{1}{2} a_{Y_-}(c) m^{(2)}_{Y_-} h^2_{Y_-}\right] - \tau_0}{\sqrt{\text{Var}(\hat{\tau}_{\text{sharp}})}} \xrightarrow{d} N(0,1). \tag{A.25}$$

Finally, we have

$$
\begin{aligned}
\frac{\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}}) - \tau_0}{\sqrt{\text{Var}(\hat{\tau}^{\text{bc}}_{\text{sharp}})}} &= \frac{\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}}) - E(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}}))}{\sqrt{\text{Var}(\hat{\tau}^{\text{bc}}_{\text{sharp}})}} \\
&\quad + \frac{E(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})) - \tau_0}{\sqrt{\text{Var}(\hat{\tau}^{\text{bc}}_{\text{sharp}})}} \\
&= \frac{\hat{\tau}^{\text{bc}}_{\text{sharp}} - E(\hat{\tau}^{\text{bc}}_{\text{sharp}})}{\sqrt{\text{Var}(\hat{\tau}^{\text{bc}}_{\text{sharp}})}} + O_p(\sqrt{n_+ h_+^7}) + O_p(\sqrt{n_- h_-^7}) \\
&\xrightarrow{d} N(0,1), \qquad \text{(A.26)}
\end{aligned}
$$

where we use the proof similar to eqs. (A.23) to (A.25) and the fact that $E(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})) - \tau_0 = O_p(h^3_{Y_+}) + O_p(h^3_{Y_-})$.

□

Proof of Theorem 4. We first note that all bias terms in eq. (34) can be obtained using Lemma 2. For terms in the adjusted variance in eq. (35), $\text{Var}(\hat{m}_{Y_+})$, $\text{Var}(\hat{m}_{T_+})$, $\text{Var}(\widehat{\text{Bias}}(\hat{m}_{Y_+}))$, $\text{Var}(\widehat{\text{Bias}}(\hat{m}_{T_+}))$, $\text{Cov}(\hat{m}_{Y_+}, \widehat{\text{Bias}}(\hat{m}_{Y_+}))$, and $\text{Cov}(\hat{m}_{T_+}, \widehat{\text{Bias}}(\hat{m}_{T_+}))$ can be obtained using results in the proof of Theorem 3; $\text{Cov}(\hat{m}_{Y_+}, \hat{m}_{T_+})$ is obtained using Lemma 3. And we list these seven terms in the following.

$$
\begin{aligned}
\text{Var}(\hat{m}_{Y_+}) &= \frac{1}{n_+ h_{Y_+}} \frac{b_{Y_+}(c)\sigma^2_{\epsilon_{Y_+}}(0)}{f_{X_+}(0)} + o_p(\frac{1}{n_+ h_{Y_+}}), \\
\text{Var}(\hat{m}_{T_+}) &= \frac{1}{n_+ h_{T_+}} \frac{b_{T_+}(c)\sigma^2_{\epsilon_{T_+}}(0)}{f_{X_+}(0)} + o_p(\frac{1}{n_+ h_{T_+}}), \\
\text{Var}(\widehat{\text{Bias}}(\hat{m}_{Y_+})) &= \frac{\sigma^2_{\epsilon_{Y_+}}(0)}{n_+ h_{Y_+} f_{X_+}(0)} a^2_{Y_+}(c) b^*_{Y_+}(c) + o_p(\frac{1}{n_+ h_{Y_+}}), \\
\text{Var}(\widehat{\text{Bias}}(\hat{m}_{T_+})) &= \frac{\sigma^2_{\epsilon_{T_+}}(0)}{n_+ h_{T_+} f_{X_+}(0)} a^2_{T_+}(c) b^*_{T_+}(c) + o_p(\frac{1}{n_+ h_{T_+}}),
\end{aligned}
$$

$$\mathrm{Cov}(\hat{m}_{Y_+},\hat{m}_{T_+}) = \frac{1}{n_+\sqrt{h_{Y_+}h_{T_+}}}\frac{\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{f_{X_+}(0)}b_{YT_+} + o_p(\frac{1}{n_+h_{Y_+}}+\frac{1}{n_+h_{T_+}}),$$
$$\mathrm{Cov}(\hat{m}_{Y_+},\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})) = \frac{a_{Y_+}(c)\sigma^2_{\epsilon_{Y_+}}(0)}{n_+h_{Y_+}qf_{X_+}(0)}e_q^T(S_{Y_+}^{-1}\Sigma_{Y_+}S_{Y_+}^{-1})_{12,2} + o_p(\frac{1}{n_+h_{Y_+}}),$$
$$\mathrm{Cov}(\hat{m}_{T_+},\widehat{\mathrm{Bias}}(\hat{m}_{T_+})) = \frac{a_{T_+}(c)\sigma^2_{\epsilon_{T_+}}(0)}{n_+h_{T_+}qf_{X_+}(0)}e_q^T(S_{T_+}^{-1}\Sigma_{T_+}S_{T_+}^{-1})_{12,2} + o_p(\frac{1}{n_+h_{T_+}}).$$

Next, we compute the remaining three covariances.

$$\begin{aligned}
\mathrm{Cov}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+}),\widehat{\mathrm{Bias}}(\hat{m}_{T_+})) &= \mathrm{Cov}(\frac{1}{2}a_{Y_+}(c)\hat{m}^{(2)}_{Y_+}h^2_{Y_+},\frac{1}{2}a_{T_+}(c)\hat{m}^{(2)}_{T_+}h^2_{T_+})\\
&= \frac{a_{Y_+}(c)a_{T_+}(c)}{\sqrt{n_+h_{Y_+}}\sqrt{n_+h_{T_+}}}\mathrm{Cov}(\hat{v}_{2,Y_+},\hat{v}_{2,T_+})\\
&= \frac{a_{Y_+}(c)a_{T_+}(c)\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{n_+\sqrt{h_{Y_+}h_{T_+}}f_{X_+}(0)}e_2^T(S_{Y_+}^{-1}\Sigma_{YT_+}S_{T_+}^{-1})_{22}e_2\\
&\quad + o_p(\frac{1}{n_+h_{Y_+}}+\frac{1}{n_+h_{T_+}}).
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{m}_{Y_+},\widehat{\mathrm{Bias}}(\hat{m}_{T_+})) &= \mathrm{Cov}(m_{Y_+}+\frac{1}{q\sqrt{n_+h_{Y_+}}}\sum_{k=1}^{q}\hat{u}_{k,Y},\frac{1}{2}a_{T_+}(c)h^2_{T_+}(m^{(2)}_{T_+}+\frac{2\hat{v}_{2,T}}{h^2_{T_+}\sqrt{n_+h_{T_+}}}))\\
&= \frac{a_{T_+}(c)\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{qn_+\sqrt{h_{Y_+}h_{T_+}}f_{X_+}(0)}e_q^T(S_{Y_+}^{-1}\Sigma_{YT_+}S_{T_+}^{-1})_{12,2} + o_p(\frac{1}{n_+h_{Y_+}}+\frac{1}{n_+h_{T_+}}).
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{m}_{T_+},\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})) &= \mathrm{Cov}(m_{T_+}+\frac{1}{q\sqrt{n_+h_{T_+}}}\sum_{k=1}^{q}\hat{u}_{k,T},\frac{1}{2}a_{Y_+}(c)h^2_{Y_+}(m^{(2)}_{Y_+}+\frac{2\hat{v}_{2,Y}}{h^2_{Y_+}\sqrt{n_+h_{Y_+}}}))\\
&= \frac{a_{Y_+}(c)\sigma_{\epsilon_{Y_+}}(0)\sigma_{\epsilon_{T_+}}(0)}{qn_+\sqrt{h_{Y_+}h_{T_+}}f_{X_+}(0)}e_q^T(S_{T_+}^{-1}\Sigma_{TY_+}S_{Y_+}^{-1})_{12,2} + o_p(\frac{1}{n_+h_{Y_+}}+\frac{1}{n_+h_{T_+}}).
\end{aligned}$$

Substituting the above results into eq. (35) gives the expression for $\mathrm{Var}((\hat{m}_{Y_+}-\tau_0\hat{m}_{T_+})-(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})-\tau_0\widehat{\mathrm{Bias}}(\hat{m}_{T_+})))$. The result for $\mathrm{Var}((\hat{m}_{Y_-}-\tau_0\hat{m}_{T_-})-(\widehat{\mathrm{Bias}}(\hat{m}_{Y_-})-\tau_0\widehat{\mathrm{Bias}}(\hat{m}_{T_-})))$ can be obtained in a similar way. Adding up the two variance results gives the adjusted variance in the fuzzy case.

To establish the asymptotic normality, note that we can use eq. (34) to write $\tilde{\tau}^{\mathrm{bc}}_{\mathrm{fuzzy}}$ as

$$\tilde{\tau}^{\mathrm{bc}}_{\mathrm{fuzzy}} = (\hat{m}_{Y_+}-\widehat{\mathrm{Bias}}(\hat{m}_{Y_+}))-\tau_0(\hat{m}_{T_+}-\widehat{\mathrm{Bias}}(\hat{m}_{T_+}))-(\hat{m}_{Y_-}-\widehat{\mathrm{Bias}}(\hat{m}_{Y_-}))+\tau_0(\hat{m}_{T_-}-\widehat{\mathrm{Bias}}(\hat{m}_{T_-})).$$

Using the similar argument in proving the asymptotic normality of $\hat{\tau}^{\text{bc}}_{\text{sharp}}$, we can establish the asymptotic distribution of $t^{\text{adj.}}_{\text{fuzzy}}$. $\square$

Proof of Theorem 5. We first expand $\text{Bias}(\hat{m}_{Y_+})$ up to $O(h^3_{Y_+})$ on the boundary. Recall $\hat{m}_{Y_+} = \sum_{k=1}^q \hat{a}_k/q$ and we have

$$
\begin{aligned}
\text{Bias}(\hat{m}_{Y_+}) &= \frac{\sigma_{\epsilon_{Y_+}}(0)}{q}\sum_{k=1}^q c_k - \frac{\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T\left[(S^{-1}_{Y_+}(c))_{11}E(w^*_{Y_+,1n}) + (S^{-1}_{Y_+}(c))_{12}E(w^*_{Y_+,2n})\right] \\
&= -\frac{\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T(S^{-1}_{Y_+}(c))_{11}E(w^*_{Y_+,1n}) - \frac{\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T(S^{-1}_{Y_+}(c))_{12}E(w^*_{Y_+,2n}) \\
&= \text{I} + \text{II}.
\end{aligned}
$$

Consider term I.

$$
\begin{aligned}
\text{I} &= \frac{-\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T(S^{-1}_{Y_+}(c))_{11}\begin{pmatrix} \frac{1}{\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}K_iE(\eta^*_{Y_+,i,1}) \\ \vdots \\ \frac{1}{\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}K_iE(\eta^*_{Y_+,i,q}) \end{pmatrix} \\
&= \frac{\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T(S^{-1}_{Y_+}(c))_{11}\begin{pmatrix} \frac{f_{\epsilon_{Y_+}}(c_1)}{\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}K_i\frac{d_{i,1}}{\sigma_{\epsilon_{Y_+,i}}} \\ \vdots \\ \frac{f_{\epsilon_{Y_+}}(c_q)}{\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}K_i\frac{d_{i,q}}{\sigma_{\epsilon_{Y_+,i}}} \end{pmatrix} + o_p(1) \\
&= \frac{1}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T(S^{-1}_{Y_+}(c))_{11}\begin{pmatrix} \frac{f_{\epsilon_{Y_+}}(c_1)}{\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}K_ir_{i,1} \\ \vdots \\ \frac{f_{\epsilon_{Y_+}}(c_q)}{\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}K_ir_{i,1} \end{pmatrix} + o_p(1) \\
&= \frac{1}{2q}e_q^T(S^{-1}_{Y_+}(c))_{11}f_{\epsilon_{Y_+}}m^{(2)}_{Y_+}\mu_{+,2}h^2_{Y_+} + \frac{1}{6q}e_q^T(S^{-1}_{Y_+}(c))_{11}f_{\epsilon_{Y_+}}m^{(3)}_{Y_+}\mu_{+,3}h^3_{Y_+} \\
&\quad + \frac{f^{(1)}_{X_+}(0)}{2qf_{X_+}(0)}e_q^T(S^{-1}_{Y_+}(c))_{11}f_{\epsilon_{Y_+}}m^{(2)}_{Y_+}\mu_{+,3}h^3_{Y_+} + o_p(h^3_{Y_+}),
\end{aligned}
$$

where the second equality follows by expanding the cumulative distribution of $\epsilon_{Y_+,i}$ around c_k and the third equality follows by noticing that all terms containing c_k, after multiplied by the coefficient $\frac{\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)}e_q^T(S^{-1}_{Y_+}(c))_{11}$, become zero after a summation. The last equality is obtained by a Taylor series expansion of m_{Y_+} at 0 up to order 3 in $r_{i,1}$, similar to the expansion in the definition of $r_{i,3}$.

Consider term II. Note that $p=1$ in the following proof when we estimate the conditional mean using degree one local polynomial.

$$
\begin{aligned}
\mathrm{II} &= \frac{-\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T (S_{Y_+}^{-1}(c))_{12} \begin{pmatrix} \frac{1}{\sqrt{n_+h_{Y_+}}}\sum_{k=1}^{q}\sum_{i=1}^{n_+} K_i X_{+,i} E(\eta^*_{Y_+,i,1}) \\ \vdots \\ \frac{1}{\sqrt{n_+h_{Y_+}}}\sum_{k=1}^{q}\sum_{i=1}^{n_+} K_i X^p_{+,i} E(\eta^*_{Y_+,i,q}) \end{pmatrix} \\
&= \frac{\sigma_{\epsilon_{Y_+}}(0)}{q\sqrt{n_+h_{Y_+}}f_{X_+}(0)} e_q^T (S_{Y_+}^{-1}(c))_{12} \begin{pmatrix} \frac{1}{\sqrt{n_+h_{Y_+}}}\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_1)\sum_{i=1}^{n_+} K_i X_{+,i} \frac{d_{i,1}}{\sigma_{\epsilon_{Y_+},i}} \\ \vdots \\ \frac{1}{\sqrt{n_+h_{Y_+}}}\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_q)\sum_{i=1}^{n_+} K_i X^p_{+,i} \frac{d_{i,q}}{\sigma_{\epsilon_{Y_+},i}} \end{pmatrix} + o_p(1) \\
&= \frac{\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k)}{2q} e_q^T (S_{Y_+}^{-1}(c))_{12} \begin{pmatrix} \mu_{+,3} \\ \vdots \\ \mu_{+,p+2} \end{pmatrix} m^{(2)}_{Y_+} h^2_{Y_+} \\
&\quad + \frac{\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k)}{6q} e_q^T (S_{Y_+}^{-1}(c))_{12} \begin{pmatrix} \mu_{+,4} \\ \vdots \\ \mu_{+,p+3} \end{pmatrix} m^{(3)}_{Y_+} h^3_{Y_+} \\
&\quad + \frac{f^{(1)}_{X_+}(0)\sum_{k=1}^{q} f_{\epsilon_{Y_+}}(c_k)}{2q f_{X_+}(0)} e_q^T (S_{Y_+}^{-1}(c))_{12} \begin{pmatrix} \mu_{+,4} \\ \vdots \\ \mu_{+,p+3} \end{pmatrix} m^{(2)}_{Y_+} h^3_{Y_+} + o_p(h^3_{Y_+}).
\end{aligned}
$$

Combining I and II yields

$$
\mathrm{Bias}(\hat{m}_{Y_+}) = \frac{1}{2} a_{Y_+}(c) m^{(2)}_{Y_+} h^2_{Y_+} + \frac{1}{6}\check{a}_{Y_+}(c) m^{(3)}_{Y_+} h^3_{Y_+} + \frac{1}{2}\frac{\tilde{a}_{Y_+}(c) f^{(1)}_{X_+}(0)}{f_{X_+}(0)} m^{(2)}_{Y_+} h^3_{Y_+} + o_p(h^3_{Y_+}),
$$

where $a_{Y_+}(c)$ is in Lemma 2, $\check{a}_{Y_+}(c) = \frac{\mu_{+,2}(c)\mu_{+,3}(c) - \mu_{+,1}(c)\mu_{+,4}(c)}{\mu_{+,0}(c)\mu_{+,2}(c) - \mu^2_{+,1}(c)}$, and $\tilde{a}_{Y_+}(c) = \frac{\mu^2_{+,2}(c) - \mu_{+,1}(c)\mu_{+,4}(c)}{\mu_{+,0}(c)\mu_{+,2}(c) - \mu^2_{+,1}(c)}$. Hence the leading term in $\mathrm{Bias}(\hat{m}_{Y_+} - \mathrm{Bias}(\hat{m}_{Y_+}))$ is $\frac{1}{6}\check{a}_{Y_+}(c) m^{(3)}_{Y_+} h^3_{Y_+} + \frac{1}{2}\frac{\tilde{a}_{Y_+}(c) f^{(1)}_{X_+}(0)}{f_{X_+}(0)} m^{(2)}_{Y_+} h^3_{Y_+}$.

Since we work with data above the cutoff in this proof, the adjusted variance is given by $\frac{1}{n_+h_{Y_+}}V_{\text{sharp}}^{\text{adj.}}$ in the proof of Theorem 3. Thus, the adjusted MSE can be written as

$$\begin{aligned}
\text{adj. MSE} &= \left[\frac{1}{6}\breve{a}_{Y_+}(c)m_{Y_+}^{(3)}h_{Y_+}^3 + \frac{1}{2}\frac{\tilde{a}_{Y_+}(c)f_{X_+}^{(1)}(0)}{f_{X_+}(0)}m_{Y_+}^{(2)}h_{Y_+}^3\right]^2 + \frac{1}{n_+h_{Y_+}}V_{\text{sharp}}^{\text{adj.}} + o_p(h_{Y_+}^6 + \frac{1}{n_+h_{Y_+}}) \\
&= C_2^2h_{Y_+}^6 + \frac{1}{n_+h_{Y_+}}C_3 + o_p(h_{Y_+}^6 + \frac{1}{n_+h_{Y_+}}),
\end{aligned}$$

where $C_2 = \frac{1}{6}\breve{a}_{Y_+}(c)m_{Y_+}^{(3)} + \frac{1}{2}\frac{\tilde{a}_{Y_+}(c)f_{X_+}^{(1)}(0)}{f_{X_+}(0)}m_{Y_+}^{(2)}$ and $C_3 = V_{\text{sharp}}^{\text{adj.}}$. The bandwidth that minimizes the adjusted MSE is given by $h = \left(\frac{C_3}{6C_2^2}\right)^{1/7} n_+^{-1/7}$. □

S.3 Lemmas and propositions for fixed-n results

This section first collects several lemmas for the development of fixed-n approximations. We then present two propositions that are the fixed-n counterparts of Theorems 3 and 4. Assume $p = 1$ in the following lemmas.

Lemma 5. Under Assumptions 1 to 6, the fixed-n bias and variance are given by

$$\begin{aligned}
\text{Bias}(\hat{m}_{Y_+}|\mathbf{X})_{\text{fixed-n}} &= \frac{1}{q}e_q^T\Bigg[(S_{nY_+}^{-1})_{11}f_{\epsilon_{Y_+}}\frac{1}{2n_+h_{Y_+}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^2}{\sigma_{\epsilon_{Y_+,i}}} \\
&\quad + (S_{nY_+}^{-1})_{12}\sum_{k=1}^{q}f_{\epsilon_{Y_+}}(c_k)\frac{1}{2n_+h_{Y_+}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^3}{\sigma_{\epsilon_{Y_+,i}}}\Bigg]m_{Y_+}^{(2)}h_{Y_+}^2 + o_p(h_{Y_+}^2), \\
\text{Var}(\hat{m}_{Y_+}|\mathbf{X})_{\text{fixed-n}} &= \frac{1}{n_+h_{Y_+}q^2}e_q^T\left(S_{nY_+}^{-1}\Sigma_{nY_+}S_{nY_+}^{-1}\right)e_q + o_p(\frac{1}{n_+h_{Y_+}}).
\end{aligned}$$

Proof of Lemma 5. We first state some results for $E(w_{Y_+,1n}^*)$ and $E(w_{Y_+,2n}^*)$ that are used in the proof of the asymptotic results in Theorem 5 in Section S.2. $E(w_{Y_+,1n}^*)$ is a $q\times 1$ vector while $E(w_{Y_+,2n}^*)$ is a $p\times 1$ vector with $p = 1$ in this case. By not letting $n\to\infty$, we have

$$\begin{aligned}
E(w_{Y_+,1n}^*) &= -f_{\epsilon_{Y_+}}\frac{1}{2\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^2}{\sigma_{\epsilon_{Y_+,i}}}m_{Y_+}^{(2)}(0)h_{Y_+}^2 + o(1), \\
E(w_{Y_+,2n}^*) &= -\sum_{k=1}^{q}f_{\epsilon_{Y_+}}(c_k)\frac{1}{2\sqrt{n_+h_{Y_+}}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^3}{\sigma_{\epsilon_{Y_+,i}}}m_{Y_+}^{(2)}(0)h_{Y_+}^2 + o(1).
\end{aligned} \tag{A.27}$$

From the proof of Theorem 5, Lemmas 2 and 3 in Kai *et al.* (2010), the loss function becomes

$$L_{n_+}(\theta) = \frac{1}{2}\theta^T S_{nY_+}\theta + W^*_{Y_+,n_+} + o_p(1),$$

the solution of which is,

$$\hat{\theta}_{n_+} = -S^{-1}_{nY_+} W^*_{Y_+,n_+} + o_p(1),$$

Rewrite the above equation as

$$\hat{\theta}_{n_+} + S^{-1}_{nY_+} E(W^*_{Y_+,n_+}|\mathbf{X}) = -S^{-1}_{nY_+}\left[W^*_{Y_+,n_+} - E(W^*_{Y_+,n_+}|\mathbf{X})\right] + o_p(1), \tag{A.28}$$

which is the base to prove Lemma 1 in the asymptotic case. Combining eq. (14) and eq. (A.28), we obtain the following expression for pre-asymptotic bias

$$\begin{aligned}\hat{m}_{Y_+} - m_{Y_+} &= \frac{1}{q\sqrt{n_+ h_{Y_+}}}\sum_{k=1}^{q}\hat{u}_k && \text{(A.29)}\\ &= -\frac{1}{q\sqrt{n_+ h_{Y_+}}} e_q^T \begin{bmatrix}(S^{-1}_{nY_+})_{11} & (S^{-1}_{nY_+})_{12}\end{bmatrix} W^*_{Y_+,n_+}. && \text{(A.30)}\end{aligned}$$

Plug the result in eq. (A.27) into eq. (A.30), and we prove the fixed-n bias result. From eq. (A.28), the variance of $\hat{\theta}_{n_+}$ becomes

$$\begin{aligned}\mathrm{Var}(\hat{\theta}_{n_+}) &= S^{-1}_{nY_+}\mathrm{Var}(W^*_{Y_+,n_+} - E(W^*_{Y_+,n_+}|\mathbf{X}))S^{-1}_{nY_+}\\ &\to S^{-1}_{nY_+}\mathrm{Var}(W_{Y_+,n_+} - E(W_{Y_+,n_+}|\mathbf{X}))S^{-1}_{nY_+}\\ &= S^{-1}_{nY_+}\mathrm{Var}(W_{Y_+,n_+})S^{-1}_{nY_+},\end{aligned}$$

where we use the result $\mathrm{Var}(W^*_{Y_+,n_+} - W_{Y_+,n_+}|\mathbf{X}) = o_p(1)$ from the proof of Theorem 5 in Kai *et al.* (2010). Similar to the proof of Lemma 3, we can show $\mathrm{Var}(W_{Y_+,n_+}) = \Sigma_{nY_+}$, which, together with eq. (14), proves the variance result in this lemma. □

Lemma 6. Under Assumptions 1 to 6, the fixed-n covariance between $\hat{m}_{Y_+}(0)$ and $\hat{m}_{T_+}(0)$ is given by

$$\mathrm{Cov}(\hat{m}_{Y_+}, \hat{m}_{T_+}|\mathbf{X})_{\text{fixed-n}} = \frac{1}{q^2 n_+\sqrt{h_{Y_+}h_{T_+}}} e_q^T \begin{bmatrix}(S^{-1}_{nY_+})_{11} & (S^{-1}_{nY_+})_{12}\end{bmatrix} \Sigma_{nYT_+} \begin{bmatrix}(S^{-1}_{nT_+})_{11} & (S^{-1}_{nT_+})_{12}\end{bmatrix}^T e_q.$$

Proof of Lemma 6. Similar to eq. (A.30), we have

$$\hat{m}_{T_+} - m_{T_+} = -\frac{1}{q\sqrt{n_+ h_{T_+}}} e_q^T \left[(S_{nT_+}^{-1})_{11} \quad (S_{nT_+}^{-1})_{12} \right] W_{T_+,n_+}^*. \tag{A.31}$$

Using the proof similar to that in Lemma 3 and the result

$$\Sigma_{nYT_+} = E\left[\begin{pmatrix} w_{Y_+,1n}^* - E(w_{Y_+,1n}^* | \mathbf{X}) \\ w_{Y_+,2n}^* - E(w_{Y_+,2n}^* | \mathbf{X}) \end{pmatrix} \begin{pmatrix} w_{T_+,1n}^* - E(w_{T_+,1n}^* | \mathbf{X}) \\ w_{T_+,2n}^* - E(w_{T_+,2n}^* | \mathbf{X}) \end{pmatrix}^T \right],$$

we have that Lemma 6 holds. □

Lemma 7. Under Assumptions 1 to 6, we have

$$\begin{aligned} \text{Bias}(\hat{\tau}_{\text{sharp}} | \mathbf{X})_{\text{fixed-n}} &= \text{Bias}(\hat{m}_{Y_+})_{\text{fixed-n}} - \text{Bias}(\hat{m}_{Y_-})_{\text{fixed-n}} + o_p(h_{Y_+}^2 + h_{Y_-}^2), \\ \text{Var}(\hat{\tau}_{\text{sharp}} | \mathbf{X})_{\text{fixed-n}} &= \text{Var}(\hat{m}_{Y_+})_{\text{fixed-n}} + \text{Var}(\hat{m}_{Y_-})_{\text{fixed-n}} + o_p(\frac{1}{n_+ h_{Y_+}} + \frac{1}{n_- h_{Y_-}}), \end{aligned}$$

where $\text{Bias}(\hat{m}_{Y_+})_{\text{fixed-n}}$ and $\text{Var}(\hat{m}_{Y_+})_{\text{fixed-n}}$ are given in Lemma 5, and $\text{Bias}(\hat{m}_{Y_-})_{\text{fixed-n}}$ and $\text{Var}(\hat{m}_{Y_-})_{\text{fixed-n}}$ are defined analogously.

Proof of Lemma 7. The results hold by applying Lemma 5 to eq. (3). □

Lemma 8. Under Assumptions 1 to 6, we have

$$\begin{aligned} \text{Bias}(\hat{\tau}_{\text{fuzzy}} | \mathbf{X})_{\text{fixed-n}} = &\frac{1}{m_{T_+} - m_{T_-}} \left[\text{Bias}(\hat{m}_{Y_+})_{\text{fixed-n}} - \text{Bias}(\hat{m}_{Y_-})_{\text{fixed-n}} \right] \\ &- \frac{m_{Y_+} - m_{Y_-}}{\left[m_{T_+} - m_{T_-} \right]^2} \left[\text{Bias}(\hat{m}_{T_+})_{\text{fixed-n}} - \text{Bias}(\hat{m}_{T_-})_{\text{fixed-n}} \right] \\ &+ o_p(h_{Y_+}^2 + h_{Y_-}^2 + h_{T_+}^2 + h_{T_-}^2). \end{aligned} \tag{A.32}$$

The variance expression is given in (A.16) by substituting the results for $\text{Var}(\hat{m}_{Y_+})$, $\text{Var}(\hat{m}_{Y_-})$, $\text{Var}(\hat{m}_{T_+})$, $\text{Var}(\hat{m}_{T_-})$, $\text{Cov}(\hat{m}_{Y_+}, \hat{m}_{T_+})$ and $\text{Cov}(\hat{m}_{Y_-}, \hat{m}_{T_-})$ in Lemmas 5 and 6.

Proof of Lemma 8. The proof follows from (A.16) and Lemmas 5 and 6. □

Let $p = 2$ in the following lemma.

Lemma 9. Under Assumptions 1 to 6, the fixed-n variance of $\hat{m}^{(2)}_{Y_+}$ is given by

$$\text{Var }(\hat{m}^{(2)}_{Y_+}|\mathbf{X}) = \frac{4}{n_+ h^5_{Y_+}} e_2^T (S^{-1}_{nY_+} \Sigma_{nY_+} S^{-1}_{nY_+})_{22} e_2 + o_p(\frac{1}{n_+ h^5_{Y_+}}).$$

Proof of Lemma 9. It reslts from combining $\text{Var}(\hat{v}_2) = e_2^T (S^{-1}_{nY_+} \Sigma_{nY_+} S^{-1}_{nY_+})_{22} e_2$ and eq. (A.21). □

Proposition 1. Under Assumptions 1 to 6, the fixed-n adjusted t-statistic for the sharp RD is given by

$$t^{\text{adj.}}_{\text{sharp, fixed-n}} = \frac{\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}} - \tau_0}{\sqrt{\text{Var}(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}})_{\text{fixed-n}}}}, \tag{A.33}$$

where the expression for fixed-n terms, $\widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}}$, $\text{Var}(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}})_{\text{fixed-n}}$, are given in the proof of this proposition.

Proof of Proposition 1. The fixed-n bias term on the numerator of eq. (A.33) is given in Lemma 7. For the denominator of eq. (A.33), recall

$$\begin{aligned}\text{Var}(\hat{\tau}_{\text{sharp}} - \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}})_{\text{fixed-n}} &= \text{Var}(\hat{\tau}_{\text{sharp}}) + \text{Var}(\widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}}) \\ &\quad - 2\text{Cov}(\hat{\tau}_{\text{sharp}}, \widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}}) \\ &= \text{I} + \text{II} + \text{III}. \end{aligned} \tag{A.34}$$

Term I is given in Lemma 5. Consider the second term II.

$$\text{Var}(\widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}}) = \text{Var}(\widehat{\text{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}) + \text{Var}(\widehat{\text{Bias}}(\hat{m}_{Y_-})_{\text{fixed-n}}). \tag{A.35}$$

Using Lemma 5 and omitting the small-order terms, we have

$$\widehat{\text{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}} = D_{nY_+,1} \hat{m}^{(2)}_{Y_+} h^2_{Y_+}, \tag{A.36}$$

$$\widehat{\text{Bias}}(\hat{m}_{Y_-})_{\text{fixed-n}} = D_{nY_-,1} \hat{m}^{(2)}_{Y_-} h^2_{Y_-}, \tag{A.37}$$

where

$$D_{nY_+,1} \tag{A.38}$$
$$=\frac{1}{q}e_q^T\left[(S_{nY_+}^{-1})_{11}f_{\epsilon_{Y_+}}\frac{1}{2n_+h_{Y_+}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^2}{\sigma_{\epsilon_{Y_+,i}}}+(S_{nY_+}^{-1})_{12}\sum_{k=1}^{q}f_{\epsilon_{Y_+}}(c_k)\frac{1}{2n_+h_{Y_+}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^3}{\sigma_{\epsilon_{Y_+,i}}}\right],$$
$$D_{nY_-,1} \tag{A.39}$$
$$=\frac{1}{q}e_q^T\left[(S_{nY_-}^{-1})_{11}f_{\epsilon_{Y_-}}\frac{1}{2n_-h_{Y_-}}\sum_{i=1}^{n_-}\frac{K_{-,i}x_{-,i}^2}{\sigma_{\epsilon_{Y_-,i}}}+(S_{nY_-}^{-1})_{12}\sum_{k=1}^{q}f_{\epsilon_{Y_-}}(c_k)\frac{1}{2n_-h_{Y_-}}\sum_{i=1}^{n_-}\frac{K_{-,i}x_{-,i}^3}{\sigma_{\epsilon_{Y_-,i}}}\right].$$

Applying Lemma 9 to $\widehat{\text{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}$, $\widehat{\text{Bias}}(\hat{m}_{Y_-})_{\text{fixed-n}}$ for eq. (A.35):

$$\text{II}=\frac{4}{n_+h_{Y_+}}D_{nY_+,1}^2e_2^T(S_{nY_+}^{-1}\Sigma_{nY_+}S_{nY_+}^{-1})_{22}e_2+\frac{4}{n_-h_{Y_-}}D_{nY_-,1}^2e_2^T(S_{nY_-}^{-1}\Sigma_{nY_-}S_{nY_-}^{-1})_{22}e_2.$$

For term III with *i.i.d.* errors, we have

$$\text{Cov}(\hat{\tau}_{\text{sharp}},\widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}})=\text{Cov}(\hat{m}_{Y_+},\widehat{\text{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}})+\text{Cov}(\hat{m}_{Y_-},\widehat{\text{Bias}}(\hat{m}_{Y_-})_{\text{fixed-n}}).$$

Using eq. (A.21) and eq. (A.29), we can show

$$\text{Cov}(\hat{m}_{Y_+},\widehat{\text{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}})=\frac{2D_{nY_+,1}}{qn_+h_{Y_+}}e_2^T(S_{nY_+}^{-1}\Sigma_{nY_+}S_{nY_+}^{-1})_{12,2},$$
$$\text{Cov}(\hat{m}_{Y_-},\widehat{\text{Bias}}(\hat{m}_{Y_-})_{\text{fixed-n}})=\frac{2D_{nY_-,1}}{qn_-h_{Y_-}}e_2^T(S_{nY_-}^{-1}\Sigma_{nY_-}S_{nY_-}^{-1})_{12,2}.$$

Putting everything together, the numerator in eq. (A.33) becomes

$$\begin{aligned}\hat{\tau}_{\text{sharp}}-\widehat{\text{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}}=\hat{\tau}_{\text{sharp}}-\Bigg\{&\frac{1}{q}e_q^T\Bigg[(S_{nY_+}^{-1})_{11}f_{\epsilon_{Y_+}}\frac{1}{2n_+h_{Y_+}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^2}{\sigma_{\epsilon_{Y_+,i}}}\\&+(S_{nY_+}^{-1})_{12}\sum_{k=1}^{q}f_{\epsilon_{Y_+}}(c_k)\frac{1}{2n_+h_{Y_+}}\sum_{i=1}^{n_+}\frac{K_{+,i}x_{+,i}^3}{\sigma_{\epsilon_{Y_+,i}}}\Bigg]m_{Y_+}^{(2)}h_{Y_+}^2\\&-\frac{1}{q}e_q^T\Bigg[(S_{nY_-}^{-1})_{11}f_{\epsilon_{Y_-}}\frac{1}{2n_-h_{Y_-}}\sum_{i=1}^{n_-}\frac{K_{-,i}x_{-,i}^2}{\sigma_{\epsilon_{Y_-,i}}}\\&+(S_{nY_-}^{-1})_{12}\sum_{k=1}^{q}f_{\epsilon_{Y_-}}(c_k)\frac{1}{2n_-h_{Y_-}}\sum_{i=1}^{n_-}\frac{K_{-,i}x_{-,i}^3}{\sigma_{\epsilon_{Y_-,i}}}\Bigg]m_{Y_-}^{(2)}h_{Y_-}^2\Bigg\}.\end{aligned}$$

The variance on the denominator is given by

$$
\begin{aligned}
&\mathrm{Var}(\hat{\tau}_{\text{sharp}} - \widehat{\mathrm{Bias}}(\hat{\tau}_{\text{sharp}})_{\text{fixed-n}}) \\
&= \frac{1}{n_+ h_{Y_+} q^2} e_q^T \left(S_{nY_+}^{-1} \Sigma_{nY_+} S_{nY_+}^{-1} \right) e_q + \frac{1}{n_- h_{Y_-} q^2} e_q^T \left(S_{nY_-}^{-1} \Sigma_{nY_-} S_{nY_-}^{-1} \right) e_q \\
&\quad + \frac{4}{n_+ h_{Y_+}} D_{nY_+,1}^2 e_2^T (S_{nY_+}^{-1} \Sigma_{nY_+} S_{nY_+}^{-1})_{22} e_2 + \frac{4}{n_- h_{Y_-}} D_{nY_-,1}^2 e_2^T (S_{nY_-}^{-1} \Sigma_{nY_-} S_{nY_-}^{-1})_{22} e_2 \\
&\quad - 2 \left[\frac{2 D_{nY_+,1}}{q n_+ h_{Y_+}} e_2^T (S_{nY_+}^{-1} \Sigma_{nY_+} S_{nY_+}^{-1})_{12,2} + \frac{2 D_{nY_-,1}}{q n_- h_{Y_-}} e_2^T (S_{nY_-}^{-1} \Sigma_{nY_-} S_{nY_-}^{-1})_{12,2} \right].
\end{aligned}
$$

□

Proposition 2. Under Assumptions 1 to 6, the fixed-n adjusted t-statistic for the fuzzy RD is given by

$$t^{\text{adj.}}_{\text{fuzzy, fixed-n}} = \frac{\tilde{\tau}_{\text{fuzzy}} - \widehat{\mathrm{Bias}}(\tilde{\tau}_{\text{fuzzy}})_{\text{fixed-n}}}{\sqrt{\mathrm{Var}(\tilde{\tau}_{\text{fuzzy}} - \widehat{\mathrm{Bias}}(\tilde{\tau}_{\text{fuzzy}})_{\text{fixed-n}})}}. \tag{A.40}$$

Proof of Proposition 2. The numerator of eq. (A.40) can be obtained by applying the fixed-n bias result in Lemma 5 to eq. (34). To compute the denominator, we again start with eq. (35). For the fixed-n result in eq. (35), expressions for $\mathrm{Var}(\hat{m}_{Y_+})_{\text{fixed-n}}$ and $\mathrm{Var}(\hat{m}_{T_+})_{\text{fixed-n}}$ are given by Lemma 5, $\mathrm{Var}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}})$, $\mathrm{Var}(\widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}})$, $\mathrm{Cov}(\hat{m}_{Y_+}, \widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}})$ and $\mathrm{Cov}(\hat{m}_{T_+}, \widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}})$ are derived in the proof of Proposition 1. We list the seven terms below and omit the small-order terms.

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_{Y_+})_{\text{fixed-n}} &= \frac{1}{n_+ h_{Y_+} q^2} e_q^T \left(S_{nY_+}^{-1} \Sigma_{nY_+} S_{nY_+}^{-1} \right) e_q, \\
\mathrm{Var}(\hat{m}_{T_+})_{\text{fixed-n}} &= \frac{1}{n_+ h_{T_+} q^2} e_q^T \left(S_{nT_+}^{-1} \Sigma_{nT_+} S_{nT_+}^{-1} \right) e_q, \\
\mathrm{Var}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}) &= \frac{4}{n_+ h_{Y_+}} D_{nY_+,1}^2 e_2^T (S_{nY_+}^{-1} \Sigma_{nY_+} S_{nY_+}^{-1})_{22} e_2, \\
\mathrm{Var}(\widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}}) &= \frac{4}{n_+ h_{T_+}} D_{nT_+,1}^2 e_2^T (S_{nT_+}^{-1} \Sigma_{nT_+} S_{nT_+}^{-1})_{22} e_2, \\
\mathrm{Cov}(\hat{m}_{Y_+}, \hat{m}_{T_+})_{\text{fixed-n}} &= \frac{1}{q^2 n_+ \sqrt{h_{Y_+} h_{T_+}}} e_q^T \left[(S_{nY_+}^{-1})_{11} \ (S_{nY_+}^{-1})_{12} \right] \Sigma_{nYT_+} \left[(S_{nT_+}^{-1})_{11} \ (S_{nT_+}^{-1})_{12} \right]^T e_q, \\
\mathrm{Cov}(\hat{m}_{Y_+}, \widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}) &= \frac{2 D_{nY_+,1}}{q n_+ h_{Y_+}} e_2^T (S_{nY_+}^{-1} \Sigma_{nY_+} S_{nY_+}^{-1})_{12,2}, \\
\mathrm{Cov}(\hat{m}_{T_+}, \widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}}) &= \frac{2 D_{nT_+,1}}{q n_+ h_{T_+}} e_2^T (S_{nT_+}^{-1} \Sigma_{nT_+} S_{nT_+}^{-1})_{12,2}.
\end{aligned}
$$

We only need to compute the remaining three terms in eq. (35). Consider the term $\mathrm{Cov}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}, \widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}})$. Using the result in eq. (A.36) and a similar result for $\widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}}$, together with the result in eq. (A.21) and a similar result for $\hat{m}^{(2)}_{T_+}$, it can be shown that

$$\mathrm{Cov}(\widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}, \widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}}) = \frac{4D_{nY_+,1}D_{nT_+,1}}{n_+\sqrt{h_{Y_+}h_{T_+}}} e_2^T (S^{-1}_{nY_+}\Sigma_{nYT_+}S^{-1}_{nT_+})_{22}e_2,$$

similar to the proof in Theorem 5.

Again, similar to the proof in Theorem 5, using eq. (A.21), eq. (A.36) and eq. (A.38), we have

$$\mathrm{Cov}(\hat{m}_{Y_+}, \widehat{\mathrm{Bias}}(\hat{m}_{T_+})_{\text{fixed-n}}) = \frac{2D_{nT_+,1}}{qn_+\sqrt{h_{Y_+}h_{T_+}}} e_2^T (S^{-1}_{nY_+}\Sigma_{nYT_+}S^{-1}_{nT_+})_{12,2},$$
$$\mathrm{Cov}(\hat{m}_{T_+}, \widehat{\mathrm{Bias}}(\hat{m}_{Y_+})_{\text{fixed-n}}) = \frac{2D_{nY_+,1}}{qn_+\sqrt{h_{Y_+}h_{T_+}}} e_2^T (S^{-1}_{nT_+}\Sigma_{nTY_+}S^{-1}_{nY_+})_{12,2}.$$

Substitute the above ten results into eq. (35) to obtain a fixed-n version of $\mathrm{Var}((\hat{m}_{Y_+} - \tau_0\hat{m}_{T_+}) - (\widehat{\mathrm{Bias}}(\hat{m}_{Y_+}) - \tau_0\widehat{\mathrm{Bias}}(\hat{m}_{T_+})))$. The fixed-$n$ result for $\mathrm{Var}((\hat{m}_{Y_-} - \tau_0\hat{m}_{T_-}) - (\widehat{\mathrm{Bias}}(\hat{m}_{Y_-}) - \tau_0\widehat{\mathrm{Bias}}(\hat{m}_{T_-})))$ can be obtained in a similar way. Adding up the two results gives the variance on the denominator of eq. (A.40). □

S.4 Additional figures and tables

S.4.1 Figure: the bias of bias-corrected estimators

To accompany Figure 1 in the main text, this subsection contains additional Figure 3 that compares the finite sample performance of LCQR and LLR in estimating the treatment effect.

S.4.2 Figure: LCQR and LLR at interior and boundary points

To motivate the use of LCQR, consider the nonlinear model in Ruppert *et al.* (1995), $Y = \sin(5\pi X) + 0.5\epsilon$, where ϵ follows a mixture normal distribution, $0.95N(0,1) + 0.05N(0,10^2)$, and X follows a uniform distribution on $[0,1]$. It is clear from Figure 4 that LCQR exhibits less "flapping" for both interior and boundary points. The relative stable behavior of LCQR on the boundary when data move away from normality is of particular importance to the estimation and inference in RD.

S.4.3 Table: coverage probability with the rule-of-thumb bandwidth

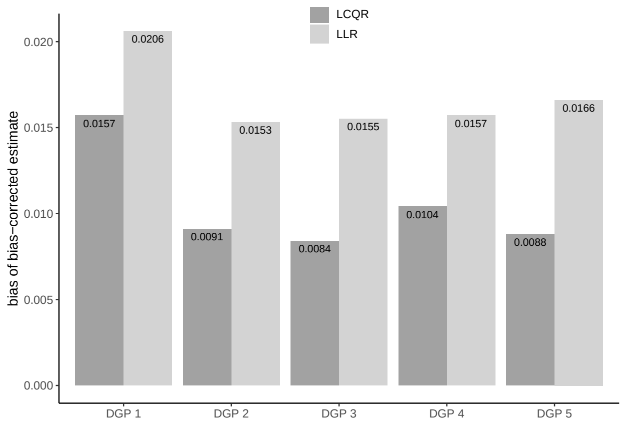
This subsection presents Table 5 that is similar to Table 4 except that $\hat{\tau}_{\mathrm{2bw}}^{\mathrm{cqr}}$ and $\hat{\tau}_{\mathrm{2bw}}^{\mathrm{cqr,bc}}$ use the rule-of-thumb bandwidth described by Equation (4.3) in Fan and Gijbels (1996). Table 5 indicates that the proposed LCQR method has some robustness to the choice of bandwidth.

S.4.4 Table: coverage probability of fixed-n LCQR with small sample

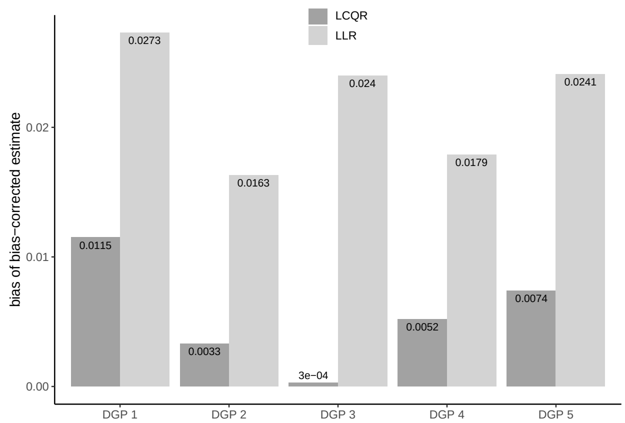
In this subsection, we decrease the sample size from $n = 500$ to 300 in the simulation study. We show that the fixed-n approach indeed can improve the coverage when the sample size is relatively small, as reported in the last row of each panel of Table 6.

S.4.5 Table: LCQR for sharp kink RD

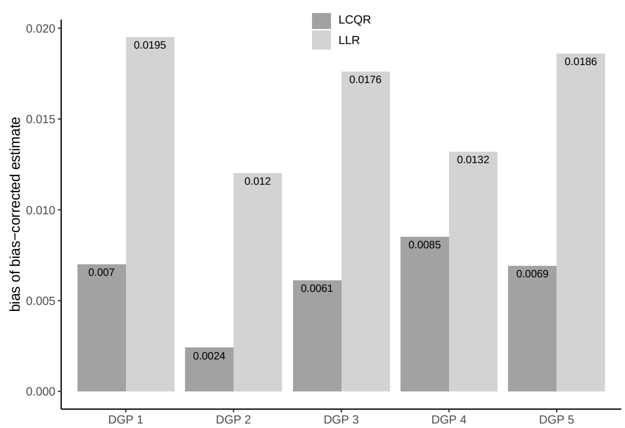
We consider the LM model used for the simulation study, but now focus on the difference in derivatives around the cutoff: $18.49 - 2.3 = 16.19$, as in a sharp kink RD design. Table 7 shows that LCQR could outperform the local polynomial regression for estimating derivatives when data are non-normal; see e.g. DGP 2 - 5.



(a) Lee with heteroskedatic errors

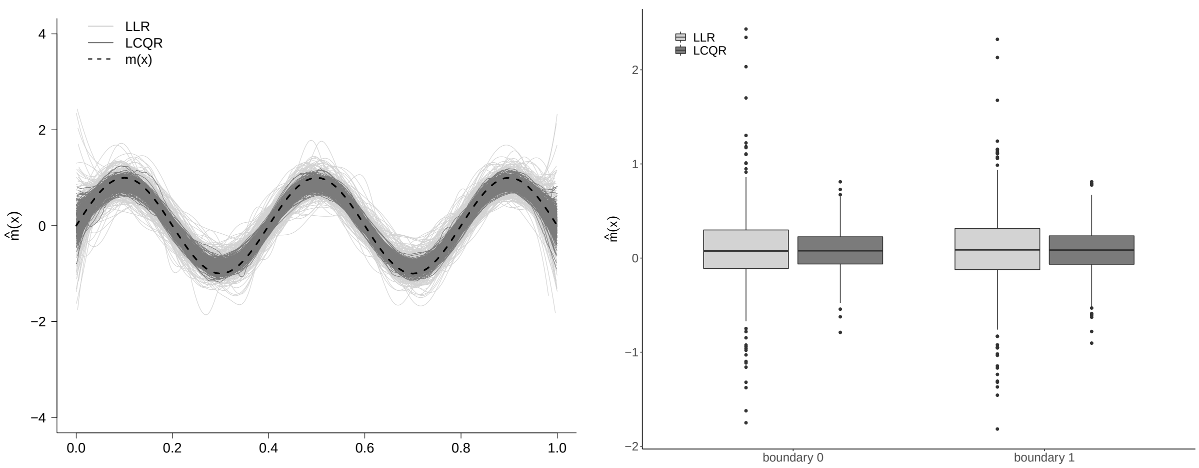


(b) LM with homoskedastic errors



(c) LM with heteroskedatic errors

Figure 3: Absolute value of average bias of the bias-corrected estimators, $\hat{\tau}_{1\text{bw}}^{\text{cqr,bc}}$ and $\hat{\tau}_{1\text{bw}}^{\text{robust,bc}}$ for the Lee and LM models. $\hat{\tau}_{1\text{bw}}^{\text{cqr,bc}}$ is the bias-corrected LCQR estimator. $\hat{\tau}_{1\text{bw}}^{\text{robust,bc}}$ is the bias-corrected LLR estimator. The result is based on 5000 replications and the true treatment effect is 0.04 for Lee and -3.45 for LM. The DGPs are described in the paper.



(a) LLR versus LCQR (b) Box plots of the estimates on the boundaries

Figure 4: Estimates of LLR and LCQR with a sample size of 400 and 400 replications. Both methods use the same direct plug-in bandwidth in Ruppert *et al.* (1995). $m(X) = \sin(5\pi X)$.

Table 5: Coverage probability of 95% confidence intervals in Lee and LM models using the rule-of-thumb bandwidth for LCQR

	A. Lee with homoskedastic errors					B. Lee with heteroskedatic errors				
Estimator	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
$\hat{\tau}_{2bw}^{cqr}$	0.915	0.917	0.909	0.916	0.917	0.901	0.896	0.887	0.897	0.895
$\hat{\tau}_{2bw}^{cqr,bc}$	0.976	0.963	0.968	0.976	0.965	0.969	0.956	0.958	0.965	0.951

	C. LM with homoskedastic errors					D. LM with heteroskedatic errors				
Estimator	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
$\hat{\tau}_{2bw}^{cqr}$	0.888	0.891	0.880	0.892	0.859	0.876	0.876	0.870	0.878	0.845
$\hat{\tau}_{2bw}^{cqr,bc}$	0.967	0.956	0.962	0.968	0.960	0.958	0.946	0.952	0.952	0.943

Notes: The reported numbers are the simulated coverage probabilities of the 95% confidence intervals associated with different estimators. The results are based on 5000 replications with a sample size $n = 500$. The s.e. and adjusted s.e. for the LCQR estimator are obtained based on the asymptotic expressions from Theorem 1 and Theorem 3. Estimators with superscript `bc` are both bias-corrected and s.e.-adjusted. The DGPs are described in the paper.

Table 6: Coverage probability of 95% confidence intervals in Lee and LM models, $n = 300$

	A. Lee with homoskedastic errors					B. Lee with heteroskedastic errors				
Estimator	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
$\hat{\tau}_{1\text{bw}}^{\text{cqr}}$	0.906	0.892	0.900	0.903	0.886	0.882	0.872	0.877	0.873	0.871
$\hat{\tau}_{2\text{bw}}^{\text{cqr}}$	0.895	0.887	0.890	0.890	0.875	0.869	0.859	0.868	0.859	0.857
$\hat{\tau}_{1\text{bw}}^{\text{llr}}$	0.926	0.926	0.929	0.917	0.943	0.923	0.922	0.928	0.918	0.945
$\hat{\tau}_{1\text{bw}}^{\text{cqr,bc}}$	0.958	0.941	0.940	0.948	0.923	0.947	0.926	0.928	0.933	0.914
$\hat{\tau}_{2\text{bw}}^{\text{cqr,bc}}$	0.954	0.937	0.937	0.945	0.918	0.939	0.924	0.929	0.929	0.906
$\hat{\tau}_{1\text{bw}}^{\text{robust,bc}}$	0.927	0.928	0.930	0.921	0.946	0.923	0.925	0.928	0.920	0.946
$\hat{\tau}_{1\text{bw,fixed-n}}^{\text{cqr,bc}}$	0.977	0.959	0.960	0.969	0.950	0.958	0.939	0.944	0.945	0.936

	C. LM with homoskedastic errors					D. LM with heteroskedastic errors				
Estimator	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
$\hat{\tau}_{1\text{bw}}^{\text{cqr}}$	0.791	0.820	0.826	0.803	0.832	0.613	0.668	0.700	0.640	0.728
$\hat{\tau}_{2\text{bw}}^{\text{cqr}}$	0.805	0.839	0.841	0.815	0.837	0.622	0.689	0.706	0.657	0.732
$\hat{\tau}_{1\text{bw}}^{\text{llr}}$	0.907	0.917	0.921	0.905	0.936	0.899	0.905	0.914	0.896	0.926
$\hat{\tau}_{1\text{bw}}^{\text{cqr,bc}}$	0.959	0.941	0.939	0.946	0.923	0.941	0.928	0.927	0.931	0.910
$\hat{\tau}_{2\text{bw}}^{\text{cqr,bc}}$	0.953	0.936	0.936	0.944	0.915	0.935	0.924	0.924	0.928	0.908
$\hat{\tau}_{1\text{bw}}^{\text{robust, bc}}$	0.926	0.930	0.934	0.923	0.946	0.927	0.930	0.933	0.923	0.946
$\hat{\tau}_{1\text{bw,fixed-n}}^{\text{cqr,bc}}$	0.975	0.956	0.959	0.966	0.951	0.956	0.939	0.942	0.946	0.938

Notes: The reported numbers are the simulated coverage probabilities of the 95% confidence intervals associated with different estimators. The results are based on 5000 replications with a sample size $n = 300$. The s.e. and adjusted s.e. for the LCQR estimator are obtained based on the asymptotic expressions from Theorem 1 and Theorem 3, except for $\hat{\tau}_{1\text{bw,fixed-n}}^{\text{cqr,bc}}$ where fixed-n approximations are used. Estimators with superscript `bc` are both bias-corrected and s.e.-adjusted. The result of $\hat{\tau}_{1\text{bw}}^{\text{robust,bc}}$ is based on the CE-optimal bandwidth. The DGPs are described in the paper.

Table 7: LCQR for sharp kink RD

	LM with $\tau_{sharp\ kink} = 16.19$									
	Homoskedastic errors					Heteroskedastic errors				
Estimator	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5	DGP 1	DGP 2	DGP 3	DGP 4	DGP 5
LCQR	15.91	16.14	15.89	16.12	16.13	15.93	16.09	15.91	16.07	16.08
(s.e.)	(10.94)	(12.90)	(13.83)	(11.47)	(11.87)	(6.27)	(7.35)	(7.88)	(6.57)	(6.79)
LPR	15.93	16.13	15.69	16.06	15.84	15.96	16.07	15.82	16.03	15.90
(s.e.)	(10.47)	(14.77)	(18.11)	(12.25)	(25.42)	(6.10)	(8.61)	(10.53)	(7.13)	(14.80)

Notes: The reported numbers are the simulated averages and standard errors (in brackets) of the associated estimators. The results are based on 5000 replications with a sample size $n = 500$. The DGPs are as described in the paper for the LM model, yet the focus here is on the difference in first derivatives. For both LCQR and LPR (local polynomial regression), we consider the 3rd-order polynomial with the fixed bandwidth $= 0.3$ and the triangular kernel. The R code to replicate this table can be downloaded from https://xhuang.netlify.app/post/r-code-to-replicate-rd-tables/.

S.4.6 Table: simulation results for sharp RD with covariates

In this subsection we use the same DGP as for Table SA-1 in Calonico *et al.* (2019). We briefly describe the DGP below. Let Z_i be the covariate. Consider a sample size of $n = 1000$ and 5000 replications. For each $i = 1, \cdots, n$, we have

$$Y_i = m_{y,j}(X_i, Z_i) + \epsilon_{y,i}, \quad Z_i = m_z(X_i) + \epsilon_{z,i}, \quad X_i \sim 2 \times \text{Beta}(2,4) - 1$$

with

$$\begin{pmatrix} \epsilon_{y,i} \\ \epsilon_{z,i} \end{pmatrix} \sim N(0, \Sigma_j), \quad \Sigma_j = \begin{pmatrix} \sigma_y^2 & \rho_j \sigma_y \sigma_z \\ \rho_j \sigma_y \sigma_z & \sigma_z^2 \end{pmatrix}$$

and $j = 1, 2, 3, 4$, corresponding to the following four models.

- **Model 1** has no covariate and is the same as eq. (45)

$$m_{y,1}(X_i, Z_i) = \begin{cases} 0.48 + 1.27X_i + 7.18X_i^2 + 20.21X_i^3 + 21.54X_i^4 + 7.33X_i^5 & \text{if } X_i < 0, \\ 0.52 + 0.84X_i - 3.00X_i^2 + 7.99X_i^3 - 9.01X_i^4 + 3.56X_i^5 & \text{if } X_i \geq 0, \end{cases}$$

  and let $\sigma_y = 0.1295$ and $\sigma_z = 0.1353$.

- **Model 2** adds one covariate, and let $\rho = 0.2692$,

$$m_{y,2}(X_i, Z_i) = \begin{cases} 0.36 + 0.96X_i + 5.47X_i^2 + 15.28X_i^3 + 15.87X_i^4 + 5.14X_i^5 + 0.22Z_i & \text{if } X_i < 0, \\ 0.38 + 0.62X_i - 2.84X_i^2 + 8.42X_i^3 - 10.24X_i^4 + 4.31X_i^5 + 0.28Z_i & \text{if } X_i \geq 0, \end{cases}$$

$$m_z(X_i) = \begin{cases} 0.49 + 1.06X_i + 5.74X_i^2 + 17.14X_i^3 + 19.75X_i^4 + 7.47X_i^5 & \text{if } X_i < 0, \\ 0.49 + 0.61X_i + 0.23X_i^2 - 3.46X_i^3 + 6.43X_i^4 - 3.48X_i^5 & \text{if } X_i \geq 0. \end{cases}$$

- **Model 3** is the same as Model 2 except for $\rho = 0$.
- **Model 4** is the same as Model 2 except for $\rho = 2 \times 0.2692$.

The true value for τ is 0.04 in Model 1 and approximately 0.05 in Models 2-4. Table 8 reports $\sqrt{\text{MSE}}$, bias as a percentage of τ and empirical coverage (EC) for the confidence intervals based on $\hat{\tau}$ and $\tilde{\tau}$. The EC for $\hat{\tau}$ is obtained using bias-corrected, s.e.-adjusted t-statistic in Theorem 3; the EC for $\tilde{\tau}$ is obtained using the same t-statistic for $\hat{\tau}$ but replacing $\hat{\tau}$ with $\tilde{\tau}$ on the numerator. See also Section 4.3 for a discussion of this *ad hoc* method for $\tilde{\tau}$. The last column in Table 8 gives reasonably good coverage probabilities, suggesting the *ad hoc* approach described in Section 4.3 works well under the considered DGP. However, more simulation studies are needed to investigate its performance.

Table 8: Simulation results using a single bandwidth

	$\hat{\tau}$ in eq. (42)			$\tilde{\tau}$ in eq. (43)		
	$\sqrt{\text{MSE}}$	Bias (%)	EC	$\sqrt{\text{MSE}}$	Bias (%)	EC
Model 1	0.046	0.369	0.953	0.046	0.368	0.952
Model 2	0.049	0.256	0.942	0.043	0.178	0.968
Model 3	0.047	0.275	0.938	0.046	0.231	0.944
Model 4	0.053	0.275	0.949	0.038	0.139	0.980

Notes: We use a single bandwidth $h = 0.15$ for estimation, bias-correction, s.e.-adjustment in all four models. This number is chosen to mimic the bandwidth used in the simulation section in Calonico *et al.* (2019). All numbers in the table are based on 5000 replications. The R code to replicate this table can be downloaded from https://xhuang.netlify.app/post/r-code-to-replicate-rd-tables/.